

Relaxations and Approximations of Chance Constraints under Finite Distributions

Shabbir Ahmed* and Weijun Xie†

May 10, 2018

Abstract

Optimization problems with constraints involving stochastic parameters that are required to be satisfied with a prespecified probability threshold arise in numerous applications. Such chance constrained optimization problems involve the dual challenges of stochasticity and nonconvexity. In the setting of a finite distribution of the stochastic parameters, an optimization problem with linear chance constraints can be formulated as a mixed integer linear program (MILP). The natural MILP formulation has a weak relaxation bound and is quite difficult to solve. In this paper, we review some recent results on improving the relaxation bounds and constructing approximate solutions for MILP formulations of chance constraints. We also discuss a recently introduced bicriteria approximation algorithm for covering type chance constrained problems. This algorithm uses a relaxation to construct a solution whose (constraint violation) risk level may be larger than the pre-specified threshold, but is within a constant factor of it, and whose objective value is also within a constant factor of the true optimal value. Finally, we present some new results that improve on the bicriteria approximation factors in the finite scenario setting and shed light on the effect of strong relaxations on the approximation ratios.

1 Introduction

We consider optimization problems with probabilistic or chance constraints of the form

$$v^* = \min_x \left\{ c^\top x : \mathbb{P} \left\{ \tilde{\xi} : x \in \mathcal{X}(\tilde{\xi}) \right\} \geq 1 - \epsilon \right\}, \quad (1)$$

where $x \in \mathbb{R}^n$ is a decision vector, $\tilde{\xi}$ is a random vector with support $\Xi \subseteq \mathbb{R}^d$ and probability distribution $\mathbb{P}$, and $\mathcal{X}(\tilde{\xi})$ denotes a system of stochastic constraints whose data is specified by the random vector $\tilde{\xi}$. Any deterministic constraints on x are assumed to be absorbed in the stochastic constraints. Problem (1) seeks a solution x that minimizes the cost $c^\top x$ and is required to satisfy the stochastic constraints $\mathcal{X}(\tilde{\xi})$ with probability at least $(1 - \epsilon)$, where $\epsilon \in (0, 1)$ is a prespecified risk level.

Since its introduction in 1958 [11], chance constrained optimization problems have been studied extensively (cf. [3, 33, 37]). Their applications can be found in numerous diverse areas such as finance [30], facility location [19], production [7, 48], telecommunication [13], healthcare [12], and

*School of Industrial & Systems Engineering, Georgia Institute of Technology

†Department of Industrial & Systems Engineering, Virginia Tech

energy systems [8, 38, 44, 46], among others. These problems are extremely challenging because of the difficulties in computing the probability of feasibility and in optimizing over their nonconvex feasible regions. In this paper our focus will be restricted to with finite distributions and linear chance constraints, i.e. we assume that

- (a) the underlying probability distribution is finite, i.e. $\tilde{\xi}$ has N realizations (or scenarios) $\{\xi^1, \dots, \xi^N\}$ with $p_i = \mathbb{P}\{\tilde{\xi} = \xi^i\}$ for $i \in [N] := \{1, 2, \dots, N\}$,
- (b) the uncertain constraint system corresponding to realization ξ^i , or scenario i in short, is a set of the form $\mathcal{X}(\xi^i) = \{x : A^i x \geq b^i\}$ for $i \in [N]$, and
- (c) the sets $\mathcal{X}(\xi^i)$ for all $i \in [N]$ are non-empty and have the same recession cone.

Even in this restricted setting, the chance constrained optimization problem (1) is strongly NP-hard [24, 34]. A natural mixed integer linear programming (MILP) formulation of (1) under the assumptions (a)-(c) is:

$$v^* = \min_{x, z} \left\{ c^\top x : \begin{array}{l} A^i x \geq z_i b^i - (1 - z_i) M_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ z \in \{0, 1\}^N \end{array} \right\}, \quad (2)$$

where, for each $i \in [N]$ and M_i is a nonnegative vector of the same dimension as b^i such that $A^i x \geq -M_i$ for all feasible x -solutions to (2). We assume that M_i exists. It is well known that this natural MILP formulation with big- M coefficients typically has a very weak linear programming (LP) relaxation. Consequently, there has been significant research on improving the relaxation bounds of (2). Another important line of research has been on approximations, i.e. efficiently constructing feasible solutions. This includes developing tractable convex inner approximations to the nonconvex feasible region of chance constraints and problem-dependent construction heuristics.

In this paper, we review some recent results on relaxations and approximations of chance constrained problems of the form (2). In particular, we discuss coefficient tightening, strong valid inequalities, and problem reformulations to improve the relaxation bounds. Then we discuss some existing approximation approaches for constructing good feasible solutions. Finally, we discuss a recently introduced bicriteria approximation algorithm for covering type chance constrained problems. This algorithm uses the solution to a convex relaxation to construct a solution whose (constraint violation) risk level may be larger than ϵ , but is within a constant factor of it, and whose objective value is within a constant factor of the true optimal value. We improve on the bicriteria approximation factors in the finite scenario setting and investigate the effect of strong relaxations on the approximation ratios.

Notation: All vectors and matrices are denoted in bold font, and their corresponding scalar components are in regular font. The vectors e and e^i denote the vector of all ones and the i -th basis vector, respectively. Their dimension will be clear from the context. For a vector x , the operations $\lfloor x \rfloor$, $\lceil x \rceil$, and $\max(x, 0)$, denotes its component-wise integer round-down, integer round-up, and positive part, respectively. The multiplication yx of a vector x with a scalar y is component-wise. For a natural number N , the set $[N] = \{1, 2, \dots, N\}$. The convex hull of a set S is denoted $\text{conv}(S)$. The set of feasible solutions of the MILP (2) is denoted by X . Additional notation will be introduced as needed.

2 Relaxations

2.1 Coefficient tightening

A straightforward attempt to improve the LP relaxation bound of (2) is by tightening the big- M coefficients. Specifically, a single row (ℓ -th row) of the constraint system corresponding to scenario i can be written as

$$(\mathbf{a}_\ell^i)^\top \mathbf{x} \geq b_\ell^i z_i - M_{i\ell}(1 - z_i)$$

where the row-specific big- M coefficient $M_{i\ell}$ satisfies

$$-M_{i\ell} \leq \inf\{(\mathbf{a}_\ell^i)^\top \mathbf{x} : \mathbf{x} \in X\}.$$

Of course computing these coefficients exactly could be as hard as solving the original problem, but it is valid to replace X by a relaxation in the above minimization problem. In particular, we can start out by very weak coefficients, improve them by minimizing $(\mathbf{a}_\ell^i)^\top \mathbf{x}$ over the current relaxation, and proceed recursively. This simple and well-known coefficient tightening idea has been effectively adapted to chance constrained problems of the form of (2) to dramatically improve their relaxation bounds (cf. [2, 31, 34, 40]).

2.2 Mixing inequalities

Next we discuss a family of valid inequalities for (2). Our presentation follows that of [47]. Let us assume that $\mathcal{X}(\xi^i) \subseteq S$ for all $i \in [N]$, for some compact set S . This holds, for example, when S is a deterministic constraint system inherited by each of the uncertain constraint systems. Given a vector $\boldsymbol{\alpha} \in \mathbb{R}^n$, let us define the following quantities

$$\beta_i^\alpha(S) := \inf\{\boldsymbol{\alpha}^\top \mathbf{x} : \mathbf{x} \in S \cap \mathcal{X}(\xi^i)\} \quad \forall i \in [N]. \quad (3)$$

Note that we have allowed the above definition to depend on S so that we can account for successive applications with changing S (in Section 2.3). The $(1 - \epsilon)$ -quantile associated with these quantities is

$$\beta_q^\alpha(S) := \min \left\{ \eta : \begin{array}{l} \eta \geq \beta_i^\alpha(S) z_i - K(1 - z_i) \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}, \quad (4)$$

where $K = \min_{i \in [N]} \beta_i^\alpha(S)$. When $p_i = 1/N$ for all $i \in [N]$, then $\beta_q^\alpha(S)$ can be computed by first sorting the values $\{\beta_i^\alpha(S)\}_{i \in [N]}$ in nondecreasing order, and then picking the the $(\lfloor \epsilon N \rfloor + 1)$ largest value.

Now consider the set

$$Y^\alpha = \left\{ (\mathbf{x}, \mathbf{z}) : \begin{array}{l} \boldsymbol{\alpha}^\top \mathbf{x} + [\beta_i^\alpha(S) - \beta_q^\alpha(S)](1 - z_i) \geq \beta_i^\alpha(S) \quad \forall i \in N^\alpha \\ \boldsymbol{\alpha}^\top \mathbf{x} \geq \beta_q^\alpha(S) \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}, \quad (5)$$

where $N^\alpha := \{i \in [N] : \beta_i^\alpha(S) \geq \beta_q^\alpha(S)\}$. From the definitions of $\beta_i^\alpha(S)$ and $\beta_q^\alpha(S)$, it follows that $Y^\alpha \supseteq X$, i.e., the mixed-integer set Y^α is a relaxation of the chance constrained system given by (2) [22]. Thus any inequality that is valid for Y^α is valid for X . The set Y^α is an example of a *mixing set* which has been extensively studied in varying degrees of generality in the MILP literature (cf. [5, 16, 17, 24, 25]). A complete polyhedral description of the convex hull of this set, i.e. $\text{conv}(Y^\alpha)$, is given by the following result.

Theorem 1 [5, 17, 24] *Given $\{i_1, i_2, \dots, i_J\} \subseteq N^\alpha$ such that $\beta_{i_1}^\alpha(S) \geq \beta_{i_2}^\alpha(S) \geq \dots \geq \beta_{i_J}^\alpha(S)$, consider the mixing inequality*

$$\alpha^\top x + \sum_{j=1}^J [\beta_{i_j}^\alpha(S) - \beta_{i_{j+1}}^\alpha(S)](1 - z_{i_j}) \geq \beta_{i_1}^\alpha(S) \quad (6)$$

where $\beta_{i_{J+1}}^\alpha(S) := \beta_q^\alpha(S)$.

- (a) *The mixing inequality (6) is valid for $\text{conv}(Y^\alpha)$.*
- (b) *If $i_1 = \arg\max\{\beta_i^\alpha(S) : i \in N^\alpha\}$ then the mixing inequality (6) defines a facet of $\text{conv}(Y^\alpha)$.*
- (c) *The set of all mixing inequalities together with the bounds on the binary variables are sufficient to describe $\text{conv}(Y^\alpha)$.*
- (d) *A violated mixing inequality can be separated in time that is polynomial in $|N^\alpha|$.*

A key issue in finding mixing inequalities for the chance constraint set X is the choice of the α -vector to use in the mixing relaxation Y^α . If the uncertainties are only on the right-hand-side, i.e. the constraint system for scenario i is of the form $Ax \geq b^i$ for a fixed A , then it is sufficient to use the rows of A as the α -vectors. Mixing inequalities derived in this manner have been successfully used in branch-and-cut approaches for chance constrained problems with only right-hand-side uncertainties [18, 24]. For general chance constraint systems, to find a violated mixing inequality we need to find the corresponding α -vector. Unfortunately, this separation is NP-hard [47]. In [22], heuristic approaches to separate α -vectors and corresponding mixing inequalities are used within a branch-and-cut scheme to effectively solve general chance constrained problems. This scheme has been extended to two-stage chance constrained problems in [21]. There has also been work on combining multiple mixing inequalities to get a strong inequality [18, 20, 49]. Note that the mixing relaxation Y^α ignores the system

$$Z = \left\{ z \in \{0, 1\}^N : \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \right\} \quad (7)$$

that is inherent in X . When the probabilities are uniform, i.e. $p_i = 1/N$, [24] propose strengthened inequalities for $\text{conv}(Y^\alpha \cap (\mathbb{R}^n \times Z))$. For general probabilities the set Z is a knapsack system and is hard to optimize/separate over. In [1], the authors extend mixing inequalities to account for the knapsack constraint.

2.3 Quantile cuts and their closure

It can be shown that [34], for any $\alpha \in \mathbb{R}^n$,

$$\text{Proj}_x(\text{conv}(Y^\alpha \cap (\mathbb{R}^n \times Z))) = \{x : \alpha^\top x \geq \beta_q^\alpha(S)\}$$

where $\text{Proj}_x(\cdot)$ denotes projection on to the x -variables. Thus the inequality $\alpha^\top x \geq \beta_q^\alpha(S)$, called a *quantile cut* [47], is a valid inequality for $\text{conv}(\text{Proj}_x(X))$, and captures the effect of all inequalities derived from the mixing system Y^α and the knapsack system Z with respect to the x -variables. Of course, the quantile cut depends on the choice of the α vector. In the recent paper [47], we

study the closure of all possible quantile cuts, i.e. the intersection of quantile cuts for all $\alpha \in \mathbb{R}^n$, for a chance constrained system. We briefly review these results next.

Recall that S is a compact set such that $\mathcal{X}(\xi^i) \subseteq S$ for all $i \in [N]$, and the β -values are computed with respect to this set as in (3) and (4). The first quantile closure is defined as

$$S^1 = \bigcap_{\alpha \in \mathbb{R}^n} \{x : \alpha^\top x \geq \beta_q^\alpha(S)\}. \quad (8)$$

Since $\text{Proj}_x(X) \subseteq S^1$, we can update each scenario system such that $\mathcal{X}(\xi^i) \subseteq S^1$ for each i . Recall from (3) and (4) that this effects the β values. We can then recursively define the r -th quantile closure as

$$S^r = \bigcap_{\alpha \in \mathbb{R}^n} \{x : \alpha^\top x \geq \beta_q^\alpha(S^{r-1})\} \text{ for } r > 1. \quad (9)$$

A natural theoretical question is whether a successive application of the quantile cuts in the above manner can eventually recover the convex hull of the chance constrained system.

Theorem 2 [47] *Assuming that the initial set S is compact, successive quantile closures converge to the convex hull of the chance constrained set (in the space of the x -variables) with respect to Hausdorff distance, i.e*

$$\lim_{r \rightarrow \infty} d_H(S^r, \text{conv}(\text{Proj}_x(X))) = 0$$

where $d_H(A, B)$ is the Hausdorff distance between two closed convex sets A and B (cf. [36]).

The above result establishes that the quantile cuts, and therefore the mixing inequalities, are essentially all that are needed to describe the convex hull of the chance constrained set (in the space of the x -variables). Unfortunately, as mentioned earlier, separating a violated quantile cut, i.e determining a α -vector, is NP-hard.

2.4 Nonanticipative relaxations

In this section, we discuss relaxations obtained by a specific dualization of problem (2), and associated extended formulations derived from them. The material in this section is from [2].

Let us reformulate the chance constrained problem (2) into an equivalent form by making N copies of the variables x , denoted as $\{x^i\}_{i \in [N]}$, and enforcing them to be equal, i.e., $x^1 = \dots = x^N$. These constraints are known as “nonanticipativity constraints” in the stochastic programming literature. For notational convenience, we let $\sum_{i \in [N]} H_i x^i = h$ denote the system of nonanticipativity constraints. Therefore, (2) is equivalent to

$$v^* = \min_{x, z} \left\{ \sum_{i \in [N]} p_i c^\top x^i : \begin{array}{l} \sum_{i \in [N]} H_i x^i = h \\ A^i x^i \geq z_i b^i - (1 - z_i) M_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ z \in \{0, 1\}^N \end{array} \right\}. \quad (10)$$

Relaxing the nonanticipativity constraints with Lagrangian multipliers λ , we obtain the following dual problem

$$v_1^{LD} = \sup_{\lambda} \left\{ \mathcal{L}_1(\lambda) - \lambda^\top h \right\} \quad (11)$$

where

$$\mathcal{L}_1(\boldsymbol{\lambda}) = \min_{\mathbf{x}, \mathbf{z}} \left\{ \sum_{i \in [N]} \left(p_i \mathbf{c}^\top \mathbf{x}^i + \boldsymbol{\lambda}^\top \mathbf{H}_i \mathbf{x}^i \right) : \begin{array}{l} \mathbf{A}^i \mathbf{x}^i \geq z_i \mathbf{b}^i - (1 - z_i) \mathbf{M}_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (12)$$

It can be shown that, for a given vector $\boldsymbol{\lambda}$, the optimization problem (12) to evaluate $\mathcal{L}_1(\boldsymbol{\lambda})$ can be decomposed over the N scenarios and solved efficiently (see [2]). Thus we can obtain a lower bound on v^* . The optimal dual bound v_1^{LD} can be obtained by solving the convex problem (11) with standard techniques such as the subgradient method.

Next we discuss a second dual problem. Note that in (10), $\max_{i \in [N]} \mathbf{c}^\top \mathbf{x}^i = \sum_{i \in [N]} p_i \mathbf{c}^\top \mathbf{x}^i$ holds for any feasible solution satisfying the nonanticipativity constraints. Thus, we can replace the objective function in (10) with $\max_{i \in [N]} \mathbf{c}^\top \mathbf{x}^i$, which can be further linearized by introducing an auxiliary variable y to obtain the following formulation

$$v^* = \min_{y, \mathbf{x}, \mathbf{z}} \left\{ y : \begin{array}{l} \mathbf{c}^\top \mathbf{x}^i \leq y \quad \forall i \in [N] \\ \sum_{i \in [N]} \mathbf{H}_i \mathbf{x}^i = \mathbf{h} \\ \mathbf{A}^i \mathbf{x}^i \geq z_i \mathbf{b}^i - (1 - z_i) \mathbf{M}_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (13)$$

As before, by relaxing the nonanticipativity constraints we can construct the following dual problem to get a lower bound on v^* :

$$v_2^{LD} = \min_y \sup_{\boldsymbol{\lambda}} \left\{ y + \mathcal{L}_2(\boldsymbol{\lambda}, y) - \boldsymbol{\lambda}^\top \mathbf{h} \right\} \quad (14)$$

where

$$\mathcal{L}_2(\boldsymbol{\lambda}, y) = \min_{\mathbf{x}, \mathbf{z}} \left\{ \sum_{i \in [N]} \boldsymbol{\lambda}^\top \mathbf{H}_i \mathbf{x}^i : \begin{array}{l} \mathbf{c}^\top \mathbf{x}^i \leq y \quad \forall i \in [N] \\ \mathbf{A}^i \mathbf{x}^i \geq z_i \mathbf{b}^i - (1 - z_i) \mathbf{M}_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (15)$$

For fixed y and $\boldsymbol{\lambda}$, problem (15) can be decomposed over the N scenarios and solved efficiently. The dual problem (14) can be solved by combining bisection to optimize over y with standard subgradient methods to optimize over $\boldsymbol{\lambda}$ (see [2] for details).

We can show the following relationship between the above derived bounds and that obtained from the linear programming relaxation of (2).

Theorem 3 [2] *Let v^M denote the optimal value of the LP relaxation of (2) (note that this value depends on the big- M coefficients used, hence the superscript). Then*

$$v_2^{LD} \geq v_1^{LD} \geq v^M.$$

Thus the bounds obtained from the nonanticipative relaxations are at least as good as that of the LP relaxation even with the best possible big- M coefficients. Using primal characterization of the Lagrangian dual of an MILP and disjunctive programming techniques, we can use the above Lagrangian duals to derive reformulations of (2) with strong relaxations bounds.

Theorem 4 [2] *The chance constrained optimization problem (2) can be reformulated as the following big- M -free MILP:*

$$\min_{\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{w}} \left\{ \mathbf{c}^\top \mathbf{x} : \begin{array}{l} \mathbf{u}^i + \mathbf{w}^i = \mathbf{x} \quad \forall i \in [N] \\ \mathbf{A}^i \mathbf{u}^i \geq z_i \mathbf{b}^i \quad \forall i \in [N] \\ \mathbf{A}^i \mathbf{w}^i \geq (1 - z_i) \mathbf{b}^i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (16)$$

The optimal value of the LP relaxation of the above formulation is at least as large as that of the LP relaxation of (2). If $p_i = 1/N$ for all $i \in [N]$ then the optimal value of the LP relaxation of the above formulation is equal to v_1^{LD} .

Theorem 5 [2] *The chance constrained optimization problem (2) can be reformulated as the following big- M -free mixed-integer nonlinear program (MINLP):*

$$\min_{\mathbf{x}, y, \mathbf{z}} \left\{ y : \begin{array}{l} \mathbf{c}^\top \mathbf{u}^i \leq y z_i \quad \forall i \in [N] \\ \mathbf{c}^\top \mathbf{w}^i \leq y(1 - z_i) \quad \forall i \in [N] \\ \mathbf{u}^i + \mathbf{w}^i = \mathbf{x} \quad \forall i \in [N] \\ \mathbf{A}^i \mathbf{u}^i \geq z_i \mathbf{b}^i \quad \forall i \in [N] \\ \mathbf{A}^i \mathbf{w}^i \geq (1 - z_i) \mathbf{b}^i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (17)$$

The optimal value of the continuous relaxation of the above formulation (i.e. when $\mathbf{z} \in [0, 1]^N$) is at least as large as that of the LP relaxation of (2). If $p_i = 1/N$ for all $i \in [N]$ then the optimal value of the continuous relaxation of the above formulation is equal to v_2^{LD} .

The continuous relaxation of (17) is a nonlinear program, and [2] propose a successive linear programming technique and a successive second-order cone programming (SOCP) technique for its solution. These schemes also lead to an MILP reformulation and a mixed integer second-order cone programming (MISOCP) reformulation of (2).

3 Approximations

Motivated by the difficulty of chance constrained optimization problems, there has been a great deal of elegant work in developing tractable methods to construct a good feasible solution. These include designing a deterministic safe convex approximation, i.e. a tractable convex set that is an inner approximation to the chance constraint set, such as robust optimization [6], conditional-value-at-risk (CVaR) approximation [28, 35], Bernstein approximation [28], and related approaches [32]. Another line of work has been on scenario approximations [9, 10, 23, 27], where the chance constraint is replaced by a sampled set of constraints, and the sample size is chosen to guarantee that a solution to the sampled system is feasible to the chance constrained problem with high probability. See [26] for a survey of safe and scenario approximations of chance constraints. While these approximation methods have guarantees on feasibility, they do not have any guarantee on optimality. That is, the approximate solutions could be arbitrarily far from being optimal. To the

best of our knowledge, all existing approximation algorithms with provable optimality guarantees have been studied for chance constrained *combinatorial* optimization problems, i.e. where the decision vector $\mathbf{x}$ is restricted to be binary (cf. [14, 15, 41]).

In this section we review the CVaR based safe approximation and the scenario approximation approach. We also discuss a recently proposed bisection heuristic [2]. For each of these approximations, we provide examples to show that the optimality guarantees could be arbitrarily bad. The worst-case examples for CVaR and scenario approximations are from [45], while that for the bisection heuristic of [2] is new.

3.1 CVaR approximation

For a scalar continuous random variable W with distribution $\mathbb{P}$ and some $\alpha \in (0, 1)$, the value-at-risk (VaR) (at level α) is the quantile defined as

$$\text{VaR}_\alpha(W) = \inf\{t : \mathbb{P}\{W \leq t\} \geq \alpha\},$$

and the corresponding conditional-value-at-risk (CVaR) is

$$\text{CVaR}_\alpha(W) = \mathbb{E}[W | W \geq \text{VaR}_\alpha(W)].$$

For notational convenience, let us use v to denote $\text{VaR}_\alpha(W)$. Notice that

$$\begin{aligned} \text{CVaR}_\alpha(W) &= \mathbb{E}[W | W \geq v] \\ &= \left(\int_{W \geq v} W d\mathbb{P} \right) / \left(\int_{W \geq v} d\mathbb{P} \right) \\ &= \frac{1}{1-\alpha} \left(\int_{W \geq v} (W + v - v) d\mathbb{P} \right) \\ &= \frac{1}{1-\alpha} \left(\int_{W \geq v} (W - v) d\mathbb{P} + \int_{W \geq v} v d\mathbb{P} \right) \\ &= v + \frac{1}{1-\alpha} \mathbb{E}[(W - v)_+], \end{aligned}$$

where $(a)_+ = \max\{0, a\}$. Thus

$$\text{CVaR}_\alpha(W) = \text{VaR}_\alpha(W) + \frac{1}{1-\alpha} \mathbb{E}[(W - \text{VaR}_\alpha(W))_+].$$

Using the above construction and standard convex analysis, [35] show the following important representation theorem for the CVaR of a continuous random variable W ,

$$\text{CVaR}_\alpha(W) = \inf_y \left\{ y + \frac{1}{1-\alpha} \mathbb{E}[(W - y)_+] \right\}. \quad (18)$$

Since $\text{CVaR}_\alpha(W) \geq \text{VaR}_\alpha(W)$, it is easy to see that for a continuous random variable

$$\begin{aligned} \mathbb{P}\{W \leq 0\} \geq \alpha &\Leftrightarrow \text{VaR}_\alpha(W) \leq 0 \\ &\Leftrightarrow \text{CVaR}_\alpha(W) \leq 0 \\ &\Leftrightarrow \exists y : y + \frac{1}{1-\alpha} \mathbb{E}[(W - y)_+] \leq 0. \end{aligned} \quad (19)$$

Now consider the chance constraint

$$\mathbb{P}\{\tilde{\mathbf{A}}\mathbf{x} \geq \tilde{\mathbf{b}}\} \geq 1 - \epsilon \quad (20)$$

where the random quantities $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$ have a continuous joint distribution. The chance constraint (20) can be written as

$$\mathbb{P}\{W(\mathbf{x}) \leq 0\} \geq 1 - \epsilon,$$

where $W(\mathbf{x}) = \|(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x})_+\|_\infty$ is a scalar random variable with a continuous distribution. Using the construction in (19) we see that the following constraint on the variables $\mathbf{x}$ and the new variable y implies the nonconvex chance constraint (20):

$$y + \frac{1}{\epsilon} \mathbb{E} \left[(\|(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x})_+\|_\infty - y)_+ \right] \leq 0. \quad (21)$$

Note that $\|(\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x})_+\|_\infty$ is convex in $\mathbf{x}$, and $(\cdot)_+$ is convex and nondecreasing, thus above constraint is convex in $\mathbf{x}$ and y . Thus we have a safe convex approximation to the chance constraint (20).

Note that the above construction works when the random variables are continuously distributed. The following result shows that it is also valid for the chance constrained optimization problem (2) with discrete random variables.

Proposition 1 *Let $\hat{\mathbf{x}}$ be a solution to the following linear program*

$$\min_{\mathbf{x}, y, \theta_1, \dots, \theta_N} \left\{ \begin{array}{ll} y + \frac{1}{\epsilon} \sum_{i \in [N]} p_i \theta_i \leq 0 \\ \mathbf{c}^\top \mathbf{x} : y + \theta_i \geq 0 & \forall i \in [N] \\ (y + \theta_i) \mathbf{e} \geq \mathbf{b}^i - \mathbf{A}^i \mathbf{x} & \forall i \in [N] \\ \theta_i \geq 0 & \forall i \in [N] \end{array} \right\} \quad (22)$$

then $\hat{\mathbf{x}}$ is a feasible solution of (2). (Note that we have introduced the θ variables to linearize the nonlinear operations $(\cdot)_+$ and $\|(\cdot)_+\|_\infty$.)

Proof: Let $(\hat{\mathbf{x}}, \hat{y}, \hat{\theta}_1, \dots, \hat{\theta}_N)$ be a solution to the linear program. We need to construct a solution $(\hat{z}_1, \dots, \hat{z}_N)$ such that $(\hat{\mathbf{x}}, \hat{z}_1, \dots, \hat{z}_N)$ is feasible to (2). Let

$$N_1 = \{i \in [N] : \mathbf{A}^i \hat{\mathbf{x}} \geq \mathbf{b}^i\} \text{ and } N_2 = \{i \in [N] : \mathbf{A}^i \hat{\mathbf{x}} \not\geq \mathbf{b}^i\}.$$

Let $\hat{z}_i = 1$ for $i \in N_1$ and $\hat{z}_i = 0$ for $i \in N_2$. Then the solution $(\hat{\mathbf{x}}, \hat{z}_1, \dots, \hat{z}_N)$ satisfies the constraints

$$\begin{aligned} \mathbf{A}^i \mathbf{x} &\geq z_i \mathbf{b}^i - (1 - z_i) \mathbf{M}_i \quad \forall i \in [N] \\ \mathbf{z} &\in \{0, 1\}^N. \end{aligned}$$

We only need to show that it satisfies $\sum_{i \in [N]} p_i z_i \geq 1 - \epsilon$. Consider the first constraint of the linear program (22):

$$\begin{aligned} 0 &\geq \hat{y} + \frac{1}{\epsilon} \sum_{i \in N_1} p_i \hat{\theta}_i + \frac{1}{\epsilon} \sum_{i \in N_2} p_i \hat{\theta}_i \\ &\geq \hat{y} + \frac{1}{\epsilon} \sum_{i \in N_2} p_i \hat{\theta}_i \\ &\geq \hat{y} + \frac{1}{\epsilon} \sum_{i \in N_2} p_i [\|(\mathbf{b}^i - \mathbf{A}^i \hat{\mathbf{x}})_+\|_\infty - \hat{y}], \end{aligned}$$

where the second inequality is due to $\hat{\theta}_i \geq 0$ for all $i \in N_1$ and the third inequality is because $\hat{y} + \hat{\theta}_i \geq \|(\mathbf{b}^i - \mathbf{A}^i \hat{\mathbf{x}})_+\|_\infty$ for all $i \in [N]$.

Note that $\|(\mathbf{b}^i - \mathbf{A}^i \hat{\mathbf{x}})_+\|_\infty > 0$ for all $i \in N_2$, and let $C = \min_{i \in N_2} \|(\mathbf{b}^i - \mathbf{A}^i \hat{\mathbf{x}})_+\|_\infty > 0$. Then

$$\sum_{i \in N_2} p_i \leq \frac{-\hat{y}}{C - \hat{y}} \epsilon \leq \epsilon$$

since $\hat{y} \leq 0$ from the first constraint of (22). Thus

$$\sum_{i \in [N]} p_i \hat{z}_i = \sum_{i \in N_1} p_i = 1 - \sum_{i \in N_2} p_i \geq 1 - \epsilon.$$

Thus $(\hat{\mathbf{x}}, \hat{z}_1, \dots, \hat{z}_N)$ satisfies all the constraints of (2), hence $\hat{\mathbf{x}}$ is a feasible solution. $\square$

The above CVaR approximation is a widely used convex approximation of chance constraints, and can be shown to be in some sense the best one (see [28] for details). Unfortunately, the CVaR approximation can be overly conservative (cf. [4]), and the following example shows that this approximation does not have any optimality guarantee.

Example 1 [45] *Given a parameter $\kappa > 1$, consider the following instance of the chance constrained problem (2):*

$$(P_\kappa) : \quad v_\kappa^* = \min_{x \geq 0} \left\{ x : \begin{array}{l} \xi_i x \geq z_i - M_i(1 - z_i) \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ z \in \{0, 1\}^N. \end{array} \right\},$$

where ϵN is an integer, $p_i = 1/N$ for all $i \in [N]$, $\{\xi_i\}_{i \in [N]}$ is such that $\xi_i = 1$ for $i \in [\epsilon N]$ and $\xi_i = \kappa$ for $i \in [N] \setminus [\epsilon N]$. Note that for this problem we can set $M_i = 0$ for all $i \in [N]$. Let v_κ^{CVaR} be the optimal value of the CVaR approximation (22) of (P_κ) . Then it can be shown (see [45]) that

$$\lim_{\kappa \rightarrow \infty} \frac{v_\kappa^{\text{CVaR}}}{v_\kappa^*} = \infty.$$

3.2 Scenario approximation

Consider the chance constrained optimization problem under general distributions

$$v^* = \min_{\mathbf{x}} \left\{ \mathbf{c}^\top \mathbf{x} : \mathbb{P}\{\tilde{\mathbf{A}}\mathbf{x} \geq \tilde{\mathbf{b}}\} \geq 1 - \epsilon \right\}. \quad (23)$$

Recall that (2) is a special case of (23) with a finite distribution. The scenario approximation (SA) approach for (23), proposed by [10], uses $\bar{N}$ i.i.d. samples of $(\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$ and considers the linear program where the constraints corresponding to each sampled scenario are enforced:

$$v^{SA} = \min_{\mathbf{x}} \left\{ \mathbf{c}^\top \mathbf{x} : \mathbf{A}^k \mathbf{x} \geq \mathbf{b}^k \quad \forall k \in [\bar{N}] \right\}. \quad (24)$$

In [10] it is shown that when the sample size $\bar{N}$ satisfies

$$\bar{N} \geq \left\lceil \frac{2}{\epsilon} \log \left(\frac{1}{\delta} \right) + \frac{2n}{\epsilon} \log \left(\frac{2}{\epsilon} \right) + 2n \right\rceil.$$

If the approximate problem (24) is feasible then with probability at least $(1 - \delta)$, where $\delta \in (0, 1)$ is a given confidence level, an optimal solution to (24) is feasible to the chance constrained problem (23), thus, $v^{SA} \geq v^*$. The following example illustrates that approximation quality from the SA approach can be arbitrarily bad by showing that (24) can be infeasible with high probability.

Example 2 [45] Consider the following instance of the chance constrained problem (23):

$$(P) : \quad v^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\{\tilde{\xi}x \geq 1\} \geq 1 - \epsilon \right\},$$

where $\tilde{\xi}$ is a Bernoulli random variable with probability $(1 - \epsilon)$. Let v_δ^{SA} denote the optimal value of the scenario approximation problem (24) corresponding to (P) for a given confidence parameter $\delta \in (0, 1)$ and sample size $\bar{N}_\delta = \left\lceil \frac{2}{\epsilon} \log\left(\frac{1}{\delta}\right) + \frac{2}{\epsilon} \log\left(\frac{2}{\epsilon}\right) + 2 \right\rceil$. Then it can be shown (see [45]) that $v^* = 1$ and

$$\mathbb{P} \left\{ \frac{v_\delta^{SA}}{v^*} = \infty \right\} \geq 1 - \delta^2 \epsilon^2,$$

i.e., with probability at least $1 - \delta^2 \epsilon^2$, a scenario approximation of (P) is infeasible.

3.3 Bisection heuristic of [2]

In [2], the authors propose a heuristic for constructing a feasible solution to the chance constrained problem (2). The key idea of this heuristic is to minimize the sum of infeasibilities for each scenario while bounding the objective value by a prescribed amount by solving the following linear program

$$\min_{\mathbf{x}, \mathbf{s} \in \mathbb{R}_+^N} \left\{ \sum_{i \in [N]} p_i s_i : \begin{array}{l} \mathbf{A}^i \mathbf{x} \geq \mathbf{b}^i - s_i \quad i \in [N], \\ \mathbf{c}^\top \mathbf{x} \leq v \end{array} \right\} \quad (25)$$

where v is a provided bound on the objective value. The heuristic searches over v (by bisection) in order to find a feasible solution. The procedure is as follows.

Algorithm 1 Bisection heuristic of [2]

- 1: Let $L > -\infty$ and $U < \infty$ be known lower and upper bounds on the optimal value of (2), let $\hat{\delta} > 0$ be a stopping tolerance parameter.
 - 2: **while** $U - L > \hat{\delta}$ **do**
 - 3: $v \leftarrow (L + U)/2$.
 - 4: Let $(\hat{\mathbf{x}}, \hat{\mathbf{s}})$ be an optimal solution of (25) and set $\hat{z}_i = \mathbb{I}(\hat{s}_i = 0)$ for all $i \in [N]$.
 - 5: **if** $\sum_{i=1}^N p_i \hat{z}_i \geq 1 - \epsilon$ **then**
 - 6: $U \leftarrow v$.
 - 7: **else**
 - 8: $L \leftarrow v$.
 - 9: **end if**
 - 10: **end while**
 - 11: **Output** $v^{heur} \leftarrow U$.
-

In [2], the above heuristic is proven to be exact for some restricted classes of chance constrained problems, and is shown to perform very well in numerical experiments with more general classes of problems. Here we show that, in general, the heuristic solution can be quite far from being optimal.

Example 3 Consider the following instance of (2)

$$v^* = \min_{\mathbf{x} \in \mathbb{R}_+^2} \left\{ \frac{1}{1-\epsilon} x_1 + x_2 : \begin{array}{l} (\boldsymbol{\xi}^i)^\top \mathbf{x} \geq z_i - (1-z_i)M_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N. \end{array} \right\} \quad (26)$$

where $\epsilon = (N-1)/N$, $\boldsymbol{\xi}^i = (1, 0)^\top$ for $i \in [\epsilon N]$, $\boldsymbol{\xi}^i = (1, 1)^\top$ for $i \in [N] \setminus [\epsilon N]$, and $p_i = 1/N$ for all $i \in [N]$, and $M_i = 0$. The following result shows that objective value of the heuristic solution for this problem is a factor N of the optimal value.

Proposition 2 Let v^* be the optimal value of the chance constrained optimization problem (26) and let v^{heur} be the optimal value returned by the value found by the bisection heuristic (Algorithm 1) applied to (26). Then $v^{heur}/v^* \geq N$.

Proof: It is easily verified that the optimal value $v^* = 1$. Problem (25) corresponding to (26) is

$$\min_{(\mathbf{x}, \mathbf{s}) \in \mathbb{R}_+^{N+2}} \left\{ (1/N) \sum_{i \in [N]} s_i : \begin{array}{ll} x_1 \geq 1 - s_i & i \in [\epsilon N] \\ x_1 + x_2 \geq 1 - s_i & i \in [N] \setminus [\epsilon N] \\ \frac{1}{1-\epsilon} x_1 + x_2 \leq v \end{array} \right\} \quad (27)$$

Suppose in Algorithm 1, we start with any $U > \frac{1}{1-\epsilon}$ and $L = 0$. Then, for any $v \in [1, \frac{1}{1-\epsilon})$, we must have $s_i = 1 - (1-\epsilon)v > 0$ for each $i \in [N]$ and $x_1 = (1-\epsilon)v$, $x_2 = 0$ in an optimal solution of the above linear program. Thus, we can get $v^{heur} = \frac{1}{1-\epsilon}$ while $v^* = 1$. Since $\epsilon = \frac{N-1}{N}$, we have $v^{heur}/v^* = N$. $\square$

4 Bicriteria approximation

Given a *violation ratio* $\sigma \geq 1$ and an *approximation ratio* $\gamma \geq 1$, a (σ, γ) -bicriteria approximation algorithm for chance constrained optimization problem (2) returns a solution $(\hat{\mathbf{x}}, \hat{\mathbf{z}}) \in \mathbb{R}^n \times \{0, 1\}^N$, such that

$$\mathbf{A}^i \hat{\mathbf{x}} \geq \hat{z}_i \mathbf{b}^i - (1 - \hat{z}_i) \mathbf{M}_i \quad \forall i \in [N], \quad \sum_{i \in [N]} p_i \hat{z}_i \geq 1 - \sigma \epsilon \quad \text{and} \quad \mathbf{c}^\top \hat{\mathbf{x}} \leq \gamma v^*$$

i.e., the solution violates the uncertain constraints with probability at most $\sigma \epsilon$ and has an objective value at most γ times the optimal value. Note that σ, γ may be dependent on the risk parameter ϵ and the underlying probability distribution $\{p_1, \dots, p_N\}$. As a special case, when the violation ratio $\sigma = 1$, we have a single criterion approximation algorithm with an approximation ratio $\gamma \geq 1$.

In the previous section we discussed examples to show that several existing approximation schemes do not come with any constant factor approximation guarantees. The following recent result formally establishes the inapproximability of these problems.

Theorem 6 [45] Suppose we have a polynomial time algorithm that returns a (σ, γ) -approximate solution to the chance constraint optimization problem (2) with a discrete distribution with N equiprobable realizations. Then, unless $P=NP$, the following holds:

- (i) if $\gamma = 1$, then we must have $\sigma = 1/\epsilon - f(N)(1 - \epsilon)/\epsilon$ for some function f such that $f(N) \rightarrow 0$ as $N \rightarrow \infty$;
- (ii) if $\sigma = 1$, then we must have $\gamma = g(N)$ for some function g such that $g(N) \rightarrow \infty$ as $N \rightarrow \infty$.

The above result shows that, unless $P = NP$, it is impossible to obtain a polynomial time algorithm for a chance constrained problem with a constant single criteria approximation in terms of the approximation ratio or the violation ratio. This result motivates the need for the bicriteria approximations that simultaneously relax both optimality and feasibility.

Recently [45] propose a bicriteria approximation for chance constrained covering problems under general distributions. In the following, we present this algorithm and the analysis for the discrete distribution setting and slightly improve the approximation ratios.

4.1 Chance constrained covering problems

For the remainder of the paper we focus on covering type instances of (2), where $\mathbf{x} \in \mathbb{R}_+^n$, $\mathbf{c} \in \mathbb{R}_+^n$ and, for each $i \in [N]$, the data for the constraint system $\mathbf{A}^i \mathbf{x} \geq z_i \mathbf{b}^i$ satisfies $\mathbf{A}^i \in \mathbb{R}_+^{m \times n}$ and $\mathbf{b}^i \in \mathbb{R}_+^m$. In this case, the big- M coefficients in (2) reduce to zero, and the formulation becomes

$$v^* = \min_{\mathbf{x} \in \mathbb{R}_+^n, \mathbf{z}} \left\{ \mathbf{c}^\top \mathbf{x} : \begin{array}{l} \mathbf{A}^i \mathbf{x} \geq z_i \mathbf{b}^i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (28)$$

Although a restricted class, chance constrained covering problems are of importance in a wide variety of applications (cf. [29, 7, 48, 42, 43, 39, 13, 38]). From the complexity point of view, even this restricted problem is NP-hard [34]. Furthermore, the inapproximability result in Theorem 6 applies to these problems. Finally, the worst-case examples in Section 3 are all covering type instances.

The bicriteria approximation algorithm of [45], called Relax-and-Scale, for chance constrained covering problems proceeds by solving a convex relaxation of the problem, and then scaling its solution appropriately to obtain the desired optimality and violation ratios. For (28), the Relax-and-Scale algorithm is as follows.

Algorithm 2 : Relax-and-Scale(σ, γ)

- 1: **input:** parameters $\sigma \geq 1, \gamma \geq 1$
 - 2: **initialize:** set $l = 1, u = \gamma$ and let $\delta > 0$ be a stopping tolerance
 - 3: solve the linear programming relaxation of (28), and let $(\hat{\mathbf{x}}, \hat{\mathbf{z}})$ be an optimal solution
 - 4: **while** $u - l > \delta$ **do**
 - 5: set $\tau \leftarrow (l + u)/2$
 - 6: set $\bar{\mathbf{x}} \leftarrow \tau \hat{\mathbf{x}}$ and $\bar{\mathbf{z}} \leftarrow \min(\lfloor \tau \hat{\mathbf{z}} \rfloor, \mathbf{e})$
 - 7: **if** $\sum_{i \in [N]} p_i \bar{z}_i \geq 1 - \sigma \epsilon$ **set** $u \leftarrow \tau$, **else set** $l \leftarrow \tau$
 - 8: **end while**
 - 9: **output:** $(\bar{\mathbf{x}}, \bar{\mathbf{z}})$ where $\bar{\mathbf{x}} = u \hat{\mathbf{x}}$ and $\bar{\mathbf{z}} = \min(\lfloor u \hat{\mathbf{z}} \rfloor, \mathbf{e})$
-

Next, we establish the choice of the parameters σ, γ , such that Relax-and-Scale(σ, γ) returns a (σ, γ) -bicriteria approximate solution to (28).

4.2 Bicriteria analysis

For the remainder of the paper we assume that $p_i = 1/N$ for all $i \in [N]$. The analysis can be extended to the general (rational) probability setting. The following result is a slight strengthening of Theorem 5 of [45] which establishes a $(\sigma, \frac{\sigma}{\sigma-1})$ -bicriteria approximation result.

Theorem 7 *If the violation ratio and the approximation ratio are chosen as*

$$\sigma \in [1, 1/\epsilon) \text{ and } \gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N},$$

respectively, then Relax-and-Scale(σ, γ) returns a (σ, γ) -bicriteria approximate solution to the chance constrained covering problem (28) with equiprobable scenarios. Furthermore, $\gamma \leq \frac{\sigma}{\sigma-1}$.

Proof: Let $(\hat{x}, \hat{z})$ be an optimal solution of the linear programming relaxation of (28). Note that the Relax-and-Scale outputs a solution $(\bar{x}, \bar{z})$ where $\bar{x} = u\hat{x}$ and $\bar{z} = \min(\lfloor u\hat{z} \rfloor, e)$ for some $u \in [1, \gamma]$ as long as $\sum_{i \in [N]} \bar{z}_i \geq (1 - \sigma\epsilon)N$. If $u\hat{z}_i < 1$ then $\bar{z}_i = 0 \leq u\hat{z}_i$, otherwise $\bar{z}_i = 1 \leq u\hat{z}_i$. Thus $\bar{z} \in \{0, 1\}^N$ and $\bar{z}_i \leq u\hat{z}_i$ for all $i \in [N]$. It then follows that for all $i \in [N]$,

$$A^i \bar{x} \geq b^i \bar{z}_i \Leftrightarrow A^i(u\hat{x}) \geq b^i(u\hat{z}_i) \Rightarrow A^i \bar{x} \geq b^i \bar{z}_i,$$

where the second implication is due to the nonnegativity of b^i .

It remains to show that $\sum_{i \in [N]} \bar{z}_i \geq (1 - \sigma\epsilon)N$. It is sufficient to do this for $u = \gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N}$. Let

$$I := \left\{ i \in [N] : \hat{z}_i \geq \frac{1}{\gamma} = 1 - \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} \right\}.$$

We claim that $|I| \geq N - \lfloor \sigma \epsilon N \rfloor$. For contradiction, suppose not. That is, $|I| \leq N - \lfloor \sigma \epsilon N \rfloor - 1$. Note that this hypothesis implies $N - |I| > 1$. Then

$$\begin{aligned} \frac{1}{N} \sum_{i \in [N]} \hat{z}_i &= \frac{1}{N} \sum_{i \in I} \hat{z}_i + \frac{1}{N} \sum_{i \in [N] \setminus I} \hat{z}_i \\ &< \frac{|I|}{N} + \left(1 - \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} \right) \frac{N - |I|}{N} \\ &= 1 - \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} + \frac{|I|}{N} \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} \\ &\leq 1 - \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} + \frac{N - \lfloor \sigma \epsilon N \rfloor - 1}{N} \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor} \\ &= 1 - \epsilon, \end{aligned}$$

where the first strict inequality is because $\hat{z}_i < 1 - \frac{\epsilon N}{1 + \lfloor \sigma \epsilon N \rfloor}$ for all $i \in [N] \setminus I$ and the fact that $N - |I| > 1$, and the second inequality follows from the hypothesis $|I| \leq N - \lfloor \sigma \epsilon N \rfloor - 1$. This contradicts the fact that $(\hat{x}, \hat{z})$ is a feasible solution to the linear programming relaxation of (28).

Thus $\gamma \hat{z}_i \geq 1$ and $\bar{z}_i = 1$ for all $i \in I$. Then

$$\sum_{i \in [N]} \bar{z}_i \geq \sum_{i \in I} \bar{z}_i = |I| \geq N - \lfloor \sigma \epsilon N \rfloor \geq (1 - \sigma\epsilon)N.$$

We have shown that the constructed solution $(\bar{x}, \bar{z}) \in \mathbb{R}_+^n \times \{0, 1\}^N$ satisfies the constraints

$$A^i \bar{x} \geq \bar{z}_i b^i \quad \forall i \in [N], \quad \sum_{i \in [N]} \bar{z}_i \geq (1 - \sigma\epsilon)N.$$

Also, since $(\hat{x}, \hat{z})$ is an optimal solution to the LP relaxation of (28), it holds that $\mathbf{c}^\top \hat{x} \leq v^*$. Thus $\mathbf{c}^\top \bar{x} = \mathbf{c}^\top(u\hat{x}) \leq uv^* \leq \gamma v^*$. Thus $(\bar{x}, \bar{z})$ is a (σ, γ) -bicriteria approximate solution to the chance constrained covering problem (28) with equiprobable scenarios.

Finally, simple calculations ensure that

$$\frac{\sigma}{\sigma - 1} \geq \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N}.$$

□

The following corollary to Theorem 7 establishes a (non-constant) single criterion approximation guarantee.

Corollary 1 *If the violation ratio $\sigma = 1$ then the Relax-and-Scale algorithm returns a γ -approximate solution to the chance constrained covering problem (28) with equiprobable scenarios with $\gamma = \epsilon N + 1$.*

We close this subsection by providing an example for which the approximation ratio of Theorem 7 is tight.

Example 4 *Given $\sigma \in [1, 1/\epsilon)$, consider the following instance of (28):*

$$v^* = \min_{x \in \mathbb{R}_+^2, z} \left\{ x_1 + x_2 : \begin{array}{ll} \gamma x_1 \geq z_i & \forall i \in [1 + \lfloor \sigma \epsilon N \rfloor] \\ x_1 \geq z_i & \forall i \in [N] \setminus [1 + \lfloor \sigma \epsilon N \rfloor] \\ x_1 + x_2 \geq 1 \\ (1/N) \sum_{i \in [N]} z_i \geq 1 - \epsilon \\ z \in \{0, 1\}^N \end{array} \right\}. \quad (29)$$

where $\gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N}$.

Proposition 3 *Relax-and-scale(σ, γ) applied to (29) (where σ and γ are equal to those of the instance) returns a solution whose violation ratio is at most σ and whose optimality ratio is equal to γ .*

Proof: Note that the optimal value of the linear programming relaxation of (29) is at least 1, and the following solution

$$\hat{x}_1 = \frac{1}{\gamma}, \hat{x}_2 = 1 - \frac{1}{\gamma}, \hat{z}_i = \begin{cases} \frac{1}{\gamma} & \forall i \in [1 + \lfloor \sigma \epsilon N \rfloor] \\ 1 & \forall i \in [N] \setminus [1 + \lfloor \sigma \epsilon N \rfloor] \end{cases}$$

is feasible and achieves the bound, and hence is optimal. For this solution

$$\sum_{i \in [N]} \min\{\lfloor u \hat{z}_i \rfloor, 1\} \geq N - \lfloor \sigma \epsilon N \rfloor$$

if and only if $u \geq \gamma$. Thus the solution returned by Relax-and-scale(σ, γ) must have an approximation ratio equal to γ . □

4.3 Effect of tighter relaxation

The Relax-and-Scale algorithm scales a solution to the linear programming relaxation of the chance constrained problem (28). In Section 2 we discussed several techniques to improve the quality of the relaxations. In this subsection, we address the natural question of whether the performance of the Relax-and-Scale algorithm can be improved by using a stronger relaxation.

We restrict our attention to covering chance constraints (28) with only right-hand-side uncertainties, and equiprobable scenarios. In this case (28) can be written as:

$$v^* = \min_{\mathbf{x} \in \mathbb{R}_+^n, \mathbf{y}, \mathbf{z}} \left\{ \mathbf{c}^\top \mathbf{x} : \begin{array}{l} \mathbf{Ax} \geq \mathbf{y} \\ \mathbf{y} \geq z_i \mathbf{b}^i \quad \forall i \in [N] \\ \sum_{i \in [N]} z_i \geq (1 - \epsilon)N \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (30)$$

Consider the following mixed integer set which is a substructure of (30):

$$Y = \left\{ (\mathbf{y}, \mathbf{z}) : \begin{array}{l} \mathbf{y} \geq z_i \mathbf{b}^i \quad \forall i \in [N] \\ \sum_{i \in [N]} z_i \geq (1 - \epsilon)N \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}.$$

Recall from Section 2.2 that the set Y is a collection of mixing sets intersected with a cardinality constraint. Although it is NP-hard to optimize over Y (and hence to separate over $\text{conv}(Y)$), various families of strong inequalities for $\text{conv}(Y)$ have been derived that have proven to be very effective in computations. These include the mixing inequalities and their variants (see Section 2.2).

Let us consider the following (hypothetical) stronger linear programming relaxation of (30) obtained by adding all possible inequalities describing $\text{conv}(Y)$:

$$\min_{\mathbf{x} \in \mathbb{R}_+^n, \mathbf{y}, \mathbf{z}} \left\{ \mathbf{c}^\top \mathbf{x} : \begin{array}{l} \mathbf{Ax} \geq \mathbf{y} \\ (\mathbf{y}, \mathbf{z}) \in \text{conv}(Y) \end{array} \right\}. \quad (31)$$

The following result shows that even such a tight relaxation does not allow us to improve the approximation ratio of Relax-and-Scale.

Theorem 8 *If the violation ratio and the approximation ratio are chosen as*

$$\sigma \in [1, 1/\epsilon) \text{ and } \gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N},$$

respectively, then Relax-and-Scale(σ, γ) using the relaxation (31) returns a (σ, γ) -bicriteria approximate solution to the chance constrained problem (30). Furthermore, the approximation ratio is tight.

Proof: The proof of the first part is analogous to that of Theorem 7 since (31) is at least as strong the standard linear programming relaxation of (30).

To prove the second part we consider the following instance for given $\sigma \in [1, 1/\epsilon)$ and $\gamma = (1 + \lfloor \sigma \epsilon N \rfloor) / (1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N)$ where ϵN is an integer:

$$\min_{\mathbf{x} \in \mathbb{R}_+^2, \mathbf{z}} \left\{ x_1 + x_2 : \begin{array}{l} x_1 + x_2 \geq 1 \\ \mathbf{e}x_1 \geq \xi^i z_i \quad \forall i \in [N] \\ (1/N) \sum_{i \in [N]} z_i \geq 1 - \epsilon \\ \mathbf{z} \in \{0, 1\}^N \end{array} \right\}. \quad (32)$$

where

$$\xi^i = \begin{cases} (1/\gamma)e^i & i \in [1 + \lfloor \sigma\epsilon N \rfloor] \\ e^i & i \in [N] \setminus [1 + \lfloor \sigma\epsilon N \rfloor], \end{cases}$$

with $e \in \mathbb{R}^N$ being a vector of ones, and $e^i \in \mathbb{R}^N$ being the i -th basis vector of $\mathbb{R}^N$. By introducing the y -variables as in (30) we can rewrite (32) as

$$\min_{x,y,z} \{x_1 + x_2 : x_1 + x_2 \geq 1, x_1 \geq y_i \ \forall i \in [N], (y, z) \in Y\},$$

and

$$Y = \left\{ (y, z) \in \mathbb{R}^{2N} : \begin{array}{l} y_i \geq \frac{1}{\gamma} z_i \ \forall i \in [1 + \lfloor \sigma\epsilon N \rfloor] \\ y_i \geq z_i \ \forall i \in [N] \setminus [1 + \lfloor \sigma\epsilon N \rfloor] \\ \sum_{i \in [N]} z_i \geq (1 - \epsilon)N \\ z \in \{0, 1\}^N \end{array} \right\}.$$

We only need to show that the convex hull of Y , i.e. $\text{conv}(Y)$, is equal to its continuous relaxation:

$$Y^{LP} = \left\{ (y, z) \in \mathbb{R}^{2N} : \begin{array}{l} y_i \geq \frac{1}{\gamma} z_i \ \forall i \in [1 + \lfloor \sigma\epsilon N \rfloor] \\ y_i \geq z_i \ \forall i \in [N] \setminus [1 + \lfloor \sigma\epsilon N \rfloor] \\ \sum_{i \in [N]} z_i \geq (1 - \epsilon)N \\ z \in [0, 1]^N \end{array} \right\}.$$

We know that $\text{conv}(Y) \subseteq Y^{LP}$. Thus, it remains to show that $\text{conv}(Y) \supseteq Y^{LP}$. Equivalently, we can show that all the extreme points of Y^{LP} belong to Y . Indeed, given an extreme point $(\hat{y}, \hat{z}) \in \mathbb{R}^N \times \mathbb{R}^N$ of Y^{LP} , we know that it should satisfy exactly $2N$ equalities. Note that except for the bounds on the z variables (i.e., $z \in [0, 1]^N$), there are only $N + 1$ additional inequalities. We consider the following two cases:

Case 1. $\sum_{i \in [N]} \hat{z}_i > N - \epsilon N$. Then we must have N of the bound constraints binding, i.e. $\hat{z} \in \{0, 1\}^N$.

Case 2. $\sum_{i \in [N]} \hat{z}_i = N - \epsilon N$. It follows that at least $N - 1$ of the bound constraints on $\hat{z}$ must be tight. If N are tight we are done. So suppose only $N - 1$ are tight, i.e. exactly one of $\{\hat{z}_i\}_{i \in [N]}$ is fractional. However recall that ϵN , and so $N - \epsilon N$, is an integer. So this contradicts $\sum_{i \in [N]} \hat{z}_i = N - \epsilon N$. Hence we again have $\hat{z} \in \{0, 1\}^N$.

Thus $\text{conv}(Y) = Y^{LP}$, hence for this instance the stronger relaxation does not offer any strengthening. To establish the tightness of approximation ratio note that, by the above discussion, the stronger relaxation (31) for instance (32) is equivalent to

$$\min_{x,y,z} \{x_1 + x_2 : x_1 + x_2 \geq 1, x_1 \geq y_i, i \in [N], (y, z) \in Y^{LP}\}.$$

By projecting out the y -variables, the above formulation reduces to the continuous relaxation of (29). Hence, by the tightness of approximation ratio in Proposition 3, it follows that the stronger relaxation (31) cannot improve the performance of Relax-and-Scale algorithm. $\square$

5 Some future research directions

In this paper, we focused on MILP based approaches for chance constrained optimization and discussed research on improving relaxation bounds and constructing approximate solutions. There are several important avenues for future research. As discussed in Section 2.2, the mixing set and the cardinality set are important substructures in MILP formulations of chance constraints. Despite significant research, a complete description of the convex hull of the intersection of these substructures is still unknown. Extended formulations of this intersection have been introduced in [18, 24], but a description in the original space will significantly help in developing strong MILP formulations of chance constraints. In Section 2.3, we saw that a successive application of quantile cuts (or projected mixing inequalities) recovers the convex hull of the chance constrained set in the limit. In general, this convergence is infinite. It would be very useful to derive a bound on the quantile rank (i.e. number of successive rounds needed) for specially structured sets, and also tractable methods to approximate the quantile closure. Some preliminary results along this directions are presented in [47]. The Lagrangian dual bounds introduced in Section 2.4 are provably stronger than the LP relaxation bounds, but are, in general, incomparable to the strengthened bounds using mixing inequalities and quantile cuts. A natural question is to establish a relative comparison of these bounds for specially structured problems. Another question in this regard is whether the dual bounds (and their corresponding) primal formulations can be further strengthened using nonlinear constructions (such as in the case of v_2^{LD}). The approximation methods discussed in Section 3 perform well in practice despite their relatively bad worst-case performance. Some important directions are average-case analysis of these methods, tailoring them to specially structured problems, and establishing corresponding performance bounds. For example, [2] identify a large class of problems for which the bisection heuristic (Section 3.3) is optimal. Finally, the bicriteria analysis of Section 4 is restricted to covering problems, and its extensions to more general problem classes would be very significant.

Acknowledgement

This research has been supported in part by the National Science Foundation Award 1633196 and the Office of Naval Research Grant N00014-18-1-2075.

References

- [1] Ahmad Abdi and Ricardo Fukasawa. On the mixing set with a knapsack constraint. *Mathematical Programming*, 157(1):191–217, 2016.
- [2] Shabbir Ahmed, James Luedtke, Yongjia Song, and Weijun Xie. Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. *Mathematical Programming*, 162(1-2):51–81, 2017.
- [3] Shabbir Ahmed and Alexander Shapiro. Solving chance-constrained stochastic programs via sampling and integer programming. In *Tutorials in Operations Research (INFORMS)*, pages 261–269, 2008.
- [4] Gordon J Alexander and Alexandre M Baptista. A comparison of var and cvar constraints on portfolio selection with the mean-variance model. *Management science*, 50(9):1261–1273, 2004.

- [5] Alper Atamtürk, George L Nemhauser, and Martin WP Savelsbergh. The mixed vertex packing problem. *Mathematical Programming*, 89:35–53, 2000.
- [6] Aaron Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.
- [7] Patrizia Beraldi and Andrzej Ruszczyński. A branch and bound method for stochastic integer problems under probabilistic constraints. *Optimization Methods and Software*, 17(3):359–382, 2002.
- [8] Daniel Bienstock, Michael Chertkov, and Sean Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495, 2014.
- [9] Giuseppe C Calafiore and Marco C Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102:25–46, 2005.
- [10] Giuseppe C Calafiore and Marco C Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.
- [11] Abraham Charnes, William W Cooper, and Gifford H Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.
- [12] Yan Deng and Siqian Shen. Decomposition algorithm for optimizing multi-server appointment scheduling with chance constraints. *Mathematical Programming*, 157:245–276, 2016.
- [13] Darinka Dentcheva, András Prékopa, and Andrzej Ruszczyński. Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming*, 89:55–77, 2000.
- [14] Vineet Goyal and R Ravi. Approximation algorithms for robust covering problems with chance constraints. Available at <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1365&context=tepper>, 2008.
- [15] Vineet Goyal and R Ravi. A PTAS for the chance-constrained knapsack problem with random item sizes. *Operations Research Letters*, 38(3):161–164, 2010.
- [16] Yongpei Guan, Shabbir Ahmed, and George L Nemhauser. Sequential pairing of mixed integer inequalities. *Discrete Optimization*, 4:21–39, 2007.
- [17] Oktay Günlük and Yves Pochet. Mixing mixed-integer inequalities. *Mathematical Programming*, 90:429–457, 2001.
- [18] Simge Küçükyavuz. On mixing sets arising in chance-constrained programming. *Mathematical programming*, 132:31–56, 2012.
- [19] Miguel A Lejeune and François Margot. Solving chance-constrained optimization problems with stochastic quadratic inequalities. *Operations Research*, 64(4):939–957, 2016.
- [20] Xiao Liu, Fatma Kılınç-Karzan, and Simge Küçükyavuz. On intersection of two mixing sets with applications to joint chance-constrained programs. *Mathematical Programming*, pages 1–40, 2017.

- [21] Xiao Liu, Simge Küçükyavuz, and James Luedtke. Decomposition algorithms for two-stage chance-constrained programs. *Mathematical Programming*, 157(1):219–243, 2016.
- [22] James Luedtke. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*, 146:219–244, 2014.
- [23] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.
- [24] James Luedtke, Shabbir Ahmed, and George L Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122:247–272, 2010.
- [25] Andrew J. Miller and Laurence A. Wolsey. Tight formulations for some simple mixed integer programs and convex objective integer programs. *Mathematical Programming*, 98:73–88, 2003.
- [26] Arkadi Nemirovski. On safe tractable approximations of chance constraints. *European Journal of Operational Research*, 219:707–718, 2012.
- [27] Arkadi Nemirovski and Alexander Shapiro. Scenario approximation of chance constraints. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design Under Uncertainty*, pages 3–48. Springer, London, 2005.
- [28] Arkadi Nemirovski and Alexander Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006.
- [29] Bernardo K Pagnoncelli, Shabbir Ahmed, and A Shapiro. Sample average approximation method for chance constrained programming: theory and applications. *Journal of optimization theory and applications*, 142(2):399–416, 2009.
- [30] Bernardo K Pagnoncelli, Shabbir Ahmed, and Alexander Shapiro. Computational study of a chance constrained portfolio selection problem. *J. Optim. Theory Appl*, 142(2):399–416, 2009.
- [31] Konstantin Pavlikov, Alexander Veremyev, and Eduardo L. Pasiliao. Optimization of value-at-risk: computational aspects of mip formulations. *Journal of the Operational Research Society*, 69:127–141, 2018.
- [32] Janos Pinter. Deterministic approximations of probability inequalities. *ZOR - Methods and Models of Operations Research*, 33:219–239, 1989.
- [33] András Prékopa. *Stochastic programming*. Springer, 1995.
- [34] Feng Qiu, Shabbir Ahmed, Santanu S Dey, and Laurence A Wolsey. Covering linear programming with violations. *INFORMS Journal on Computing*, 26(3):531–546, 2014.
- [35] R. Tyrell Rockafellar and Stansilav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.
- [36] Rolf Schneider. *Convex bodies: the Brunn–Minkowski theory*, volume 151. Cambridge University Press, 2013.
- [37] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*, volume 9. SIAM, 2009.

- [38] Takayuki Shiina. Numerical solution technique for joint chance-constrained programming problem: An application to electric power capacity expansion. *Journal of the Operations Research Society of Japan*, 42(2):128–140, 1999.
- [39] Lawrence V Snyder and Mark S Daskin. Models for reliable supply chain network design. In *Critical Infrastructure*, pages 257–289. Springer, 2007.
- [40] Yongjia Song, James Luedtke, and Simge Küçükyavuz. Chance-constrained binary packing problems. *INFORMS Journal on Computing*, 26:735 – 747, 2014.
- [41] Chaitanya Swamy. Risk-averse stochastic optimization: probabilistically-constrained models and algorithms for black-box distributions. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1627–1646. SIAM, 2011.
- [42] Andrews K Takyi and Barbara J Lence. Surface water quality management using a multiple-realization chance constraint method. *Water Resources Research*, 35(5):1657–1670, 1999.
- [43] Srinivas Talluri, Ram Narasimhan, and Anand Nair. Vendor performance with supply risk: A chance-constrained dea approach. *International Journal of Production Economics*, 100(2):212–222, 2006.
- [44] Wim van Ackooij, Riadh Zorgati, René Henrion, and Andris Möller. Chance constrained programming and its applications to energy management. In *Stochastic Optimization - Seeing the Optimal for the Uncertain*, Ioannis Dritsas (Ed.), InTech, 2011.
- [45] Weijun Xie and Shabbir Ahmed. Bicriteria approximation of chance constrained covering problems. Submitted for publication. Preprint available at Optimization Online, 2018.
- [46] Weijun Xie and Shabbir Ahmed. Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. (To appear in) *IEEE Transactions on Power Systems*, 2018.
- [47] Weijun Xie and Shabbir Ahmed. On quantile cuts and their closure for chance constrained optimization problems. (To appear in) *Mathematical Programming*, 2018.
- [48] Minjiao Zhang, Simge Küçükyavuz, and Saumya Goel. A branch-and-cut method for dynamic decision making under joint chance constraints. *Management Science*, 60(5):1317–1333, 2014.
- [49] Ming Zhao, Kai Huang, and Bo Zeng. A polyhedral study on chance constrained program with random right-hand side. *Mathematical Programming*, 166(1-2):19–64, 2017.