Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

From CVaR to Uncertainty Set: Implications in Joint Chance-Constrained Optimization

Wenqing Chen, Melvyn Sim, Jie Sun, Chung-Piaw Teo,

Operations Research
Vol. 58, No. 2, March–April 2010, pp. 470–485
ISSN 0030-364X | EISSN 1526-5463 | 10 | 5802 | 0470
DOI 10.1287/opre.1090.0712

From CVaR to Uncertainty Set: Implications in Joint Chance-Constrained Optimization

Wenqing Chen
NUS Business School, National University of Singapore, Singapore, chenwenqing@gmail.com

Melvyn Sim, Jie Sun
NUS Business School and NUS Risk Management Institute, National University of Singapore, Singapore, {dscsimm@nus.edu.sg, jsun@nus.edu.sg}

Chung-Piaw Teo
NUS Business School, National University of Singapore, Singapore, bizteocp@nus.edu.sg

We review and develop different tractable approximations to individual chance-constrained problems in robust optimization on a variety of uncertainty sets and show their interesting connections with bounds on the conditional-value-at-risk (CVaR) measure. We extend the idea to joint chance-constrained problems and provide a new formulation that improves upon the standard approach. Our approach builds on a classical worst-case bound for order statistics problems and is applicable even if the constraints are correlated. We provide an application of the model on a network resource allocation problem with uncertain demand.

Subject classifications: decision analysis: risk; probability: application; programming: stochastic, nonlinear.
Area of review: Optimization.
History: Received December 2007; revisions received August 2008, January 2009; accepted January 2009. Published online in *Articles in Advance* October 7, 2009.

1. Introduction

Data uncertainties prevail in many real-world linear optimization models. If ignored, the so-called "optimal solution" obtained by solving a model using the "nominal data" or point estimates can become infeasible in the model when the true data differs from the nominal one. To overcome such infeasibility, Soyster (1973) introduced a worst-case model that ensures feasibility of its solution for all possible realization of the uncertain data. Let $\mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \geqslant \mathbf{b}(\tilde{\mathbf{z}})$ be an $m \times n$ linear constraint system that depends on a random vector $\tilde{\mathbf{z}}$. Soyster proposed the following model (Soyster 1973):

$$\begin{aligned} \min\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \geqslant \mathbf{b}(\tilde{\mathbf{z}}) \quad \forall\, \tilde{\mathbf{z}} \in \mathscr{W}, \end{aligned} \tag{1}$$

where

$$\mathscr{W} = \{\mathbf{z}\colon\ -\underline{\mathbf{z}} \leqslant \mathbf{z} \leqslant \bar{\mathbf{z}}\} \quad \text{for some } \underline{\mathbf{z}}, \bar{\mathbf{z}} > \mathbf{0}$$

is the support of the primitive uncertainty vector $\tilde{\mathbf{z}}$. Soyster (1973) showed that the model can be represented as a polynomially sized linear optimization model. However, this model can be extremely conservative in addressing models where the violation of constraints may be tolerated as a trade-off for better attainment in objective.

Perhaps the most natural way of safeguarding a constraint is to restrict its violation probability. Such a constraint is known as a probabilistic or a chance constraint, which was introduced by Charnes et al. (1958). A chance-constrained model is defined as follows:

$$\begin{aligned} Z_\epsilon = \min\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \mathrm{P}(\mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \geqslant \mathbf{b}(\tilde{\mathbf{z}})) \geqslant 1 - \epsilon, \\ & \mathbf{x} \in X, \end{aligned} \tag{2}$$

where $x \in X$ represents a set of additional deterministic constraints. Problem (2) requires all the m linear constraints $\mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \geqslant \mathbf{b}(\tilde{\mathbf{z}})$ to be jointly feasible with probability at least $1 - \epsilon$, where $\epsilon \in (0, 1)$ is a desired safety factor.

Chance-constrained problems can be classified as an individual chance-constrained problem when $m = 1$, and a joint chance-constrained problem when $m > 1$. One fundamental issue on the chance-constrained problem is to determine the distributional condition under which the problem is convex. It is well known that under multivariate normal distribution, an individual chance-constrained problem is second-order cone representable. In other words, the optimization model becomes a second-order cone optimization problem (SOCP), which is computationally tractable, both in theory and practice (see, for example, Alizadeh and Goldfarb 2003). In addition, Lagoa (1999) proved that the individual chance-constrained problem is convex under the condition that the distribution of the random parameters is uniform over a convex symmetric set. More generally, Calafiore and El Ghaoui (2006) showed that the individual chance constraint can be converted to second-order cone constraints when the random parameters are under radial distributions. However, for general distributions, chance-constrained problems are computationally intractable. For example, Nemirovski and Shapiro (2006) noted that evaluating the distribution of a weighted sum of uniformly distributed independent random variables is already NP-hard.

Needless to say, joint chance-constrained problems are clearly harder than individual chance-constraint problems. For example, with only right-hand side disturbances, we can transform an individual chance-constrained problem to an equivalent linearly constrained problem. In contrast, with only right-hand side disturbances, a joint chance-constrained problem is known to be convex only when the distributions are log-concave (cf. Prékopa 1995).

It is possible to incorporate joint probabilistic constraints using a discrete representation obtained by Monte Carlo sampling (for example, in Ruszczynski 2002). Indeed, sampling approximation of the chance-constrained problem has been studied theoretically in Calafiore and Campi (2005, 2006), Ergodan and Iyengar (2006), and Luedtke and Ahmed (2008). These methods require roughly about $O(n/\epsilon)$ constraint duplications to yield a highly reliable solution with respect to its feasibility (see Calafiore and Campi 2006) as well as optimality (see Luedtke and Ahmed 2008). However, it may be computationally prohibitive to solve large problems or to solve problems under high feasibility requirement. The effectiveness of sampling approximation has also been challenged in the computation studies of Nemirovski and Shapiro (2006) and Chen et al. (2007).

The intractability of a chance-constrained problem using exact probability distributions has spurred recent interests in robust optimization in which data uncertainties are described using uncertainty sets. Moreover, robust optimization often requires only a mild assumption on probability distributions such as known supports, covariances, and/or other forms of deviation measures, notably the directional deviations derived from moment-generating functions proposed by Chen et al. (2007). For some practitioners, this could be viewed as an advantage over having to obtain the entire joint probability distributions of the uncertain data. One of the goals of robust optimization is to provide a tractable approach for obtaining a solution that remains feasible in the chance-constrained model (2) for all distributions that conform to the mild distributional assumption. Hence, such solutions are viewed as "safe" approximations to the chance-constrained problem.

Robust optimization has been fairly successful in constructing safe approximation of individual chance-constrained problems. Given an uncertainty set $\mathcal{U}$, the robust counterpart of an individual linear constraint with affinely dependent primitive uncertainty vector $\tilde{\mathbf{z}}$ is defined as

$$\mathbf{a}(\tilde{\mathbf{z}})'\mathbf{x} \geqslant b(\tilde{\mathbf{z}}) \quad \forall \tilde{\mathbf{z}} \in \mathcal{U}.$$

Clearly, Soyster's model (1) is a special case of the robust counterpart in which the uncertainty set $\mathcal{U}$ is chosen to be the support set $\mathcal{W}$. For computational tractability, the chosen uncertainty set $\mathcal{U}$ is usually in the form of tractable convex representable sets such as these with second-order cone and linear constraints. Various symmetric uncertainty sets have been proposed by Ben-Tal and Nemirovski (1998, 1999), El Ghaoui et al. (1998), and Bertsimas and Sim (2004). Calafiore and El Ghaoui (2006) also provided explicit results for enforcement of the individual chance constraint based on moments, bounds, or symmetry information. More recently, Chen et al. (2007) proposed an asymmetrical uncertainty set that generalizes the symmetric ones. All these models are computationally attractive in the form of SOCPs or even in the form of linear programs (LPs). In the recent work of Nemirovski and Shapiro (2006), the moment-generating functions are incorporated for providing safe and tractable approximations of an individual chance-constrained problem. Despite the improved approximation, the approximation is not readily second-order cone representable, and hence computationally more expensive. Other forms of deterministic approximation of an individual chance-constrained problem includes using Chebyshev's inequality, Bernstein's inequality, or Hoeffding's inequality to bound the probability of violating individual constraints. See, for example, Pintér (1989).

Although robust optimization has been pretty successful in approximating individual chance-constrained problems, it is rather unsatisfactory in approximating joint chance-constrained problems. The "standard method" for approximating a joint constrained problem is to decompose a joint chance-constrained problem into a problem with m individual chance constraints. Clearly, by Bonferroni's inequality, a sufficient condition for ensuring feasibility in the joint chance-constrained problem is to ensure that the total sum of violation probabilities of the individual chance constraints is less than ϵ. The natural choice proposed in Chen et al. (2007) and Nemirovski and Shapiro (2006) is to divide the violation probability equally among the m individual chance constraints. To the best of our knowledge, prior to this work, we do not know of any systematic approach for selecting better allocation of the safety factors among the individual chance constraints. Unfortunately, even when the individual chance constraints are independent, the Bonferroni inequality is only an approximation at best.[1] In the events when the individual chance constraints are correlated, the approximation obtained using Bonferroni's inequality could be even more conservative.

The above motivates our research to achieve better approximations of joint chance-constrained problems. We build instead on a classical result on order statistics (cf. Meilijson and Nadas 1979) to bound the probability of violation for the joint chance constraint. We show that by choosing the right multipliers in conjunction with this classical inequality, we can derive an improved approximation to the above method for the joint chance constraint-problem.

Our specific contributions in this paper include the following.

1. We review the different tractable approximations to individual chance-constraint problems used in robust optimization and, by using the bounds of $\mathrm{E}((\cdot)^+)$ developed in Chen and Sim (2009), show their interesting connections with bounds on the conditional-value-at-risk (CVaR) measure.
2. We propose a new formulation for approximating joint chance-constrained problems that improves upon the standard approach using Bonferroni's inequality.
3. We provide an application of the model on a network resource allocation problem with uncertain demand and study the performance of the new chance-constrained formulation over the standard approach.

The rest of this paper is organized as follows. In §2, we focus on robust optimization approximation of individual chance-constrained problems. Our work is closely related to Chen and Sim (2009), which is discussed in §2.1. In particular, we adopt the same model of data uncertainty (in Assumption U) and the bounding functions of $\mathrm{E}((\cdot)^+)$. In §3, we propose a new approximation of the joint chance-constrained problem. In §4, we analyze the efficacy of this approximation through a computational study of emergency supply allocation network. Finally, we conclude this paper in §5.

Notations. We denote random variables with the tilde sign, such as $\tilde{x}$. Boldface lower-case letters represent vectors such as $\mathbf{x}$, and boldface upper-case letters represent matrices such as $\mathbf{A}$. In addition, we denote $x^+ = \max\{x, 0\}$ and use $\mathrm{E}(\cdot)$ to stand for the expectation.

2. Individual Chance-Constrained Problems

In this section, we will establish the relation between bounds on the CVaR measure popularized by Rockafellar and Uryasev (2000) and the different tractable approximations of individual chance-constrained problems used in robust optimization. For simplicity, we consider a linear individual chance constraint as follows:

$$\mathrm{P}(y(\tilde{\mathbf{z}}) \leqslant 0) \geqslant 1 - \epsilon, \tag{3}$$

where $y(\tilde{\mathbf{z}})$ are affinely dependent of $\tilde{\mathbf{z}}$,

$$y(\tilde{\mathbf{z}}) = y_0 + \sum_{k=1}^{N} y_k \tilde{z}_k,$$

and $(y_0, y_1, \ldots, y_N)$ are the decision variables. To illustrate the generality, we can represent the following chance constraint

$$\mathrm{P}(\mathbf{a}(\tilde{\mathbf{z}})'\mathbf{x} \geqslant b(\tilde{\mathbf{z}})) \geqslant 1 - \epsilon,$$

where

$$\mathbf{a}(\tilde{\mathbf{z}}) = \mathbf{a}^0 + \sum_{k=1}^{N} \mathbf{a}^k \tilde{z}_k,$$

$$b(\tilde{\mathbf{z}}) = b_0 + \sum_{k=1}^{N} b_k \tilde{z}_k,$$

by enforcing the following affine relations

$$y_k = -(\mathbf{a}^k)'\mathbf{x} + b_k \quad \forall\, k = 0, \ldots, N.$$

The chance constraint (3) is not necessarily convex in its decision variables, $(y_0, y_1, \ldots, y_N)$. A step toward tractability is by convexifying the individual chance constraint (3) using the CVaR measure, $\rho_{1-\epsilon}(\tilde{v})$, which is a functional on a random variable $\tilde{v}$ defined as follows:

$$\rho_{1-\epsilon}(\tilde{v}) \triangleq \min_{\beta} \left\{ \beta + \frac{1}{\epsilon} \mathrm{E}((\tilde{v} - \beta)^+) \right\}. \tag{4}$$

The CVaR measure is a special class of optimized certainty equivalent (OCE) risk measure introduced by Ben-Tal and Teboulle (1986) and is popularized by Rockafellar and Uryasev (2000) as a tractable alternative for solving value-at-risk problems in financial applications. Recent works of Bertsimas and Brown (2009) and Natarajan et al. (2009) have uncovered the relation between financial risk measures and uncertainty sets in robust optimization. The CVaR constraint,

$$\rho_{1-\epsilon}(y(\tilde{\mathbf{z}})) \leqslant 0, \tag{5}$$

is a convex approximation of an individual chance constraint. Indeed, if the random variable $y(\tilde{\mathbf{z}})$ satisfies inequality (5), then there exists $\beta \leqslant 0$ such that

$$\beta + \frac{1}{\epsilon} \mathrm{E}((y(\tilde{\mathbf{z}}) - \beta)^+) \leqslant 0.$$

Note that if $\beta = 0$, then the inequality necessarily implies $\mathrm{E}((y(\tilde{\mathbf{z}}))^+) \leqslant 0$ and hence,

$$\mathrm{P}(y(\tilde{\mathbf{z}}) \leqslant 0) = 1 \geqslant 1 - \epsilon.$$

On the other hand, if $\beta < 0$, then by Markov inequality, we have

$$\begin{aligned} \mathrm{P}(y(\tilde{\mathbf{z}}) > 0) &= \mathrm{P}(y(\tilde{\mathbf{z}}) - \beta > -\beta) \\ &\leqslant \mathrm{P}((y(\tilde{\mathbf{z}}) - \beta)^+ > -\beta) \\ &\leqslant \mathrm{E}((y(\tilde{\mathbf{z}}) - \beta)^+)/(-\beta) \\ &\leqslant \epsilon. \end{aligned}$$

Therefore, the random variable $y(\tilde{\mathbf{z}})$ will satisfy (5). It is also well known (e.g., Föllmer and Schied 2004, Nemirovski and Shapiro 2006) that CVaR is the tightest convex approximation to the individual chance constraint (3).

Despite its convexity, however, it is generally difficult to evaluate the CVaR measure because the expectation $\mathrm{E}((\cdot)^+)$ involves multidimensional integration. Such evaluation is computationally prohibitive above the fourth dimension. Although it is possible to approximate CVaR using sampling average approximation, the solution obtained may not be a safe approximation of the chance-constrained problem (3). Furthermore, sampling average approximation of the CVaR measure relies on full knowledge of the underlying distributions, $\tilde{\mathbf{z}}$, which may become a practical concern due to the limited availability of independent stationary historical data.

2.1. Bounding $\mathrm{E}((\cdot)^+)$

Providing bounds on $\mathrm{E}((\cdot)^+)$ is pivotal in developing tractable approximations to individual and joint chance-constrained problems. We show next that different ways of bounding $\mathrm{E}((\cdot)^+)$ using mild distributional information of $\tilde{\mathbf{z}}$, such as supports, covariances, and deviation measures. The results in bounding $\mathrm{E}((\cdot)^+)$ have also been presented in Chen and Sim (2009). For ease of reference, we list some of the known bounds on $\mathrm{E}((\cdot)^+)$.

The primitive uncertainties $\tilde{\mathbf{z}}$ may be partially characterized using the forward and backward deviations (together, they are called directional deviations), which were recently introduced by Chen et al. (2007).

Definition 2.1. Given a random variable $\tilde{z}$ with zero mean, the forward deviation is defined as

$$\sigma_f(\tilde{z}) \triangleq \sup_{\theta>0}\left\{\sqrt{2\ln(\mathrm{E}(\exp(\theta\tilde{z})))/\theta^2}\right\} \tag{6}$$

and backward deviation is defined as

$$\sigma_b(\tilde{z}) \triangleq \sup_{\theta>0}\left\{\sqrt{2\ln(\mathrm{E}(\exp(-\theta\tilde{z})))/\theta^2}\right\}. \tag{7}$$

The directional deviations are derived from the moment-generating functions of $\tilde{z}$ and may not be finite. Nevertheless, for a random variable with finite support, the respective deviations can be bounded as follows.

Theorem 2.2 (Chen et al. 2007). *If $\tilde{z}$ has zero mean and is distributed in $[-\underline{z}, \bar{z}]$, $\underline{z}, \bar{z} > 0$, then*

$$\sigma_f(\tilde{z}) \leqslant \bar{\sigma}_f(\tilde{z}) = \frac{\underline{z}+\bar{z}}{2}\sqrt{g\left(\frac{\underline{z}-\bar{z}}{\underline{z}+\bar{z}}\right)}$$

and

$$\sigma_b(\tilde{z}) \leqslant \bar{\sigma}_b(\tilde{z}) = \frac{\underline{z}+\bar{z}}{2}\sqrt{g\left(\frac{\bar{z}-\underline{z}}{\underline{z}+\bar{z}}\right)},$$

where

$$g(\mu) = 2\max_{s>0}\left\{\frac{\phi_\mu(s)-\mu s}{s^2}\right\}$$

and

$$\phi_\mu(s) = \ln\left(\frac{e^s+e^{-s}}{2}+\frac{e^s-e^{-s}}{2}\mu\right).$$

Moreover, the bounds are tight in the sense that there exists a probability distribution on $\tilde{z}$ such that $\sigma_f(\tilde{z}) = \bar{\sigma}_f(\tilde{z})$ and $\sigma_b(\tilde{z}) = \bar{\sigma}_b(\tilde{z})$.

Assumption U. *We assume that the uncertainties $\{\tilde{z}_j\}_{j=1:N}$ are zero mean random variables, with a positive definite covariance matrix $\mathbf{\Sigma}$. Let $\mathscr{W}$ be the smallest closed convex set containing the support of $\tilde{\mathbf{z}}$. We denote a subset, $\mathscr{J} \subseteq \{1,\ldots,N\}$, which can be an empty set, such that $\tilde{z}_j$, $j\in\mathscr{J}$ are stochastically independent. Moreover, the corresponding forward and backward deviations (or their bounds used in Theorem 2.2) are given by $p_j = \sigma_f(\tilde{z}_j)$ and $q_j = \sigma_b(\tilde{z}_j)$ respectively, for $j\in\mathscr{J}$ and $p_j = q_j = \infty$ for $j\notin\mathscr{J}$.*

The choice of the set $\mathscr{W}$ (with a little abuse of terminology, we call it the "support set") can influence the computational tractability of the problem. Henceforth, we assume that the support set is a second-order conic representable set (a.k.a. conic quadratic representable set) proposed in Ben-Tal and Nemirovski (1998), which includes polyhedral and ellipsoidal sets. A common support set is the interval set, given by $\mathscr{W} = [-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$, in which $\underline{\mathbf{z}}, \bar{\mathbf{z}} > \mathbf{0}$.

For notational convenience, we define the following sets:

$$\mathscr{J}_1 \triangleq \{i\colon p_j < \infty\}, \quad \bar{\mathscr{J}}_1 \triangleq \{i\colon p_j = \infty\},$$
$$\mathscr{J}_2 \triangleq \{i\colon q_j < \infty\}, \quad \bar{\mathscr{J}}_2 \triangleq \{i\colon q_j = \infty\}.$$

We also denote $\mathbf{P} = \mathrm{diag}(p_1,\ldots,p_N)$ and $\mathbf{Q} = \mathrm{diag}(q_1,\ldots,q_N)$. If $p_j = \infty$ (respectively, $q_j = \infty$), then we stipulate $p_j^{-1} = 0$ (respectively, $q_j^{-1} = 0$). Moreover, the product of any p_j with zero remains zero, i.e., $p_j \times 0 = 0$ (respectively, $q_j \times 0 = 0$).

Theorem 2.3 (Chen and Sim 2009). *Suppose that the primitive uncertainty $\tilde{\mathbf{z}}$ satisfies Assumption U. The following functions are upper bounds of $\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+)$, where $\mathbf{y} = (y_1,\ldots,y_N)'$:*

(a) $\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+) \leqslant \pi^1(y_0,\mathbf{y}) \triangleq \left(y_0+\max_{\mathbf{z}\in\mathscr{W}}\mathbf{z}'\mathbf{y}\right)^+.$

The bound is tight whenever $y_0+\mathbf{y}'\mathbf{z}\leqslant 0$ for all $\mathbf{z}\in\mathscr{W}$.

(b)
$$\begin{aligned}\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+) &= y_0+\mathrm{E}((-y_0-\mathbf{y}'\tilde{\mathbf{z}})^+)\\ &\leqslant \pi^2(y_0,\mathbf{y}) \triangleq y_0+\left(-y_0+\max_{\mathbf{z}\in\mathscr{W}}(-\mathbf{y})'\mathbf{z}\right)^+.\end{aligned}$$

The bound is tight whenever $y_0+\mathbf{y}'\mathbf{z}\geqslant 0$ for all $\mathbf{z}\in\mathscr{W}$.

(c)
$$\begin{aligned}\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+) &= \tfrac{1}{2}(y_0+\mathrm{E}(|y_0+\mathbf{y}'\tilde{\mathbf{z}})|)\\ &\leqslant \pi^3(y_0,\mathbf{y}) \triangleq \tfrac{1}{2}y_0+\tfrac{1}{2}\sqrt{y_0^2+\mathbf{y}'\mathbf{\Sigma}\mathbf{y}}.\end{aligned}$$

(d) $$\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+) \leqslant \inf_{\mu>0} \frac{\mu}{e}\mathrm{E}\left(\exp\left(\frac{y_0+\mathbf{y}'\tilde{\mathbf{z}}}{\mu}\right)\right) \leqslant \pi^4(y_0,\mathbf{y}) \triangleq \inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(\frac{y_0}{\mu}+\frac{\|\mathbf{u}\|_2^2}{2\mu^2}\right)\right\},$$

where $u_j=\max\{p_jy_j,-q_jy_j\}$, $j=1,\dots,N$. *The bound is finite if and only if* $y_j\leqslant 0\ \forall j\in\bar{\mathcal{I}}_1$ *and* $y_j\geqslant 0\ \forall j\in\bar{\mathcal{I}}_2$.

(e) $$\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+) \leqslant y_0+\inf_{\mu>0} \frac{\mu}{e}\mathrm{E}\left(\exp\left(\frac{-y_0-\mathbf{y}'\tilde{\mathbf{z}}}{\mu}\right)\right) \leqslant \pi^5(y_0,\mathbf{y}) \triangleq y_0+\inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(-\frac{y_0}{\mu}+\frac{\|\mathbf{v}\|_2^2}{2\mu^2}\right)\right\},$$

where $v_j=\max\{-p_jy_j,q_jy_j\}$, $j=1,\dots,N$. *The bound is finite if and only if* $y_j\geqslant 0\ \forall j\in\bar{\mathcal{I}}_1$ *and* $y_j\leqslant 0\ \forall j\in\bar{\mathcal{I}}_2$.

Remark. Observe that $\pi^i(y_0,\mathbf{y})$, $i=1,\dots,5$ are convex, proper (i.e., the function is nowhere $-\infty$ and is not everywhere $+\infty$), and closed (i.e., lower semicontinuous). A closed convex function is necessarily continuous on its domain $\operatorname{dom} f\triangleq\{x\colon f(x)<+\infty\}$ (cf. Rockafellar 1970). In addition, $\pi^i(y_0,\mathbf{y})$, $i=1,\dots,5$ are positively homogeneous functions, that is,

$$\pi^i(ky_0,k\mathbf{y})=k\pi^i(y_0,\mathbf{y})\quad\forall k\geqslant 0. \tag{8}$$

Furthermore,

$$\pi^i(y_0,\mathbf{0})=y_0^+. \tag{9}$$

Chen and Sim (2009) showed that the epigraph of $\pi^i(y_0,\mathbf{y})$ is second-order cone representable and that the bound can be strengthened further by suitably decomposing $(y_0,\mathbf{y})$ into $(y_0^i,\mathbf{y}^\mathbf{i})$ and by using a linear combination of the bounds $\pi^i(y_0^i,\mathbf{y}^\mathbf{i})$.

Theorem 2.4 (Chen and Sim 2009). *Suppose that* $\pi^i(y_0,\mathbf{y})$ *for all* $i\in\mathcal{L}$ *is an upper bound to* $\mathrm{E}((y_0+\mathbf{y}'z)^+)$, *and* $\pi^i(y_0,\mathbf{y})$ *is convex and positively homogeneous. Define*

$$\begin{aligned}\pi^{\mathcal{L}}(y_0,\mathbf{y})\triangleq\min_{y_{l0},\mathbf{y}_l}\ &\sum_{l\in\mathcal{L}}\pi^l(y_{l0},\mathbf{y}_l)\\ \text{s.t.}\ &\sum_{l\in\mathcal{L}}y_{l0}=y_0,\\ &\sum_{l\in\mathcal{L}}\mathbf{y}_l=\mathbf{y}.\end{aligned}$$

Then,

$$0\leqslant\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+)\leqslant\pi^{\mathcal{L}}(y_0,\mathbf{y})\leqslant\min_{l\in\mathcal{L}}\pi^l(y_0,\mathbf{y}). \tag{10}$$

Moreover, $\pi^{\mathcal{L}}(y_0,\mathbf{y})$ *inherits the second-order cone representability and positively homogeneous properties of the individual functions* $\pi^i(y_0,\mathbf{y})$, $i\in\mathcal{L}$.

For details, the interested reader may refer to Chen and Sim (2009).

Proposition 2.5. *Under Assumption* U *and supposing that* $\pi(y_0,\mathbf{y})$ *is an upper bound to* $\mathrm{E}((y_0+\mathbf{y}'\mathbf{z})^+)$ *for all* $(y_0,\mathbf{y})\in\Re^{N+1}$, *then*

$$\pi(y_0,\mathbf{y})=0 \tag{11}$$

only if

$$y_0+\max_{\mathbf{z}\in\mathcal{W}}\mathbf{y}'\mathbf{z}\leqslant 0. \tag{12}$$

Proof. Note that

$$0=\pi(y_0,\mathbf{y})\geqslant\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+)\geqslant 0.$$

Suppose that

$$y_0+\mathbf{y}'\mathbf{z}^*=y_0+\max_{\mathbf{z}\in\mathcal{W}}\mathbf{y}'\mathbf{z}>0$$

for some $\mathbf{z}^*\in\mathcal{W}$. Because the objective function is linear, we can assume WLOG that $\mathbf{z}^*$ is an extreme point in $\mathcal{W}$.

Let $B_\epsilon(\mathbf{z}^*)$ denote an open ball with radius ϵ around $\mathbf{z}^*$, with

$$y_0+\mathbf{y}'\mathbf{z}>0\quad\text{for all } z\in B_\epsilon(\mathbf{z}^*).$$

Because $\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+)=0$, we must have

$$\mathrm{P}(\{\tilde{\mathbf{z}}\in B_\epsilon(\mathbf{z}^*)\})=0.$$

Thus, the support for $\tilde{\mathbf{z}}$ lies in the convex hull $\mathcal{W}'$ of the (closed) set $\mathcal{W}\backslash B_\epsilon(\mathbf{z}^*)$. Because $\mathbf{z}^*$ is an extreme point in $\mathcal{W}$, we have $\mathbf{z}^*\notin\mathcal{W}'$. This contradicts our earlier assumption that $\mathcal{W}$ denotes the smallest convex set containing the support for $\tilde{\mathbf{z}}$. □

2.2. Bounds on CVaR and Robust Optimization

There are several attractive proposals for approximating individual chance-constrained problems, in which the solution $(y_0,\mathbf{y})$ to the following problem

$$y_0+\max_{\mathbf{z}\in\mathcal{U}}\mathbf{y}'\mathbf{z}\leqslant 0$$

guarantees that

$$\mathrm{P}(y_0+\mathbf{y}'\tilde{\mathbf{z}}\leqslant 0)\geqslant 1-\epsilon. \tag{13}$$

Clearly, the choice of uncertainty set $\mathcal{U}$ depends on the underlying assumption on primitive uncertainty.

Another approach of approximating the chance constraint is to provide an upper bound of the CVaR function $\rho_{1-\epsilon}(y_0+\mathbf{y}'\tilde{\mathbf{z}})$, so that if the bound is nonnegative, the chance constraint (13) will also be satisfied. For a given upper bound $\pi(y_0,\mathbf{y})$ to $\mathrm{E}((\cdot)^+)$, we define

$$\eta_{1-\epsilon}(y_0,\mathbf{y})\triangleq\min_\beta\left\{\beta+\frac{1}{\epsilon}\pi(y_0-\beta,\mathbf{y})\right\}.$$

Clearly,

$$\rho_{1-\epsilon}(y_0+\mathbf{y}'\tilde{\mathbf{z}}) = \min_{\beta}\left\{\beta+\frac{1}{\epsilon}\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}}-\beta)^+)\right\} \leqslant \eta_{1-\epsilon}(y_0,\mathbf{y}),$$

and a sufficient condition for satisfying (13) is

$$\eta_{1-\epsilon}(y_0,\mathbf{y}) \leqslant 0. \tag{14}$$

Note that if the epigraph of $\pi(\cdot,\cdot)$ can be approximated by a second-order cone, constraint (14) is also approximately second-order cone representable.

We next show that the two approaches are essentially equivalent.

Theorem 2.6. *Suppose that $\pi(y_0,\mathbf{y})$ is convex, closed, and positively homogeneous, and is an upper bound to $\mathrm{E}((y_0+\mathbf{y}'z)^+)$ with $\pi(y_0,\mathbf{0})=y_0^+$. Then, under Assumption U and given $\epsilon\in(0,1)$, it holds that for all $(y_0,\mathbf{y})$ such that $\pi(y_0,\mathbf{y})<\infty$, we have*

$$\eta_{1-\epsilon}(y_0,\mathbf{y}) = y_0+\max_{\mathbf{z}\in\mathcal{U}(\epsilon)}\mathbf{y}'\mathbf{z}$$

for some convex uncertainty set $\mathcal{U}(\epsilon)$.

The proof of Theorem 2.6 is based on the following strong duality theorem, which can be found, e.g., in Ben-Tal and Nemirovski (2001).

Theorem 2.7 (Theorem 2.4.1, Ben-Tal and Nemirovski 2001). *Let $\mathcal{K}$ be a nonempty closed convex cone. Consider the primal problem*

$$p^* = \min_{\mathbf{x}}\{\mathbf{c}'\mathbf{x}\colon \mathbf{A}\mathbf{x}-\mathbf{b}\in\mathcal{K}\}$$

and its dual

$$d^* = \max_{\boldsymbol{\lambda}}\{\mathbf{b}'\boldsymbol{\lambda}\colon \mathbf{A}'\boldsymbol{\lambda}=\mathbf{c},\ \boldsymbol{\lambda}\in\mathcal{K}^*\}.$$

If the primal problem is bounded below and is strictly feasible (i.e., $\mathbf{A}\mathbf{x}-\mathbf{b}\in\mathrm{ri}(\mathcal{K})$ for some x, where $\mathrm{ri}(\mathcal{K})$ denotes the relative interior of the cone $\mathcal{K}$), then the dual problem is solvable and $-\infty<d^*=p^*<+\infty$.[2]

We now prove Theorem 2.6.

Proof. The set $\mathcal{K}\triangleq\{(u,y_0,\mathbf{y})\colon u\geqslant\pi(y_0,\mathbf{y})\}$ is a nonempty closed convex cone because it is the epigraph of a convex, closed, and positively homogeneous function with $\pi(y_0,\mathbf{0})=y_0^+$. Let $(y_0,\mathbf{y})\in\mathrm{dom}\,\pi$. Define

$$\mathbf{c}=(\epsilon^{-1},-1)',\quad \mathbf{b}=(0,-y_0,-\mathbf{y}')',$$
$$\mathbf{x}=(u,-\beta)',\quad \boldsymbol{\lambda}=(\gamma,-z_0,-\mathbf{z}')',$$

and

$$\mathbf{A}=\begin{pmatrix}1&0\\0&1\\\mathbf{0}&\mathbf{0}\end{pmatrix}.$$

We apply Theorem 2.7 with the above $\mathbf{c}$, $\mathbf{x}$, $\mathbf{b}$, $\boldsymbol{\lambda}$, $\mathbf{A}$, and $\mathcal{K}$. The primal problem is strictly feasible because there exists u such that $u>\pi(y_0,\mathbf{y})$ and that $\pi(y_0,\mathbf{y})<\infty$; i.e., $\exists\mathbf{x}$ such that $\mathbf{A}\mathbf{x}-\mathbf{b}\in\mathrm{ri}(\mathcal{K})$. If $\beta\geqslant0$, then because $u\geqslant\pi(y_0-\beta,\mathbf{y})\geqslant\mathrm{E}((y_0-\beta+\tilde{\mathbf{z}}'\mathbf{y})^+)\geqslant0$, one has $\mathbf{c}'\mathbf{x}=\epsilon^{-1}u+\beta\geqslant0$; whereas if $\beta<0$, then $\mathbf{c}'\mathbf{x}=\epsilon^{-1}u+\beta\geqslant\epsilon^{-1}\pi(y_0-\beta,\mathbf{y})+\beta\geqslant\epsilon^{-1}\mathrm{E}(y_0-\beta+\mathbf{y}'\tilde{\mathbf{z}})+\beta=\epsilon^{-1}y_0-(\epsilon^{-1}-1)\beta>\epsilon^{-1}y_0$, which shows that the primal problem is bounded below. Thus, by Theorem 2.7,

$$\begin{aligned}\eta_{1-\epsilon}(y_0,\mathbf{y}) &= \min_{\beta,u}\{\beta+u/\epsilon\colon (u,y_0-\beta,\mathbf{y})\in\mathcal{K}\}\\ &= \min_{\mathbf{x}}\{\mathbf{c}'\mathbf{x}\colon \mathbf{A}\mathbf{x}-\mathbf{b}\in\mathcal{K}\}\\ &= \max_{\boldsymbol{\lambda}}\{\mathbf{b}'\boldsymbol{\lambda}\colon \mathbf{A}'\boldsymbol{\lambda}=\mathbf{c},\ \boldsymbol{\lambda}\in\mathcal{K}^*\}\\ &= \max_{\gamma,z_0,\mathbf{z}}\big\{y_0z_0+\mathbf{y}'\mathbf{z}\colon (\gamma,-z_0,-\mathbf{z}')' = (1/\epsilon,-1,-\mathbf{z}')'\in\mathcal{K}^*\big\}.\end{aligned}$$

Hence,

$$\eta_{1-\epsilon}(y_0,\mathbf{y}) = y_0+\max_{\mathbf{z}\in\mathcal{U}(\epsilon)}\mathbf{y}'\mathbf{z},$$

with

$$\mathcal{U}(\epsilon)\triangleq\{\mathbf{z}\colon (1/\epsilon,-1,-\mathbf{z}')\in\mathcal{K}^*\}. \quad\square$$

For the functions $\pi^i(y_0,\mathbf{y})$, $i=1,\ldots,5$, the corresponding uncertainty sets can be computed explicitly. Consider the following uncertainty sets:

$$\mathcal{U}_1(\epsilon)\triangleq\mathcal{W},$$
$$\mathcal{U}_2(\epsilon)\triangleq\{\mathbf{z}\colon \mathbf{z}=(1-1/\epsilon)\boldsymbol{\zeta}\ \text{ for some }\boldsymbol{\zeta}\in\mathcal{W}\},$$
$$\mathcal{U}_3(\epsilon)\triangleq\left\{\mathbf{z}\colon \|\boldsymbol{\Sigma}^{-1/2}\mathbf{z}\|_2\leqslant\sqrt{\frac{1-\epsilon}{\epsilon}}\right\},$$
$$\mathcal{U}_4(\epsilon)\triangleq\left\{\mathbf{z}\colon \exists\mathbf{s},\mathbf{t}\in\Re^N,\ \mathbf{z}=\mathbf{s}-\mathbf{t},\ \|\mathbf{P}^{-1}\mathbf{s}+\mathbf{Q}^{-1}\mathbf{t}\|_2\leqslant\sqrt{-2\ln\epsilon}\right\},$$
$$\mathcal{U}_5(\epsilon)\triangleq\left\{\mathbf{z}\colon \exists\mathbf{s},\mathbf{t}\in\Re^N,\ \mathbf{z}=\mathbf{s}-\mathbf{t},\ \|\mathbf{Q}^{-1}\mathbf{s}+\mathbf{P}^{-1}\mathbf{t}\|_2\leqslant\frac{1-\epsilon}{\epsilon}\sqrt{-2\ln(1-\epsilon)}\right\}.$$

Note that, in general, the matrixes $\mathbf{P}^{-1}$ and $\mathbf{Q}^{-1}$ may not be positive definite. Hence, except for $\mathcal{U}_3$, the rest of the uncertainty sets may be unbounded.

Corollary 2.8.

$$\eta^i_{1-\epsilon}(y_0,\mathbf{y})\triangleq\min_{\beta}\left\{\beta+\frac{1}{\epsilon}\pi^i(y_0-\beta,\mathbf{y})\right\} = y_0+\max_{\mathbf{z}\in\mathcal{U}_i(\epsilon)}\mathbf{y}'\mathbf{z}.$$

Proof.

Uncertainty Set $\mathcal{U}_1(\epsilon)$:

$$\begin{aligned}\eta^1_{1-\epsilon}(y_0,\mathbf{y}) &= \min_{\beta}\left(\beta+\frac{\pi^1(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\ &= \min_{\beta}\left(\beta+\frac{1}{\epsilon}\Big(y_0-\beta+\max_{\mathbf{z}\in\mathcal{W}}\mathbf{y}'\mathbf{z}\Big)^+\right)\\ &= y_0+\max_{\mathbf{z}\in\mathcal{U}_1(\epsilon)}\mathbf{y}'\mathbf{z}.\end{aligned}$$

Uncertainty Set $\mathcal{U}_2(\epsilon)$:

$$\begin{aligned}
&\eta^2_{1-\epsilon}(y_0,\mathbf{y})\\
&=\min_\beta\left(\beta+\frac{\pi^2(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\
&=y_0+\min_\beta\left(\beta+\frac{\pi^2(-\beta,\mathbf{y})}{\epsilon}\right)\\
&=y_0+\min_\beta\left\{\beta+\frac{1}{\epsilon}\left(\left(\max_{\mathbf{z}\in\mathcal{W}}(-\mathbf{y})'\mathbf{z}+\beta\right)^+-\beta\right)\right\}\\
&=y_0+\min_\beta\left\{\beta(1-1/\epsilon)+\frac{1}{\epsilon}\left(\left(\max_{\mathbf{z}\in\mathcal{W}}(-\mathbf{y})'\mathbf{z}+\beta\right)^+\right)\right\}\\
&=y_0+(1/\epsilon-1)\min_\beta\left\{-\beta+\frac{1}{1-\epsilon}\left(\left(\max_{\mathbf{z}\in\mathcal{W}}(-\mathbf{y})'\mathbf{z}+\beta\right)^+\right)\right\}\\
&=y_0+(1/\epsilon-1)\max_{\mathbf{z}\in\mathcal{W}}\mathbf{y}'(-\mathbf{z})\\
&\quad+(1/\epsilon-1)\min_\beta\left(-\beta+\frac{1}{1-\epsilon}(\beta)^+\right)\\
&=y_0+\max_{\mathbf{z}\in\mathcal{U}_2(\epsilon)}\mathbf{y}'\mathbf{z}.
\end{aligned}$$

Uncertainty Set $\mathcal{U}_3(\epsilon)$:

$$\begin{aligned}
\eta^3_{1-\epsilon}(y_0,\mathbf{y})&=\min_\beta\left(\beta+\frac{\pi^3(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\
&=\min_\beta\left(\beta+\frac{y_0-\beta+\sqrt{(y_0-\beta)^2+\mathbf{y}'\boldsymbol{\Sigma}\mathbf{y}}}{2\epsilon}\right)\\
&=y_0+\sqrt{\frac{1-\epsilon}{\epsilon}}\sqrt{\mathbf{y}'\boldsymbol{\Sigma}\mathbf{y}}\\
&=y_0+\max_{\mathbf{z}\in\mathcal{U}_3(\epsilon)}\mathbf{y}'\mathbf{z},
\end{aligned}$$

where the second equality follows from choosing the optimum β,

$$\beta^*=y_0+\frac{\sqrt{\mathbf{y}'\boldsymbol{\Sigma}\mathbf{y}}(1-2\epsilon)}{2\sqrt{\epsilon(1-\epsilon)}}.$$

Uncertainty Set $\mathcal{U}_4(\epsilon)$: Observe that

$u_j\geqslant\max\{p_jy_j,-q_jy_j\}$ if and only if $p_j^{-1}u_j\geqslant y_j$ and $q_j^{-1}u_j\geqslant -y_j$.

$$\begin{aligned}
&\eta^4_{1-\epsilon}(y_0,\mathbf{y})\\
&=\min_\beta\left(\beta+\frac{\pi^4(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\
&=\min_{\beta,\mu,\mathbf{u}}\left(\beta+\frac{(\mu/e)\exp((y_0-\beta)/\mu+\|\mathbf{u}\|_2^2/(2\mu^2))}{\epsilon}\,\middle|\,\mathbf{P}^{-1}\mathbf{u}\geqslant\mathbf{y},\mathbf{Q}^{-1}\mathbf{u}\geqslant-\mathbf{y}\right)\\
&=\min_{\mu,\mathbf{u}}\left(y_0+\frac{\|\mathbf{u}\|_2^2}{2\mu}-\mu\ln\epsilon\,\middle|\,\mathbf{P}^{-1}\mathbf{u}\geqslant\mathbf{y},\mathbf{Q}^{-1}\mathbf{u}\geqslant-\mathbf{y}\right)\\
&=\min_{\mathbf{u}}\left(y_0+\sqrt{-2\ln\epsilon}\,u_0\,\middle|\,\mathbf{P}^{-1}\mathbf{u}\geqslant\mathbf{y},\mathbf{Q}^{-1}\mathbf{u}\geqslant-\mathbf{y},\|\mathbf{u}\|_2\leqslant u_0\right)\\
&=y_0+\max_{\mathbf{z}\in\mathcal{U}_4(\epsilon)}\mathbf{y}'\mathbf{z},
\end{aligned}$$

where the second and third equalities follow from choosing the tightest β^* and μ^*, that is,

$$\beta^*=y_0+\frac{\|\mathbf{u}\|_2^2}{2\mu}-\mu\ln\epsilon-\mu,$$

$$\mu^*=\frac{\|\mathbf{u}\|_2}{\sqrt{-2\ln\epsilon}}.$$

The last equality is the result of conic duality. See, for example, Chen et al. (2007).

Uncertainty Set $\mathcal{U}_5(\epsilon)$: Following from the above exposition,

$$\begin{aligned}
&\eta^5_{1-\epsilon}(y_0,\mathbf{y})\\
&=\min_\beta\left(\beta+\frac{\pi^5(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\
&=\min_{\beta,\mu,\mathbf{v}}\left(\beta+\frac{y_0-\beta+(\mu/e)\exp(-(y_0-\beta)/\mu+\|\mathbf{v}\|_2^2/(2\mu^2))}{\epsilon}\,\middle|\,\mathbf{P}^{-1}\mathbf{v}\geqslant-\mathbf{y},\mathbf{Q}^{-1}\mathbf{v}\geqslant\mathbf{y}\right)\\
&=\min_{\mu,\mathbf{v}}\left(y_0+\left(\frac{1}{\epsilon}-1\right)\left(\frac{\|\mathbf{v}\|_2^2}{2\mu}-\mu\ln(1-\epsilon)\right)\,\middle|\,\mathbf{P}^{-1}\mathbf{v}\geqslant-\mathbf{y},\mathbf{Q}^{-1}\mathbf{v}\geqslant\mathbf{y}\right)\\
&=\min_{\mathbf{v}}\left(y_0+\frac{1-\epsilon}{\epsilon}\sqrt{-2\ln(1-\epsilon)}\|\mathbf{v}\|_2\,\middle|\,\mathbf{P}^{-1}\mathbf{v}\geqslant-\mathbf{y},\mathbf{Q}^{-1}\mathbf{v}\geqslant\mathbf{y}\right)\\
&=y_0+\max_{\mathbf{z}\in\mathcal{U}_5(\epsilon)}\mathbf{y}'\mathbf{z}.\quad\square
\end{aligned}$$

We show next that the uncertainty set corresponding to the stronger bound $\pi^{\mathcal{L}}(y_0,\mathbf{y})$ can also be obtained in a similar way.

Theorem 2.9. *Suppose that* $\tilde{\mathbf{z}}$ *satisfies Assumption* U. *Let*

$$\mathcal{U}_{\mathcal{L}}(\epsilon)\triangleq\bigcap_{l\in\mathcal{L}}\mathcal{U}_l(\epsilon),$$

and suppose that $\mathcal{U}_{\mathcal{L}}(\epsilon)$ *is compact and has a nonempty interior. Then,*

$$\eta^{\mathcal{L}}_{1-\epsilon}(y_0,\mathbf{y})=y_0+\max_{\mathbf{z}\in\mathcal{U}_{\mathcal{L}}(\epsilon)}\mathbf{y}'\mathbf{z}.$$

Proof.

$$\begin{aligned}
&\eta^{\mathcal{L}}(y_0,\mathbf{y})\\
&=\min_\beta\left(\beta+\frac{\pi^{\mathcal{L}}(y_0-\beta,\mathbf{y})}{\epsilon}\right)\\
&=\min_{\boldsymbol{\beta},y_{l0},\mathbf{y}_l,l\in\mathcal{L}}\left(\beta+\sum_{l\in\mathcal{L}}\left(\frac{\pi^l(y_{l0}-\beta_l,\mathbf{y}_l)}{\epsilon}\right)\,\middle|\,\sum_{l\in\mathcal{L}}\mathbf{y}_l=\mathbf{y},\sum_{l\in\mathcal{L}}y_{l0}=y_0,\sum_{l\in\mathcal{L}}\beta_l=\beta\right)
\end{aligned}$$

$$= \min_{y_{l0},\mathbf{y}_l,\,l\in\mathscr{L}} \left(\sum_{l\in\mathscr{L}} \min_{\beta_l} \left(\beta_l + \frac{\pi^l(y_{l0}-\beta_l,\mathbf{y}_l)}{\epsilon} \right) \,\middle|\, \sum_{l\in\mathscr{L}} \mathbf{y}_l = \mathbf{y}, \sum_{l\in\mathscr{L}} y_{l0} = y_0 \right)$$

$$= \min_{y_{l0},\mathbf{y}_l,\,l\in\mathscr{L}} \left(\sum_{l\in\mathscr{L}} \left(y_{l0} + \max_{\mathbf{z}\in\mathscr{U}_l(\epsilon)} \mathbf{y}_l'\mathbf{z} \right) \,\middle|\, \sum_{l\in\mathscr{L}} \mathbf{y}_l = \mathbf{y}, \sum_{l\in\mathscr{L}} y_{l0} = y_0 \right)$$

$$= y_0 + \min_{\mathbf{y}_l,\,l\in\mathscr{L}} \left(\sum_{l\in\mathscr{L}} \left(\max_{\mathbf{z}\in\mathscr{U}_l(\epsilon)} \mathbf{y}_l'\mathbf{z} \right) \,\middle|\, \sum_{l\in\mathscr{L}} \mathbf{y}_l = \mathbf{y} \right)$$

$$= y_0 + \max_{\mathbf{z}\in\mathscr{U}_{\mathscr{L}}(\epsilon)} \mathbf{y}'\mathbf{z},$$

where the last inequality is due to infimum convolution of support functions. See Corollary 16.4.1 of Rockafellar (1970). □

Hence, the different approximations to individual chance-constrained problems used in robust optimization are the consequences of applying different bounds on $\mathrm{E}((\cdot)^+)$. Notably, when the primitive uncertainties are characterized only by their means and covariance, the corresponding uncertainty set is an ellipsoid of the form $\mathscr{U}_3(\epsilon)$. See, for example, Bertsimas et al. (2004) and El Ghaoui et al. (2003). When $I = N$, that is, all the primitive uncertainties are independently distributed, Chen et al. (2007) proposed the asymmetrical uncertainty set

$$\mathscr{U}_A(\epsilon) = \underbrace{\mathscr{W}}_{=\mathscr{U}_1(\epsilon)} \cap \mathscr{U}_4(\epsilon),$$

which generalizes the uncertainty set proposed by Ben-Tal and Nemirovski (2000). With the independence assumption, it suffices to consider the interval set given by $\mathscr{W} = [-\underline{\mathbf{z}}, \bar{\mathbf{z}}]$.

Noting that $\mathscr{U}_A(\epsilon) \supseteq \mathscr{U}_{\{1,2,4,5\}}(\epsilon)$, we can therefore improve upon the approximation using the uncertainty set $\mathscr{U}_{\{1,2,4,5\}}(\epsilon)$. However, in most application of chance-constrained problems, the safety factor ϵ is relatively small, in which case the uncertainty sets of $\mathscr{U}_2(\epsilon)$ and $\mathscr{U}_5(\epsilon)$ are usually exploded to engulf the uncertainty sets of $\mathscr{W}$ and $\mathscr{U}_4(\epsilon)$, respectively. For example, under symmetric distributions, that is $\mathbf{P} = \mathbf{Q}$ and $\bar{\mathbf{z}} = \underline{\mathbf{z}}$, it is easy to establish that for $\epsilon < 0.5$, we have

$$\mathscr{U}_{\{1,2,4,5\}}(\epsilon) = \underbrace{\mathscr{U}_1(\epsilon)}_{=\mathscr{W}} \cap \underbrace{\mathscr{U}_2(\epsilon)}_{\supseteq\mathscr{W}} \cap \mathscr{U}_4(\epsilon) \cap \underbrace{\mathscr{U}_5(\epsilon)}_{\supseteq\mathscr{U}_4} = \mathscr{U}_A(\epsilon).$$

3. Joint Chance-Constrained Problems

Unfortunately, the notion of uncertainty set in classical robust optimization does not carry forward as well in addressing joint chance-constrained problems. We consider a linear joint chance constraint as follows:

$$\mathrm{P}(y_j(\tilde{\mathbf{z}}) \leqslant 0,\, j\in\mathscr{M}) \geqslant 1-\epsilon, \qquad (15)$$

where $\mathscr{M} = \{1,\ldots,m\}$, $y_j(\tilde{\mathbf{z}})$ are affinely dependent of $\tilde{\mathbf{z}}$,

$$y_j(\tilde{\mathbf{z}}) = y_0^j + \sum_{k=1}^N y_k^j \tilde{z}_k, \quad j\in\mathscr{M},$$

and $(y_0^1,\ldots,y_N^1,\ldots,y_0^m,\ldots,y_N^m)$ being the decision variables. For notational convenience, we represent

$$\mathbf{y}_j = (y_1^j,\ldots,y_N^j),$$

so that $y_i(\tilde{\mathbf{z}}) = y_0^i + \mathbf{y}_i'\tilde{\mathbf{z}}$, and denote

$$\mathbf{Y} = (y_0^1,\ldots,y_N^1,\ldots,y_0^m,\ldots,y_N^m)$$

as the collection of decision variables in the joint chance constraint. By suitable affine constraints imposed on the decision variables $\mathbf{Y}$ and $\mathbf{x}$, we can represent the joint chance constraint in model (2) in the form of constraint (15).

It is not surprising that a joint chance constraint is more difficult to solve than an individual one. For computational tractability, the common approach is to decompose the joint constraint into a problem with m individual constraints of the form

$$\mathrm{P}(y_i(\tilde{\mathbf{z}}) \leqslant 0) \geqslant 1-\epsilon_i, \quad i\in\mathscr{M}. \qquad (16)$$

By enforcing Bonferroni's inequality on their safety factors,

$$\sum_{i\in\mathscr{M}} \epsilon_i \leqslant \epsilon, \qquad (17)$$

any feasible solution that satisfies (16) and (17) will also satisfy (15). See, for example, Chen et al. (2007) and Nemirovski and Shapiro (2006). Consequently, using the techniques discussed in the previous section, we can then build tractable safe approximations as follows:

$$\eta_{1-\epsilon_i}(y_0^i, \mathbf{y}_i) \leqslant 0, \quad i\in\mathscr{M}. \qquad (18)$$

The main issue with using Bonferroni's inequality is the choice of ϵ_i. Unfortunately, the problem becomes nonconvex and possibly intractable if ϵ_i, $i\in\mathscr{M}$, are made decision variables and (17) is enforced as a constraint in the optimization model. As such, it is natural to choose, $\epsilon_i = \epsilon/m$ as proposed in Chen et al. (2007) and Nemirovski and Shapiro (2006).

In some instances, this approach may be rather conservative even for an optimal choice of ϵ_i. For example, suppose that $y_i(\tilde{\mathbf{z}})$ are completely correlated, such as

$$y_i(\tilde{\mathbf{z}}) = \delta_i(a_0 + \mathbf{a}'\tilde{\mathbf{z}}), \quad i\in\mathscr{M} \qquad (19)$$

for some $\delta_i > 0$; the least conservative choice of ϵ_i is $\epsilon_i = \epsilon$ for all $i\in\mathscr{M}$, which would violate condition (17) imposed by Bonferroni's inequality. As a matter of fact, it is easy to see in this case that the least conservative choice of ϵ_i, while satisfying Bonferroni's inequality, is $\epsilon_i = \epsilon/m$ for all

$i=1,\ldots,m$. Hence, if $y_i(\tilde{\mathbf{z}})$ are correlated, the efficacy of Bonferroni's inequality will possibly diminish.

We propose a new tractable way of approximating the joint chance constraint. Given a vector of positive constants, $\boldsymbol{\alpha}\in\Re^N$, $\boldsymbol{\alpha}>\mathbf{0}$, an index set $\mathcal{J}\subseteq\mathcal{M}$, an upper bound $\pi(y_0,\mathbf{y})$ for $\mathrm{E}((y_0+\mathbf{y}'\tilde{\mathbf{z}})^+)$, we define the following function:

$$\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J}) \triangleq \min_{w_0,\mathbf{w}}\Bigg\{\underbrace{\min_{\beta}\left[\beta+\frac{1}{\epsilon}\pi(w_0-\beta,\mathbf{w})\right]}_{=\eta_{1-\epsilon}(w_0,\mathbf{w})} + \frac{1}{\epsilon}\left[\sum_{i\in\mathcal{J}}\pi(\alpha_i y_0^i-w_0,\alpha_i\mathbf{y}_i-\mathbf{w})\right]\Bigg\}.$$

The next result shows that we can use the above function to approximate a joint chance constraint.

Theorem 3.1. (a) *Suppose that* $\tilde{\mathbf{z}}$ *satisfies Assumption* U. *If*

$$\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})\leqslant 0 \tag{20}$$

and

$$y_0^i+\max_{\mathbf{z}\in\mathcal{W}}\mathbf{y}_i'\mathbf{z}\leqslant 0\quad \forall\, i\in\mathcal{M}\backslash\mathcal{J}, \tag{21}$$

then

$$\rho_{1-\epsilon}\Big(\max_{i\in\mathcal{J}}\{\alpha_i y_i(\tilde{\mathbf{z}})\}\Big)\leqslant\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J}).$$

Consequently, the joint chance constraint (15) *is satisfied.*

(b) *For fixed* $(\boldsymbol{\alpha},\mathcal{J})$, *the epigraph of the function* $\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})$ *with respect to* $\mathbf{Y}$ *is second-order cone representable and positively homogeneous. Similarly, for fixed* $(\mathbf{Y},\mathcal{J})$, *the epigraph of the function* $\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})$ *with respect to* $\boldsymbol{\alpha}$ *is second-order cone representable and positively homogeneous.*

Proof. (a) Under Assumption U, the set $\mathcal{W}$ contains the support of the primitive uncertainty, $\tilde{\mathbf{z}}$, hence, the robust counterpart (21) implies

$$\mathrm{P}(y_0^i+\mathbf{y}_i'\tilde{\mathbf{z}}>0)=0\quad\forall\, i\in\mathcal{M}\backslash\mathcal{J}.$$

Hence, with $\boldsymbol{\alpha}>0$, we have

$$\mathrm{P}(y_0^i+\mathbf{y}_i'\tilde{\mathbf{z}}\leqslant 0,\ i\in\mathcal{M})=\mathrm{P}(y_0^i+\mathbf{y}_i'\tilde{\mathbf{z}}\leqslant 0,\ i\in\mathcal{J}) =\mathrm{P}\Big(\max_{i\in\mathcal{J}}\{\alpha_i y_0^i+\alpha_i\mathbf{y}_i'\tilde{\mathbf{z}}\}\leqslant 0\Big).$$

Therefore, it suffices to show that if $\mathbf{Y}$ is feasible in constraint (20), then the CVaR measure

$$\rho_{1-\epsilon}\Big(\max_{i\in\mathcal{J}}\{\alpha_i y_i(\tilde{\mathbf{z}})\}\Big)\leqslant 0.$$

Using the classical inequality (cf. Meilijson and Nadas 1979) that

$$\mathrm{E}\Big(\max_{i=1,\ldots,n}X_i-\beta\Big)^+\leqslant\mathrm{E}(Y-\beta)^+ +\sum_{i=1}^n\mathrm{E}(X_i-Y)^+\quad\text{for any r.v. } Y, \tag{22}$$

we have

$$\begin{aligned}
&\rho_{1-\epsilon}\Big(\max_{i\in\mathcal{J}}\{\alpha_i(y_0^i+\mathbf{y}_i'\tilde{\mathbf{z}})\}\Big)\\
&=\min_{\beta}\left\{\beta+\frac{1}{\epsilon}\mathrm{E}\left[\Big(\max_{i\in\mathcal{J}}\{\alpha_i(y_0^i+\mathbf{y}_i'\tilde{\mathbf{z}})\}-\beta\Big)^+\right]\right\}\\
&\leqslant\min_{\beta,w_0,\mathbf{w}}\left\{\beta+\frac{1}{\epsilon}\left[\mathrm{E}((w_0-\beta+\mathbf{w}'\tilde{\mathbf{z}})^+) +\sum_{i\in\mathcal{J}}\mathrm{E}((\alpha_i y_0^i-w_0+(\alpha_i\mathbf{y}_i-\mathbf{w})'\tilde{\mathbf{z}})^+)\right]\right\}\\
&\leqslant\min_{\beta,w_0,\mathbf{w}}\left\{\beta+\frac{1}{\epsilon}\left[\pi(w_0-\beta,\mathbf{w})+\sum_{i\in\mathcal{J}}\pi(\alpha_i y_0^i-w_0,\alpha_i\mathbf{y}_i-\mathbf{w})\right]\right\}\\
&=\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})\leqslant 0.
\end{aligned}$$

(b) For a fixed $\boldsymbol{\alpha}$, the corresponding epigraph can be expressed as

$$Y_1=\{(\mathbf{Y},t)\colon\ \gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})\leqslant t\} =\left\{(\mathbf{Y},t)\colon\ \begin{array}{l}\exists w_0,r_0,\ldots,r_m\in\Re,\quad \mathbf{w}\in\Re^N,\\ r_0+\dfrac{1}{\epsilon}\displaystyle\sum_{i\in\mathcal{J}}r_i\leqslant t,\\ \eta_{1-\epsilon}(w_0,\mathbf{w})\leqslant r_0,\\ \pi(\alpha_i y_0^i-w_0,\alpha_i\mathbf{y}_i-\mathbf{w})\leqslant r_i\quad\forall\, i\in\mathcal{J}.\end{array}\right\}.$$

Because the epigraphs of $\eta_{1-\epsilon}(\cdot,\cdot)$ and $\pi(\cdot,\cdot)$ are second-order cone representable, the set Y_1 is also second-order cone representable. For positive homogeneity, we observe that because $\pi(\cdot,\cdot)$ is positively homogeneous, we have that for all $k\geqslant 0$,

$$\begin{aligned}
\gamma_{1-\epsilon}(k\mathbf{Y},\boldsymbol{\alpha},\mathcal{J})&=\min_{\beta,w_0,\mathbf{w}}\left\{\beta+\frac{1}{\epsilon}\left[\pi(w_0-\beta,\mathbf{w}) +\sum_{i\in\mathcal{J}}\pi(k\alpha_i y_0^i-w_0,k\alpha_i\mathbf{y}_i-\mathbf{w})\right]\right\}\\
&=k\min_{\beta,w_0,\mathbf{w}}\left\{\frac{1}{k}\beta+\frac{1}{\epsilon}\left[\pi\Big(\frac{1}{k}w_0-\frac{1}{k}\beta,\frac{1}{k}\mathbf{w}\Big) +\sum_{i\in\mathcal{J}}\pi\Big(\alpha_i y_0^i-\frac{1}{k}w_0,\alpha_i\mathbf{y}_i-\frac{1}{k}\mathbf{w}\Big)\right]\right\}\\
&=k\min_{\beta,w_0,\mathbf{w}}\left\{\beta+\frac{1}{\epsilon}\left[\pi(w_0-\beta,\mathbf{w}) +\sum_{i\in\mathcal{J}}\pi(\alpha_i y_0^i-w_0,\alpha_i\mathbf{y}_i-\mathbf{w})\right]\right\}\\
&=k\gamma_{1-\epsilon}(\mathbf{Y},\boldsymbol{\alpha},\mathcal{J}).
\end{aligned}$$

Similarly, the same exposition applies when $\mathbf{Y}$ is fixed and $\boldsymbol{\alpha}$ is the decision variable. □

REMARK. Note that constraints (21) do not depend on the values of α_j for all $j \in \mathcal{M}\backslash\mathcal{J}$. Speaking intuitively, we can perceive $\alpha_j = \infty$ for all $j \in \mathcal{M}\backslash\mathcal{J}$. However, to avoid dealing with infinite entities, we define the set $\mathcal{J}$ as part of the input to the function $\gamma_{1-\epsilon}(\cdot,\cdot,\cdot)$. Throughout this paper, we will restrict the focus of $\boldsymbol{\alpha}$ to only elements corresponding to the indices in the set $\mathcal{J}$. Unfortunately, the function $\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$ is not jointly convex in both $\mathbf{Y}$ and $\boldsymbol{\alpha}$. Nevertheless, for a given $\mathbf{Y}$, it is a tractable convex function with respect to $\boldsymbol{\alpha}$ and is in the attractive form of SOCP. We will later exploit this property for improving the choice of $\boldsymbol{\alpha}$.

If the sets

$$\mathcal{S}_i \triangleq \{\tilde{\mathbf{z}}: y_i(\tilde{\mathbf{z}}) \geqslant \beta\}, \qquad i = 1, \ldots, n$$

are mutually disjoint, then

$$\mathrm{E}\Big(\max_i y_i(\tilde{\mathbf{z}}) - \beta\Big)^+ = \sum_{i=1}^n \mathrm{E}(y_i(\tilde{\mathbf{z}}) - \beta)^+,$$

and hence inequality (22) cannot be tightened further substantially. Interestingly, by introducing the parameters $\boldsymbol{\alpha}$ and random variable $w_0 + \mathbf{w}'\mathbf{z}$, our approach is also able to handle the situation when the variables are positively correlated. In the example (19) where $y_i(\tilde{\mathbf{z}})$, $i \in \mathcal{M}$ are completely positively correlated, the following condition

$$\eta_{1-\epsilon}(a^0, \mathbf{a}) \leqslant 0$$

is also sufficient to guarantee feasibility in the joint chance constraint. Choosing $\alpha_i = 1/\delta_i > 0$, we see that

$$\begin{aligned}
&\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{M})\\
&= \min_{w_0, \mathbf{w}} \left\{ \eta_{1-\epsilon}(w_0, \mathbf{w}) + \frac{1}{\epsilon}\left[\sum_{i\in\mathcal{M}} \pi(\alpha_i y_0^i - w_0, \alpha_i \mathbf{y}_i - \mathbf{w})\right]\right\}\\
&= \min_{w_0, \mathbf{w}} \left\{ \eta_{1-\epsilon}(w_0, \mathbf{w}) + \frac{1}{\epsilon}\left[\sum_{i\in\mathcal{M}} \pi(\alpha_i \delta_i a^0 - w_0, \alpha_i \delta_i \mathbf{a} - \mathbf{w})\right]\right\}\\
&\leqslant \eta_{1-\epsilon}(a^0, \mathbf{a}) + \frac{1}{\epsilon}\left\{\sum_{i\in\mathcal{M}} \pi(a^0 - a_0, \mathbf{a} - \mathbf{a})\right\}\\
&= \eta_{1-\epsilon}(a^0, \mathbf{a}) \leqslant 0.
\end{aligned}$$

Therefore, we see that the new bound is potentially better than the application of Bonferroni's inequality on individual chance constraints. By choosing the right combination of $(\boldsymbol{\alpha}, \mathcal{J})$, we can prove a stronger result as follows.

THEOREM 3.2. *Let $\epsilon_i \in (0,1), i \in \mathcal{M}$, and $\sum_{i\in\mathcal{M}} \epsilon_i \leqslant \epsilon$. Under Assumption* U, *suppose that* $\mathbf{Y}$ *satisfies*

$$\eta_{1-\epsilon_i}(y_0^i, \mathbf{y}_i) \leqslant 0 \quad \forall\, i \in \mathcal{M}.$$

Then, there exists $\boldsymbol{\alpha} > 0$, and a set $\mathcal{J} \subseteq \mathcal{M}$ such that $(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$ are feasible in constraints (20) *and* (21).

PROOF. Let β_i be the optimal solution to

$$\underbrace{\min_\beta \left(\beta + \frac{1}{\epsilon_i}(\pi(y_0^i - \beta, \mathbf{y}_i))\right)}_{=\eta_{1-\epsilon_i}(y_0^i, \mathbf{y}_i)}.$$

Because $\eta_{1-\epsilon_i}(y_0^i, \mathbf{y}_i) \leqslant 0$ and that

$$\pi(y_0^i - \beta_i, \mathbf{y}_i) \geqslant \mathrm{E}((y_0^i - \beta_i + \mathbf{y}_i'\tilde{\mathbf{z}})^+) \geqslant 0,$$

we must have $\beta_i \leqslant 0$. Let $\mathcal{J} = \{i \mid \beta_i < 0\}$ and

$$\alpha_j = -\frac{1}{\beta_j} \quad \forall\, j \in \mathcal{J}.$$

Because $\beta_j = 0$ for all $j \in \mathcal{M}\backslash\mathcal{J}$, we have

$$0 \leqslant \pi(y_0^i, \mathbf{y}_i) \leqslant 0 \quad \forall\, i \in \mathcal{M}\backslash\mathcal{J}.$$

From Proposition 2.5, it follows that

$$y_0^i + \mathbf{y}_i'\mathbf{z} \leqslant 0 \quad \forall\, \mathbf{z} \in \mathcal{W},\ \forall\, i \in \mathcal{M}\backslash\mathcal{J},$$

which satisfies the set of inequalities in (21).

For $i \in \mathcal{J}$, the constraint $\eta_{1-\epsilon_i}(y_0^i, \mathbf{y}_i) \leqslant 0$ is equivalent to

$$\frac{1}{-\beta_i}\pi(y_0^i - \beta_i, \mathbf{y}_i) \leqslant \epsilon_i.$$

Because the function $\pi(\cdot,\cdot)$ is positive homogeneous, we have

$$\begin{aligned}
\frac{1}{-\beta_i}\pi(y_0^i - \beta_i, \mathbf{y}_i) &= \pi\left(\frac{1}{-\beta_i}y_0^i + 1, \frac{1}{-\beta_i}\mathbf{y}_i\right)\\
&= \pi(\alpha_i y_0^i + 1, \alpha_i y_i)\\
&\leqslant \epsilon_i \quad \forall\, i \in \mathcal{J}.
\end{aligned}$$

Finally,

$$\begin{aligned}
&\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})\\
&= \min_{\beta, w_0, \mathbf{w}} \left\{\beta + \frac{1}{\epsilon}\left[\pi(w_0 - \beta, \mathbf{w}) + \sum_{i\in\mathcal{J}} \pi(\alpha_i y_0^i - w_0, \alpha_i \mathbf{y}_i - \mathbf{w})\right]\right\}\\
&\leqslant -1 + \frac{1}{\epsilon}\left\{\pi(-1+1, \mathbf{0}) + \sum_{i\in\mathcal{J}} \pi(\alpha_i y_0 + 1, \alpha_i \mathbf{y} - \mathbf{0})\right\}\\
&= -1 + \frac{1}{\epsilon}\sum_{i\in\mathcal{J}} \pi(\alpha_i y_0 + 1, \alpha_i \mathbf{y})\\
&\leqslant -1 + \frac{1}{\epsilon}\sum_{i\in\mathcal{J}} \epsilon_i \leqslant 0,
\end{aligned}$$

where the first inequality is due to the choice of $\beta = -1$, $w_0 = -1$, $\mathbf{w} = \mathbf{0}$ and the last inequality follows from $\sum_{i\in\mathcal{M}} \epsilon_i \leqslant \epsilon$. □

3.1. Optimizing Over $\boldsymbol{\alpha}$

Consider a joint chance-constrained model as follows:

$$\begin{aligned} Z_\epsilon = \min\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \mathrm{P}(y_i(\tilde{\mathbf{z}}) \leqslant 0,\ i \in \mathcal{M}) \geqslant 1-\epsilon, \\ & (\mathbf{x}, \mathbf{Y}) \in X, \end{aligned} \tag{23}$$

in which X is an efficiently computable convex set, such as a polyhedron or a second-order cone representable set. Given a set of a constant, $\boldsymbol{\alpha} > 0$, and a set $\mathcal{J}$, we consider the following optimization model:

$$\begin{aligned} Z^1_\epsilon(\boldsymbol{\alpha}, \mathcal{J}) = \min\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leqslant 0, \\ & y_0^i + \max_{\mathbf{z}\in\mathcal{W}} \mathbf{y}_i'\mathbf{z} \leqslant 0 \quad \forall\, i \in \mathcal{M}\backslash\mathcal{J}, \\ & (\mathbf{x}, \mathbf{Y}) \in X. \end{aligned} \tag{24}$$

In view of Theorem 3.1, suppose that model (24) is feasible, and the solution $(\mathbf{x}, \mathbf{Y})$ is also feasible in model (23), albeit more conservatively.

The main concern here is how to choose $\boldsymbol{\alpha}$ and $\mathcal{J}$. A likely choice is, say, $\alpha_j = 1/m$ for all $j \in \mathcal{M}$ and $\mathcal{J} = \mathcal{M}$. Alternatively, we may use the classical approach by decomposing into m individual chance constraints with $\epsilon_i = \epsilon/m$. In virtue of Theorem 3.2, we can find a feasible $\boldsymbol{\alpha} > \mathbf{0}$ and set $\mathcal{J}$ such that model (24) is also feasible.

Our aim is to improve upon the objective by minimizing $\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$ over $\boldsymbol{\alpha}$ and $\mathcal{J}$, resulting in greater slack in model (24). Hence, this approach will lead to improvement in the objective, or at least will not increase the value.

Given a feasible solution $\mathbf{Y}$ in model (24), our aim is to improve upon the objective by readjusting the set $\mathcal{J}$ and the weights α_j, $j \in \mathcal{J}$, which will result in greater slack in model (24) over the solution, $\mathbf{Y}$. We define the following set:

$$\mathcal{H}(\mathbf{Y}) \triangleq \left\{ i\colon\ y_0^i + \max_{\mathbf{z}\in\mathcal{W}} \mathbf{y}_i'\mathbf{z} > 0 \right\}.$$

Note that we can obtain the set $\mathcal{H}(\mathbf{Y})$ by solving the following linear optimization problem:

$$\begin{aligned} \min\ & \sum_{i=1}^m s_i \\ \text{s.t.}\ & y_0^i + \max_{\mathbf{z}\in\mathcal{W}} \mathbf{y}_i'\mathbf{z} \leqslant s_i, \end{aligned} \tag{25}$$

so that $\mathcal{H}(\mathbf{Y}) = \{i\colon s_i^* > 0\}$, $\mathbf{s}^*$ being its optimal solution.

Because $\mathbf{Y}$ is feasible in model (24), we must have $\mathcal{H}(\mathbf{Y}) \subseteq \mathcal{J}$. If the set $\mathcal{H}(\mathbf{Y})$ is nonempty, we consider the following optimization problem over α_j, $j \in \mathcal{H}(\mathbf{Y})$:

$$\begin{aligned} Z^1_\alpha(\mathbf{Y}) = \min\ & \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{H}(\mathbf{Y})) \\ \text{s.t.}\ & \sum_{j\in\mathcal{H}(\mathbf{Y})} \alpha_j = 1, \\ & \alpha_j \geqslant 0 \quad \forall\, j \in \mathcal{H}(\mathbf{Y}). \end{aligned} \tag{26}$$

By choosing $\pi(y_0, \mathbf{y}) \leqslant \pi^1(y_0, \mathbf{y})$, we can ensure that the objective function of problem (26) is finite. Moreover, because the feasible region of problem (26) is compact, the optimal solution for α_j, $j \in \mathcal{H}(\mathbf{Y})$ is therefore attained.

Proposition 3.3. *Assume that there exists* $(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$, *such that* $\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leqslant 0$. *Let* $\boldsymbol{\alpha}^*$ *be the optimum solution of problem* (26).

(a) $Z^1_\alpha(\mathbf{Y}) \leqslant 0$.

(b) *Moreover, the solution* $\boldsymbol{\alpha}^*$ *satisfies*

$$\alpha_i^* > 0 \quad \forall\, i \in \mathcal{H}(\mathbf{Y}).$$

Proof. (a) Because $\mathcal{H}(\mathbf{Y}) \subseteq \mathcal{J}$, and under the assumption that there exists $(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$ such that $\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leqslant 0$, by using the same $\boldsymbol{\alpha}$, we observe that

$$\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{H}) \leqslant \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leqslant 0.$$

Due to the positively homogeneous property of Theorem 3.1(b), we scale $\boldsymbol{\alpha}$ by a positive constant so that it is feasible in problem (26). Hence, the result follows.

(b) Note that under the constraints of problem (26), there exists $\alpha_j^* > 0$ for some $j \in \mathcal{H}(\mathbf{Y})$. Suppose that there exists a nonempty set $\mathcal{G} \subset \mathcal{H}(\mathbf{Y})$ (strict inclusion) such that $\alpha_i^* = 0$, $\forall\, i \in \mathcal{G}$. We will show that the following holds:

$$y_0^i + \max_{\mathbf{z}\in\mathcal{W}} \mathbf{y}_i'\mathbf{z} \leqslant 0 \quad \forall\, i \in \mathcal{H}(\mathbf{Y})\backslash\mathcal{G},$$

which is a contradiction. We have argued that $Z^1_\alpha(\mathbf{Y}) \leqslant 0$. Let $k \in \mathcal{G}$, that is, $\alpha_k^* = 0$. Observe that for some suitably chosen $(\beta, w_0, \mathbf{w})$,

$$\begin{aligned} 0 &\geqslant \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}^*, \mathcal{H}(\mathbf{Y})) \\ &= \beta + \frac{1}{\epsilon}\left\{ \pi(w_0-\beta, \mathbf{w}) + \sum_{i\in\mathcal{H}(\mathbf{Y})} \pi(\alpha_i^* y_0^i - w_0, \alpha_i^*\mathbf{y}_i - \mathbf{w}) \right\} \\ &= \beta + \frac{1}{\epsilon}\{\pi(w_0-\beta, \mathbf{w}) + \pi(-w_0, -\mathbf{w})\} \\ &\quad + \frac{1}{\epsilon} \sum_{i\in\mathcal{H}(\mathbf{Y})\backslash\{k\}} \pi(\alpha_i^* y_0^i - w_0, \alpha_i^*\mathbf{y}_i - \mathbf{w}) \\ &\geqslant \beta + \frac{1}{\epsilon}\{E(w_0 + \mathbf{w}'\mathbf{z} - \beta)^+ + E(-w_0 - \mathbf{w}'\mathbf{z})^+\} \\ &\geqslant \beta + \frac{1}{\epsilon}(-\beta)^+, \end{aligned}$$

where the second equality is due to $\alpha_k^* = 0$. Because $\epsilon \in (0, 1)$, the inequality $\beta + (1/\epsilon)(-\beta)^+ \leqslant 0$ is satisfied if and only if $\beta = 0$. We now argue that

$$\pi(y_0^i, \mathbf{y}_i) = 0 \quad \forall\, i \in \mathcal{H}(\mathbf{Y})\backslash\mathcal{G} \tag{27}$$

which, from Proposition 2.5, implies

$$y_0^i + \max_{\mathbf{z}\in\mathcal{W}} \mathbf{y}_i'\mathbf{z} \leqslant 0 \quad \forall\, i \in \mathcal{H}(\mathbf{Y})\backslash\mathcal{G}.$$

Indeed, for any $l \in \mathcal{H}(\mathbf{Y})\backslash\mathcal{G}$, we observe that

$$\begin{aligned}
0 &\geqslant \beta + \frac{1}{\epsilon}\left\{\pi(w_0-\beta,\mathbf{w}) + \sum_{i\in\mathcal{H}(\mathbf{Y})}\pi(\alpha_i^* y_0^i - w_0, \alpha_i^*\mathbf{y}_i - \mathbf{w})\right\}\\
&= \frac{1}{\epsilon}\left\{\pi(w_0,\mathbf{w}) + \sum_{i\in\mathcal{H}(\mathbf{Y})}\pi(\alpha_i^* y_0^i - w_0, \alpha_i^*\mathbf{y}_i - \mathbf{w})\right\}, \quad \text{substituting } \beta = 0,\\
&\geqslant \frac{1}{\epsilon}\{\pi(w_0,\mathbf{w}) + \pi(\alpha_l^* y_0^l - w_0, \alpha_l^*\mathbf{y}_l - \mathbf{w})\}\\
&\geqslant \frac{1}{\epsilon}\{\pi(\alpha_l^* y_0^l, \alpha_l^*\mathbf{y}_l)\}\\
&= \frac{\alpha_l^*}{\epsilon}\pi(y_0^l,\mathbf{y}_l) \geqslant 0.
\end{aligned}$$

Hence, equality (27) is achieved by noting that $\alpha_l^* > 0$. □

We propose an algorithm for improving the choice of $\boldsymbol{\alpha}$ and the set $\mathcal{J}$. Again, we assume that we can find an initial feasible solution of model (24).

Algorithm 3.4
Input: Y

Step 1. Solve problem (25) with Input $\mathbf{Y}$. Obtain optimal solution $\mathbf{s}^*$.

Step 2. Set $\mathcal{H}(\mathbf{Y}) := \{i \mid s_j^* > 0,\ j \in \mathcal{M}\}$.

Step 3. Solve problem (26) with Input $\mathbf{Y}$. Obtain optimal solution $\boldsymbol{\alpha}^*$. Set $\mathcal{J} := \mathcal{H}(\mathbf{Y})$.

Step 4. Solve model (24) with Input $(\boldsymbol{\alpha}, \mathcal{J})$. Obtain optimal solution $(\mathbf{x}^*, \mathbf{Y}^*)$. Set $\mathbf{Y} := \mathbf{Y}^*$.

Step 5. Repeat Steps 1–4 until a termination criterion is met.

Theorem 3.5. *In Algorithm* 3.4, *the sequence of objectives obtained by solving model* (24) *is nonincreasing. The algorithm will either have $\mathcal{J} = \varnothing$ or, consecutively, have $\mathcal{J}^k = \mathcal{J}^* \neq \varnothing$ for a certain index set $\mathcal{J}^*$ in a finite number of iterations. If in addition the set X is bounded, then the algorithm will produce a bounded infinite sequence $\{(x^k, \mathbf{Y}^k)\}$ with $c'x^k \downarrow \tau$, a certain limit.*

Proof. Starting with a feasible solution of model (24), we are assured that there exists $(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$, such that $\gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leqslant 0$. With Proposition 3.3(b), the condition in Step 3 ensures that $\alpha_j^* > 0$ for all $j \in \mathcal{J}$. Moreover, Proposition 3.3(a) ensures that the updates on $\boldsymbol{\alpha}$ and $\mathcal{J}$ do not affect the feasibility of its previous solution $(\mathbf{x}, \mathbf{Y})$ in model (24). Hence, its objective value will not increase.

Note that the set sequence $\{\mathcal{J}^k\}$ is nonexpanding, i.e., one has $\mathcal{J}^{k+1} \subseteq \mathcal{J}^k$ because $\mathcal{J}^{k+1} = \mathcal{H}(\mathbf{Y}^{k+1}) \subseteq \mathcal{J}^k$. Therefore, either $\mathcal{J}^k = \varnothing$ in a finite number of iterations, or there is a k_0 such that $\mathcal{J}^k = \mathcal{H}(\mathbf{Y}^k) \equiv \mathcal{J}^*$ for all $k \geqslant k_0$. In either case, the infinite sequence $\{(x^k, \mathbf{Y}^k)\}$ is bounded because X is bounded. This, together with the monotonicity of the sequence $\{c'x^k\}$, implies the last part of the theorem. □

The implementation of Algorithm 3.4 may involve perpetual updates of the set $\mathcal{J}$ and result in reformulating problem (24). A practical solution is to ignore the set $\mathcal{J}$ and solve the following model:

$$\begin{aligned}
Z_\epsilon^2(\boldsymbol{\alpha}) = \min\ & \mathbf{c}'\mathbf{x}\\
\text{s.t. } & \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{M}) \leqslant 0,\\
& (\mathbf{x}, \mathbf{Y}) \in X,
\end{aligned} \tag{28}$$

for a given $\boldsymbol{\alpha} \geqslant \mathbf{1}$ such that $\mathbf{1}'\boldsymbol{\alpha} = M$, where M is a large number. The updates of $\boldsymbol{\alpha}$ are done by solving

$$\begin{aligned}
Z_\alpha^2(\mathbf{Y}) = \min\ & \gamma_{1-\epsilon}(\mathbf{Y}, \boldsymbol{\alpha}, \mathcal{M})\\
\text{s.t. } & \sum_{j\in\mathcal{J}} \alpha_j = M,\\
& \boldsymbol{\alpha} \geqslant 1.
\end{aligned} \tag{29}$$

The algorithm is also simplified as follows.

Algorithm 3.5
Input: Y

Step 1. Solve problem (29) with Input $\mathbf{Y}$. Obtain optimal solution $\boldsymbol{\alpha}^*$. Set $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$

Step 2. Solve model (28) with Input $\boldsymbol{\alpha}$. Obtain optimal solution $(\mathbf{x}^*, \mathbf{Y}^*)$. Set $\mathbf{Y} = \mathbf{Y}^*$.

Step 3. Repeat Steps 1–2 until a termination criterion is met.

Step 4. Output solution $(\mathbf{x}^*, \mathbf{Y}^*)$.

The following result is straightforward.

Theorem 3.7. *Assume that* $\mathbf{Y}$ *is feasible in model* (28) *for some* $\boldsymbol{\alpha} \geqslant \mathbf{1}$ *and* $\mathbf{1}'\boldsymbol{\alpha} = M$. *Then, the sequence of objectives obtained by solving model* (24) *in Algorithm* 3.5 *is nonincreasing. If in addition the set X is bounded, then the algorithm will generate a bounded infinite sequence $\{(x^k, \mathbf{Y}^k)\}$ with $c'x^k \downarrow \tau$, a certain limit.*

Like most "Big M approaches," the quality of the solution improves with larger values of M. However, M cannot be so large that it results in numerical instability of the optimization problem. Although the Big M approach does not provide the theoretical guaranteed improvement over the classical approach using Bonferroni's inequality, it seems to perform very well from our numerical studies, as demonstrated in the next section.

4. Computational Studies

We analyze a resource allocation problem on a network with uncertain node demands and allowing transshipment of resources to neighboring nodes when necessary. We consider a directed graph with node set $\mathcal{V}$, $|\mathcal{V}| = n$ and arc set $\mathcal{E}$, $|\mathcal{E}| = r$. At each node i, $i \in \mathcal{V}$, we decide on the quantity of resource x_i to stock up, which will incur a cost of c_i per unit resource. When the demands $\tilde{d}_i$, $i \in \mathcal{V}$ are realized, resources at the nodes or from neighboring nodes are used to meet the demands. The goal is to minimize the total allocation cost subjected to a service-level constraint of meeting all demands with probability at least $1-\epsilon$. We assume that the resource at each node i can only be transshipped across to its outgoing neighboring nodes defined as

$$\mathcal{N}^-(i) \triangleq \{j\colon (i,j)\in\mathcal{E}\},$$

and received from its incoming neighboring nodes defined as

$$\mathcal{N}^+(i) \triangleq \{j\colon (j,i)\in\mathcal{E}\}.$$

Transshipment of resources received from other nodes is prohibited.

In our model, we ignore operating costs such as the transshipment costs. One of such applications is with regard to allocation of equipment such as ambulances or time-critical medical supplies for emergency response to local or neighboring demands. The costs associated with their procurement is more significant than the operating cost of transshipment, which may occur rather infrequently. We list the notations of the model as follows:

- c_i: unit cost of having one resource at node i, $i\in\mathcal{V}$;
- $d_i(\tilde{\mathbf{z}})$: demand at node i, $i\in\mathcal{V}$ as a function of the primitive uncertainties $\tilde{\mathbf{z}}$;
- x_i: quantity at resource at node i, $i\in\mathcal{V}$; and
- $w_{ij}(\tilde{\mathbf{z}})$: transshipment quantity from node i to node j, $(i,j)\in\mathcal{E}$ in respond to realization of $\tilde{\mathbf{z}}$.

The problem can be formulated as a joint chance-constrained problem as follows:

$$\begin{aligned}
\min\ & \mathbf{c}'\mathbf{x}\\
\text{s.t. }& \mathrm{P}\left(\begin{array}{l}
x_i+\sum_{j\in\mathcal{N}^+(i)} w_{ji}(\tilde{\mathbf{z}})-\sum_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\mathbf{z}})\geqslant d_i(\tilde{\mathbf{z}}),\\
\qquad\qquad i=1,\ldots,n,\\
x_i\geqslant \sum_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\mathbf{z}}),\quad i=1,\ldots,n,\\
\mathbf{w}(\tilde{\mathbf{z}})\geqslant \mathbf{0},
\end{array}\right)\\
&\geqslant 1-\epsilon,\\
&\mathbf{x}\geqslant\mathbf{0},\ \mathbf{w}(\tilde{\mathbf{z}}).
\end{aligned} \tag{30}$$

We assume that the demands are independently distributed and represented as

$$d_j(\tilde{\mathbf{z}}) = d_j^0+\tilde{z}_j,$$

where $\tilde{z}_j$ are independent zero mean random variables with unknown distribution.

By introducing new variables, we can transform model (30) to the "standard form" model as follows:

$$\begin{aligned}
\min\ & \mathbf{c}'\mathbf{x}\\
\text{s.t. }& x_i+\sum_{j\in\mathcal{N}^+(i)} w_{ji}(\tilde{\mathbf{z}})-\sum_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\mathbf{z}})+\mathbf{r}(\tilde{\mathbf{z}})=d_i(\tilde{\mathbf{z}}),\\
&\qquad\qquad i=1,\ldots,n,\\
& x_i+s_i(\tilde{\mathbf{z}})=\sum_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\mathbf{z}}),\quad i=1,\ldots,n,\\
& \mathbf{w}(\tilde{\mathbf{z}})+\mathbf{t}(\tilde{\mathbf{z}})=\mathbf{0},\\
& \mathbf{y}(\tilde{\mathbf{z}})=\begin{pmatrix}\mathbf{r}(\tilde{\mathbf{z}})\\ \mathbf{s}(\tilde{\mathbf{z}})\\ \mathbf{t}(\tilde{\mathbf{z}})\end{pmatrix},\\
& \mathrm{P}(\mathbf{y}(\tilde{\mathbf{z}})\leqslant\mathbf{0})\geqslant 1-\epsilon,\\
& \mathbf{x}\geqslant\mathbf{0},\mathbf{r}(\tilde{\mathbf{z}}),\ \mathbf{s}(\tilde{\mathbf{z}}),\ \mathbf{t}(\tilde{\mathbf{z}}),\ \mathbf{y}(\tilde{\mathbf{z}}),\ \mathbf{w}(\tilde{\mathbf{z}}).
\end{aligned} \tag{31}$$

Note that the dimension of $\mathbf{y}(\tilde{\mathbf{z}})$ is $m=2n+r$.

The transshipment variables $\mathbf{w}(\tilde{\mathbf{z}})$ are an arbitrary function of $\tilde{\mathbf{z}}$. To obtain a bound on problem (30), we apply the linear decision rule on the transshipment variables $\mathbf{w}(\tilde{\mathbf{z}})$ advocated in Ben-Tal et al. (2004) and Chen et al. (2007) as follows:

$$\mathbf{w}(\tilde{\mathbf{z}})=\mathbf{w}^0+\sum_{j=1}^n \mathbf{w}^j\tilde{z}_j.$$

Under the assumption of linear decision on $\mathbf{w}(\tilde{\mathbf{z}})$ and with suitable affine mapping, we have

$$\mathbf{r}(\tilde{\mathbf{z}})=\mathbf{r}^0+\sum_{j=1}^n \mathbf{r}^j\tilde{z}_j,$$

$$\mathbf{s}(\tilde{\mathbf{z}})=\mathbf{s}^0+\sum_{j=1}^n \mathbf{s}^j\tilde{z}_j,$$

$$\mathbf{t}(\tilde{\mathbf{z}})=\mathbf{t}^0+\sum_{j=1}^n \mathbf{t}^j\tilde{z}_j,$$

$$\mathbf{y}(\tilde{\mathbf{z}})=\mathbf{y}^0+\sum_{j=1}^n \mathbf{y}^j\tilde{z}_j,$$

which are affine functions with respect to the primitive uncertainty, $\tilde{\mathbf{z}}$. Hence, we transform the problem from one with infinite variables (optimizing over a functional) to a restricted one with a polynomial number of variables. Therefore, we can apply our proposed framework to obtain an approximate solution to problem (31).

The use of the linear decision rule is subject to criticism. As Shapiro and Nemirovski (2005, Remark 2) argued,

> The only reason for restricting ourselves with affine decision rules[3] stems from the desire to end up with a computationally tractable problem. We do not pretend that affine decision rules approximate well the optimal ones—whether it is so or not, it depends on the problem, and we usually have no possibility to understand how good in this respect is a particular problem we should solve. The rationale behind restricting to affine decision rules is the belief that in actual applications it is better to pose a modest and achievable goal rather than an ambitious goal which we do not know how to achieve.

Indeed, other than using the linear decision rule, we do not know of any other methods of addressing the joint chance-constrained problem with recourse and under incomplete distributional assumption.

In our test problem, we generate 15 nodes randomly positioned on a square grid and restrict to the r shortest arcs on the grid in terms of Euclidean distances. We assume that $c_i = 1$. For the demand uncertainty, we assume that $d_j^0 = 10$ and the demand at each node, $d_j(\tilde{\mathbf{z}})$, takes a value from 0 to 100. Therefore, we have $\tilde{z}_j \in [-10, 90]$. Using Theorem 2.2, we can determine the bounds on the forward and backward deviations, which are, respectively, $p_j = 42.67$ and $q_j = 30$.

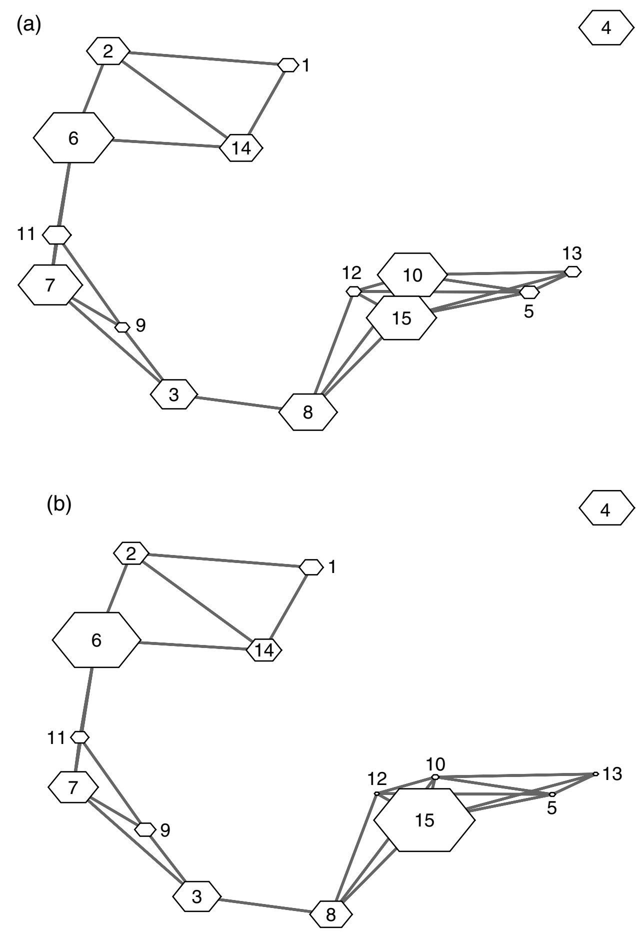
For the evaluation of bounds, we use $\mathscr{L} = \{1, 2, 4, 5\}$. We formulate the model using an in-house developed software, PROF (Platform for Robust Optimization Formulation). The Matlab-based software is essentially an SOCP modeling environment that contains reusable functions for modeling multiperiod robust optimization using decision rules. We have implemented bounds for the CVaR measure and expected positivity of a weighted sum of random variables. The software calls upon CPLEX 11.0 to solve the underlying SOCP.

In the computational experiment, we impose a service level of 99% or $\epsilon = 0.01$. We first solve the problem using the classical approach by decomposing the joint chance-constrained problem into m constraints of the form (18), with $\epsilon_i = \epsilon/(2n + r)$. We denote the optimal solution as $\mathbf{x}^B$ and its objective as Z^B. Subsequently, we use Algorithm 3.5, the big M approach, with $M = 10^6$, to improve upon the solution. We report results at the end of 20 iterations. Here we denote the optimal solution as $\mathbf{x}^N$ and its objective as Z^N. We also benchmark against the worst-case solution, which corresponds to all the demands at its maximum value. Hence, the worst-case solution is $x_i^W = 100$ for all $i \in \mathscr{V}$ and $Z^W = 1{,}500$.

Figure 1 illustrates the solution. The size of the hexagon on each location i, corresponds to the quantity x_i. Each link refers to two directed arcs in opposite directions. We present the solutions in Table 1. It is interesting to note that the solution obtained using the classical approach has significant resources allocated at nodes 5, 10, 12, and 13, which are all linked to node 15. After several iterations, the new solution centrally locates the resources at node 15, diminishing the requirements at nodes 5, 10, 12, and 13.

In Table 2, we compare the relative improvement of Z^N against Z^B and Z^N against Z^W. The new method has 8%–12% improvement compared with the classical approach of applying Bonferroni's inequality and has 30%–42% improvement compared with the worst-case solution. We also note that the improvement generally increases over the classical approach when the number of connectivity increases. This is probably due to the increase in correlation among the constraints as connectivity increases. Even though minimum distributional information is provided, this experiment shows that the new method solves the joint chance-constrained problem more efficiently.

Figure 1. Inventory allocation: 15 nodes, 50 arcs. (a) solution using Bonferroni's inequality; (b) solution using the new method.



We also evaluate the effectiveness of Algorithm 3.5, which may depend on the initial solution. We study the convergence of the algorithm with a random starting solution for a network with 15 nodes and 90 arcs (the last row in Table 2). We choose 100 sets of parameters ϵ_i at random in the simplex and solve the corresponding chance-constrained problems based on the Bonferroni inequality. For each solution, we apply Algorithm 3.5 and track the changes in objective values at every iteration. In Table 3, we present the distribution of the initial objective values as well as their values after completing 1, 5, 9, 13, and 18 iterations of Algorithm 3.5. The first column indicates the range within which the objective values fall. For example, at the end of Iteration 5, 19% of the solution has object

Table 1. Resource allocation: 15 nodes, 50 arcs (rounded to nearest integer).

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x^B	14	61	73	100	13	213	136	112	7	161	27	8	9	61	161
x^N	18	41	77	100	1	257	82	59	15	2	11	0	0	41	337

Table 2. Comparisons among worst-case solution Z^W, solution using Bonferroni's inequality Z^B, and solution using the new approximation Z^N.

No. of nodes	No. of arcs	Z^W	Z^B	Z^N	$(Z^W - Z^N)/Z^W$ (%)	$(Z^B - Z^N)/Z^B$ (%)
15	50	1,500	1,158.1	1,043.3	30.45	9.91
15	60	1,500	1,059.7	968.1	35.46	8.64
15	70	1,500	1,027.3	929.5	38.03	9.52
15	80	1,500	1,009.3	890.1	40.66	11.81
15	90	1,500	989.1	865.7	42.29	12.48

values in [905, 910). It is interesting to note that even with one iteration, Algorithm 3.5 is able to generate solutions that improve the best solution achieved by the Bonferroni inequality approach. At the end of 18 iterations, more than 75% of the solutions have objective values in [860, 870], which is at least a 10% improvement over the best solution obtained via the Bonferroni inequality.

5. Conclusion

We propose a general technique to deal with joint chance-constrained optimization problems. The standard approach decomposes the joint chance constraint into a problem with m individual chance constraints and then applies safe robust optimization approximation on each one of them. Our approach builds on a classical worst-case bound for the order statistics problem, where the bound is tight when the random variables are negatively correlated. By introducing new parameters $(\boldsymbol{\alpha}, w_0, \mathbf{w}, \mathcal{J})$ into the worst-case bound, we enlarge the search space so that our approach can also deal with positively correlated variables, and improves upon the solution obtained by using the standard approach via Bonferroni's inequality.

The quality of solution obtained by using this approach depends largely on the availability of a good upper bound $\pi(y_0, \mathbf{y})$ for the function $\mathrm{E}((y_0 + \mathbf{y}'\tilde{\mathbf{z}})^+)$. As a by-product of this study, we show that any such bound satisfying convexity, positively homogeneity, and with $\pi(y_0, \mathbf{0}) = y_0^+$, can be used to construct an uncertainty set to develop a robust optimization framework for (single) chance-constrained problems. This provides a unified perspective on the choice of uncertainty set in the development of robust optimization methodology.

Table 3. Distribution of objective values.

	Initial	Iter. 1	Iter. 5	Iter. 9	Iter. 13	Iter. 18
[995, 1,000)	1	0	0	0	0	0
[990, 995)	10	0	0	0	0	0
[985, 990)	27	0	0	0	0	0
[980, 985)	44	0	0	0	0	0
[975, 980)	14	0	0	0	0	0
[970, 975)	4	0	0	0	0	0
[965, 970)	0	1	0	0	0	0
[960, 965)	0	19	0	0	0	0
[955, 960)	0	21	1	1	0	0
[950, 955)	0	10	2	0	0	0
[945, 950)	0	11	3	0	1	1
[940, 945)	0	14	2	2	0	0
[935, 940)	0	12	0	0	0	0
[930, 935)	0	3	0	0	0	0
[925, 930)	0	2	0	0	0	0
[920, 925)	0	3	1	0	0	0
[915, 920)	0	2	1	0	1	0
[910, 915)	0	2	8	2	0	0
[905, 910)	0	0	19	6	4	3
[900, 905)	0	0	12	7	2	1
[895, 900)	0	0	24	7	6	4
[890, 895)	0	0	11	11	4	0
[885, 890)	0	0	11	15	8	4
[880, 885)	0	0	1	15	5	3
[875, 880)	0	0	4	5	6	3
[870, 875)	0	0	0	11	8	5
[865, 870)	0	0	0	18	55	56
[860, 865)	0	0	0	0	0	20

Endnotes

1. Take, for example, the joint chance constraint $\mathrm{P}(\tilde{\mathbf{a}}x \geqslant 1, \tilde{\mathbf{b}}y \geqslant 1) \geqslant 1 - \epsilon$, when $X = \{(x, y)\colon\ x \geqslant 1,\ \ y \geqslant 1\}$ and $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ are independent uniform distributions over [0, 1]. The approach using Bonferroni's inequality, with the decomposition $\epsilon_1 + \epsilon_2 = \epsilon$, reduces the feasible region to $\mathrm{P}(\tilde{\mathbf{a}}x \geqslant 1) \geqslant 1 - \epsilon_1$ and $\mathrm{P}(\tilde{\mathbf{b}}y \geqslant 1) \geqslant 1 - \epsilon_2$. The above holds only when $1/x \leqslant \epsilon_1$, $1/y \leqslant \epsilon_2$. Hence, the approach using Bonferroni's inequality reduces the set of feasible solutions to the region $\{(x, y)\colon 1/x + 1/y \leqslant \epsilon\}$. Note that the exact solution is $\{(x, y)\colon 1/x + 1/y - 1/(xy) \leqslant \epsilon\}$.
2. Although it is not explicitly stated in the theorem, it is assumed that $\mathcal{K}$ is a full-dimensional pointed cone in Ben-Tal and Nemirovski (2001). However, a classical theorem of Rockafellar (1970, Theorem 28.2) indicates that this assumption can be removed if $\mathcal{K}$ is defined by a finite number of convex inequalities, which is obviously the case considered in this paper.
3. An affine decision rule is equivalent to a linear decision rule in our context.

Acknowledgments

The authors thank the associate editor and two referees for their constructive comments. In particular, one referee suggested the setup for the last numerical experiment reported in §4. This research was supported by Singapore-MIT Alliance, NUS Risk Management Institute, and NUS academic research grant R-314-000-068-122.

References

Alizadeh, F., D. Goldfarb. 2003. Second-order cone programming. *Math. Programming* **95** 3–51.

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Math. Oper. Res.* **23** 769–805.

Ben-Tal, A., A. Nemirovski. 1999. Robust solutions to uncertain programs. *Oper. Res. Lett.* **25** 1–13.

Ben-Tal, A., A. Nemirovski. 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Math. Programming* **88** 411–424.

Ben-Tal, A., A. Nemirovski. 2001. Lectures on modern convex optimization: Analysis, algorithms, and engineering applications. *MPR-SIAM Series on Optimization*. SIAM, Philadelphia.

Ben-Tal, A., M. Teboulle. 1986. Expected utility, penalty functions and duality in stochastic nonlinear programming. *Management Sci.* **32** 1445–1466.

Ben-Tal, A., A. Goryashko, E. Guslitzer, A. Nemirovski. 2004. Adjusting robust solutions of uncertain linear programs. *Math. Programming* **99**(2) 351–376.

Bertsimas, D., D. B. Brown. 2009. Constructing uncertainty sets for robust linear optimization. *Oper. Res.*, ePub ahead of print April 24, http://or.journal.informs.org/cgi/content/abstract/opre.1080.0646v1.

Bertsimas, D., M. Sim. 2004. Price of robustness. *Oper. Res.* **52**(1) 35–53.

Bertsimas, D., D. Pachamanova, M. Sim. 2004. Robust linear optimization under general norms. *Oper. Res. Lett.* **32** 510–516.

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

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

Calafiore, G., L. El Ghaoui. 2006. Distributionally robust chance-constrained linear programs with applications. *J. Optim. Theory Appl.* **130**(1) 1–22.

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

Chen, W., M. Sim. 2009. Goal-driven optimization. *Oper. Res.* **57**(2) 342–357.

Chen, X., M. Sim, P. Sun. 2007. A robust optimization perspective on stochastic programming. *Oper. Res.* **55**(6) 1058–1107.

El Ghaoui, L., M. Oks, F. Oustry. 2003. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Oper. Res.* **51**(4) 543–556.

El Ghaoui, L., L. F. Oustry, H. Lebret. 1998. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.* **9** 33–52.

Ergodan, G., G. Iyengar. 2006. Ambiguous chance constrained problems and robust optimization. *Math. Programming, Ser. B* **107** 37–61.

Föllmer, H., A. Schied. 2004. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, Berlin.

Lagoa, C. M. 1999. On the convexity of probabilistically constrained linear programs. *Proc. 38th Conf. Decision and Control*, Vol. 1. Phoenix, AZ, 516–521.

Luedtke, J., S. Ahmed. 2008. A sample approximation approach for optimization with probabilistic constraints. *SIAM J. Optim.* **19**(2) 674–699.

Meilijson, I., A. Nadas. 1979. Convex majorization with an application to the length of critical path. *J. Appl. Probab.* **16** 671–677.

Natarajan, K., D. Pachmanova, M. Sim. 2009. Constructing risk measures from uncertainty sets. *Oper. Res.* **57**(5).

Nemirovski, A., A. Shapiro. 2006. Convex approximation of chance constrained programs. *SIAM J. Optim.* **17**(4) 969–996.

Pintér, J. 1989. Deterministic approximations of probability inequalities. *ZOR—Methods Models Oper. Res.* **33** 219–239.

Prékopa, A. 1995. *Stochastic Programming*. Kluwer, Dordrecht, The Netherlands.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.

Rockafellar, R. T., S. Uryasev. 2000. Optimization of conditional value-at-risk. *J. Risk* **2** 21–41.

Ruszczynski, A. 2002. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Math. Programming* **93**(2) 195–215.

Shapiro, A., A. Nemirovski. 2005. On complexity of stochastic programming problems. V. Jeyakumar, A. M. Rubinov, eds. *Continuous Optimization: Current Trends and Applications*. Springer, 111–144.

Soyster, A. L. 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Oper. Res.* **21** 1154–1157.