Chance-Constrained Optimization: A Review of Mixed-Integer Conic Formulations and Applications

Simge Küçükyavuz
Northwestern University

Chance-Constrained Optimization: A Review of Mixed-Integer Conic Formulations and Applications

Simge Küçükyavuz* Ruiwei Jiang †

January 21, 2021

Abstract

Chance-constrained programming (CCP) is one of the most difficult classes of optimization problems that has attracted the attention of researchers since the 1950s. In this survey, we first review recent developments in mixed-integer linear formulations of chance-constrained programs that arise from finite discrete distributions (or sample average approximation). We highlight successful reformulations and decomposition techniques that enable the solution of large-scale instances. We then review active research in distributionally robust CCP, which is a framework to address the ambiguity in the distribution of the random data. The focal point of our review is scalable formulations that can be readily implemented with state-of-the-art optimization software. However, we also discuss alternative approaches and specialized algorithms. Furthermore, we highlight the prevalence of CCPs with a review of applications across multiple domains.

1 Introduction

Most optimization models in practice involve problem parameters that are uncertain. Furthermore, in some cases these uncertain parameters involve risky outcomes with low probability. Therefore, requiring feasibility of a solution for every possible outcome may lead to overly conservative solutions. To remedy this, chance-constrained programming (CCP) has emerged as a powerful paradigm to model system failure/reliability considerations and to address the conservatism of a solution given a certain tolerance for risky outcomes.

For example, in power systems, production levels need to be determined so as to meet peak load (demand) [93]. This problem is complicated by uncertainties in both generator availabilities (especially with renewables) and loads. The utility company's aim is to minimize the expected cost of power production while ensuring that the loss-of-load probability (i.e., the probability that the available generator capacity is insufficient to meet the peak load) is below an acceptable reliability level [163]. In supply chain problems, service level constraints are introduced to limit the probability of stock-outs [40]. In portfolio optimization problems, there is interest to restrict the downside risk at a certain threshold (value-at-risk) [53]. Finally, in communications network design problems, a certain quality of

*Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA, simge@northwestern.edu
†Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA, ruiwei@umich.edu

service (QoS) with respect to packet losses needs to be ensured [148]. Such risk, service, or reliability constraints are modeled using CCPs. We will discuss more applications of CCPs in Section 4.

1.1 Problem Definition

Formally, for a given probability space $(\Omega, \mathcal{F}, \mathbb{P}^0)$, a chance-constrained program (CCP) is given by

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & \mathbb{P}^0(x \in \mathcal{P}(\omega)) \geq 1 - \epsilon, && \text{(1a)} \\ & x \in \mathcal{X}, && \text{(1b)} \end{aligned}$$

where $c \in \mathbb{R}^n$ is a cost vector, $\mathcal{X} \subset \mathbb{R}^n$ represents a compact set defined by deterministic constraints on the decision variables x, possibly including integrality restrictions on some variables, $\omega \in \Omega \subset \mathbb{R}^d$ is a random vector with a true distribution $\mathbb{P}^0$, for a given ω, $\mathcal{P}(\omega)$ represents the set of solutions that are safe or desirable, and $\epsilon \in (0, 1)$ is the risk tolerance for the decision vector x being unsafe. For risk-averse decision makers typical choices for the risk level are small values, e.g., $\epsilon \leq 0.05$. In this survey, we mainly focus on *linear* chance constraints, i.e., polyhedral $\mathcal{P}(\omega)$. More precisely, let

$$\mathcal{P}(\omega) := \{x : T(\omega)x \geq r(\omega)\}, \tag{2}$$

where $T(\omega)$ is an $m \times n$ matrix of random constraint coefficients, and $r(\omega) \in \mathbb{R}^m$ is a vector of random right-hand sides.

Next, we introduce the taxonomy of CCPs. Constraint (1a) is said to be an *individual* chance constraint for $m = 1$, and a *joint* chance constraint for $m > 1$. If, for all $\omega \in \Omega$, we have $T(\omega) = T$ for some deterministic $m \times n$ matrix T, and only $r(\omega)$ is random, we say that the CCP has *right-hand side (RHS) uncertainty.* In contrast, if the so-called *technology matrix* $T(\omega)$ is random, we say that the CCP has *left-hand side (LHS) uncertainty*, regardless of whether $r(\omega)$ is a fixed vector or is random. Most of the work in CCP can be seen as *single-stage* (i.e., static) decision-making problems where the decisions are made here and now, and there are no recourse actions once the uncertainty is revealed. In Section 2.4, we discuss extensions to *two-stage CCPs*. Finally, in many problems of interest, the decision vector x is pure binary and this structure can be exploited to obtain stronger formulations and specialized algorithms. We refer to such CCPs with pure binary variables as *chance-constrained combinatorial optimization* problems.

CCP dates back to the early work of Charnes and Cooper [38], Charnes et al. [39], Miller and Wagner [152], Prékopa [182], and Prékopa [183], who first consider problems with individual or joint chance constraints. We refer the reader to [25, 59, 104, 185, 186, 202] for textbook treatment and detailed reviews that describe the earlier developments in this area. This survey is aimed at reviewing the developments in the past two decades primarily from a mixed-integer conic reformulations perspective.

Despite long-standing interest and ubiquity in practice, CCP remains one of the most challenging class of problems in general. There are two main challenges with CCPs.

1. **Difficulty of evaluating the probability of an undesirable solution.** In practice, the distribution $\mathbb{P}^0$ in the chance constraint is not fully specified. In rare cases when $\mathbb{P}^0$ is a known continuous distribution, calculating the joint probability of several events requires evaluation of a multi-dimensional integral, which is hard to compute accurately [4]. Ben-Tal and Nemirovski [19], Calafiore and Campi [29, 30], and Nemirovski and Shapiro [161, 162] approximate the non-convex chance constraint with convex constraints such that the solution to this approximation is feasible with high probability. However, such methods could yield highly conservative solutions [4] (see Section 2.5). Finally, a black-box simulation model or an oracle may be available to evaluate $\mathbb{P}^0$ for a given solution x, however it is not straightforward to integrate such an oracle within the optimization model and the number of feasible solutions to evaluate is typically huge [228]. In this survey, we focus on two main approaches to address this difficulty, namely the Sample Average Approximation (SAA) approach (Section 2) and the distributionally robust approach (Section 3).

2. **Non-convexity of the feasible set.** For certain special cases such as joint CCPs with RHS uncertainty involving quasi-concave or log-concave distributions [182, 185, 226, 227], or individual chance constraints with LHS uncertainty under a certain log-concave distribution and choice of ϵ [116], such as normal [105], there is an equivalent convex representation of the corresponding CCP. In general, however, chance constraints even in the case with continuous x, polyhedral $\mathcal{P}$, and only RHS uncertainty result non-convex feasible regions in their original variable space. We illustrate this challenge with an example.

   Example 1. [Adapted from [198]] Let ω_1 and ω_2 are dependent random variables with joint probability density function given in Table 1. Consider the CCP with RHS uncertainty

$$\begin{array}{lc} \min & x_1 + x_2 \\ \text{s.t.} & \mathbb{P}^0 \left\{ \begin{array}{rcl} 2x_1 - x_2 & \geq & \omega_1 \\ x_1 + 2x_2 & \geq & \omega_2 \end{array} \right\} \geq 0.6 \\ & x \geq 0. \end{array}$$

   The feasible region of this problem is non-convex as illustrated in Figure 1.

   Table 1: Joint probability density function of ω

Scenario	1	2	3	4	5	6	7	8	9
ω_1	0.75	0.5	0.5	0.25	0.25	0.25	0	0	0
ω_2	1.25	1.5	1.25	1.75	1.5	1.25	2	1.5	1.25
Probability	0.2	0.14	0.06	0.06	0.06	0.3	0.04	0.04	0.1

   $\square$

   Indeed, the resulting problems are NP-hard, in general [145, 162].

There has been a renewed and growing interest in CCP since the early 2000s [61, 196] to tackle these challenges. Capitalizing on the enormous success of mixed-integer programming (MIP) and conic optimization solvers since the early 2000s, our focal point is on reformulations that aim to circumvent the aforementioned challenges and enable progress towards the solution of this difficult class of problems.

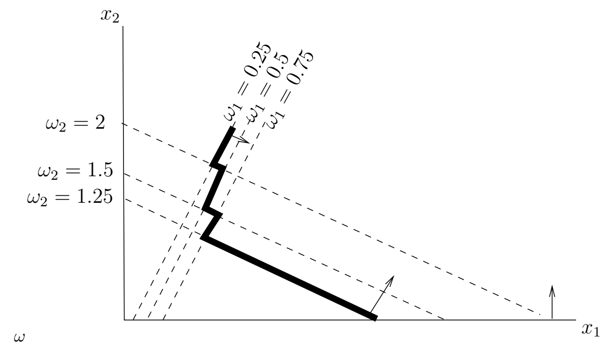


Figure 1: The feasible region of the example CCP.

1.2 Preliminaries

We next present two relevant definitions pertaining to the risk associated with a univariate random variable that will be used in our discussion. We refer the reader to [176, 177, 192] for a more detailed treatment of these risk measures.

Definition 1. For a univariate random variable X, with cumulative distribution function F_X, the *value-at-risk* (VaR) at confidence level $(1-\epsilon)$, also known as $(1-\epsilon)$-quantile, is given by:

$$\mathrm{VaR}_{1-\epsilon}(X) = \min\{\eta \ : \ F_X(\eta) \geq 1-\epsilon\}. \tag{3}$$

□

It follows from (3) that, for any $x \in \mathbb{R}$, the inequalities $\mathrm{VaR}_{1-\epsilon}(X) \leq x$ and $\mathbb{P}(X \leq x) \geq 1-\epsilon$ are equivalent. That is, a chance constraint on random variable X can be equivalently represented as a constraint on its VaR.

Definition 2 ([193, 194])**.** The *conditional value-at-risk* (CVaR) at confidence level $(1-\epsilon) \in (0,1]$ is given by

$$\mathrm{CVaR}_{1-\epsilon}(X) = \min\left\{\eta + \frac{1}{\epsilon}\mathbb{E}\left([X-\eta]_+\right) \ : \ \eta \in \mathbb{R}\right\}, \tag{4}$$

where $(a)_+ := \max\{0, a\}$. □

It is well known that the minimum in definition (4) is attained at the VaR at confidence level $(1-\epsilon)$. CVaR, introduced by Rockafellar and Uryasev [193], satisfies the axioms of coherent risk measures, such as law invariance and sub-additivity, as defined in [9]. It has other desirable properties, such as tractability—for finite distributions, CVaR can be formulated as a linear program and embedded in an optimization model [192]. More precisely, suppose X is a random variable with realizations $X_1, \ldots, X_N$ and corresponding probabilities $p_1, \ldots, p_N$. Throughout, for $a \in \mathbb{Z}_+$, let $[a] := \{1, \ldots, a\}$. The optimization problem in (4) can equivalently be formulated as the linear program (LP):

$$\min\left\{\eta + \frac{1}{\epsilon}\sum_{i\in[N]} p_i w_i \ : \ w_i \geq X_i - \eta, \ \forall\, i \in [N], \quad w \in \mathbb{R}^N_+\right\}. \tag{5}$$

Furthermore, let ρ denote an ordering of the realizations such that $X_{\rho_1} \leq X_{\rho_2} \leq \cdots \leq X_{\rho_N}$. Then, for a given confidence level $\epsilon \in (0, 1]$ we have

$$\mathrm{VaR}_{1-\epsilon}(X) = X_{\rho_q}, \text{ where } q = \min\left\{ j \in [N] : \sum_{i \in [j]} p_{\rho_i} \geq 1 - \epsilon \right\}. \tag{6}$$

1.3 Outline

Our survey is organized as follows. In the first part of this survey, in Section 2, we consider CCPs under a finite discrete distribution. We consider a natural MIP formulation and valid inequalities for both RHS and LHS uncertainty in Sections 2.1 and 2.2, respectively. In Section 2.3, we review alternative formulations and specialized methods for CCPs under a finite distribution. In Section 2.4, we describe a two-stage CCP and a Benders decomposition method for its solution. In Section 2.5 we describe approximations of CCPs. In the second part of this survey, in Section 3, we consider distributionally robust CCPs, primarily under two types of uncertainty sets: moment-based (Section 3.1) and Wasserstein ambiguity sets (Section 3.2). We give an overview of a wide range of applications in Section 4, and conclude in Section 5.

2 CCPs under Finite Discrete Distributions

In this section, we consider CCPs under a finite discrete probability space $(\Omega, 2^\Omega, \mathbb{P}_N)$, where $\Omega = \{\omega_1, \ldots, \omega_N\}$, where $p_i = \mathbb{P}_N(\omega = \omega_i)$. Of particular interest are such CCPs that result from the Sample Average Approximation (SAA) approach [144, 173], which approximates $\mathbb{P}^0$ via a finite empirical distribution, $\mathbb{P}_N$.

For ease of exposition, we will assume that the samples are independent and identically distributed (i.i.d.) and consider the SAA formulation of CCP (i.e., $p_i = \frac{1}{N}, i \in [N]$). The methods we discuss can be adapted to the case of non-i.i.d. scenarios, for example those that are obtained via importance sampling [17].

The SAA formulation of (1) is

$$\min_x \quad c^\top x \tag{7a}$$
$$\text{s.t.} \quad \frac{1}{N} \sum_{i \in [N]} \mathbf{1}(x \notin \mathcal{P}(\omega_i)) \leq \epsilon, \tag{7b}$$
$$x \in \mathcal{X}, \tag{7c}$$

where $\mathbf{1}(\cdot)$ is the indicator function. From this formulation, it is apparent that the use of finite discrete distribution circumvents the first difficulty of evaluating high-dimensional integrals. Under non-equal probability scenarios, constraint (7b) is simply

$$\sum_{i \in [N]} p_i \mathbf{1}(x \notin \mathcal{P}(\omega_i)) \leq \epsilon.$$

When $\mathcal{P}(\cdot)$ is polyhedral as given by (2), formulation (7) for CCP under a discrete distribution lends itself to an equivalent mixed-integer linear program (MIP) via the introduction of binary variables and big-M constraints. Hence, the non-convex feasible region in the original space of variables can be represented as a MIP with additional binary variables. This addresses the second difficulty of non-convexity by enabling the immediate use of off-the-shelf MIP solvers. Next we present such MIP formulations for the RHS and LHS uncertainty cases.

2.1 RHS uncertainty

First, let us consider the problem with RHS uncertainty. In this setting, the joint linear CCP (7) with RHS uncertainty is reformulated as a mixed-integer linear program [196]

$$\min_{x,t,z} \quad c^\top x \tag{8a}$$

$$\text{s.t.} \quad x \in \mathcal{X}, \qquad Tx = \bar{r} + t, \tag{8b}$$

$$t_j \geq r_{i,j}(1 - z_i), \quad \forall i \in [N],\ \forall j \in [m], \tag{8c}$$

$$\frac{1}{N} \sum_{i \in [N]} z_i \leq \epsilon, \tag{8d}$$

$$t \in \mathbb{R}^m_+,\ z \in \{0,1\}^N, \tag{8e}$$

where $\bar{r} \in \mathbb{R}^m$ is chosen vector satisfying $r(\omega_i) \geq \bar{r}$ for all i and $r_i = (r_{i,1}, \ldots, r_{i,m})^\top$ denotes $r(\omega_i) - \bar{r}$. The choice of $\bar{r}$ ensures that the data vector r_i is nonnegative for all $i \in [N]$. For $\epsilon < 1$, we have $Tx \geq \bar{r}$ from (8c)-(8d), hence $t \geq 0$. The binary variable z_i encodes the indicator function in (7b) to model the event $Tx \geq r(\omega_i)$. In particular, if $z_i = 0$, then constraints (8c) enforce that $t \geq r_i$ holds and thus $Tx \geq r(\omega_i)$ is satisfied. Otherwise, $z_i = 1$, and constraints (8c) reduce to the trivial relation $t \geq 0$. Finally, (8d) enforces that the probability of $x \notin \mathcal{P}(\omega)$ is within the risk threshold ϵ. Note that this constraint is equivalent to a cardinality constraint on the binary variables $\sum_{i\in[N]} z_i \leq \lfloor \epsilon N \rfloor =: k$. In the non-equiprobable case, it is a knapsack constraint $\sum_{i\in[N]} p_i z_i \leq \epsilon$.

In the case of individual chance constraints, when $m = 1$, we can linearize the single inequality in the chance constraint as $Tx \geq F_\omega^{-1}(1-\epsilon)$ to lower bound the LHS with the $(1-\epsilon)$-quantile. Therefore, under RHS uncertainty, problems with joint chance constraints $(m > 1)$ are more challenging. In fact, Luedtke et al. [145] show that the problem is NP-hard for $m > 1$. Constraints (8c) are referred to as big-M constraints. Often, formulations with big-M constraints result in weak LP relaxation bounds, which hinder the convergence of the branch-and-bound methods. Therefore, MIP approaches have focused on obtaining strong formulations for the SAA formulation to scale up the problem sizes that can be solved. To this end, an important substructure in the formulation (8) is given by the constraints (8c) and (8e) for a fixed j. This particular substructure is a special case of the *mixing set* studied in [83] that involve general integer variables. Its specific form involving only binary variables is first considered in Atamtürk et al. [14] in the context of vertex covering.

We first consider strengthening based on an individual inequality in the chance constraint. More precisely, consider (8c) and (8e) for a fixed j. We will drop the dependence on j for notational convenience. The resulting system is nothing but a mixing set with binary variables given by

$$\mathcal{M} := \left\{(t, z) \in \mathbb{R}_+ \times \{0,1\}^N : \ t + r_i z_i \geq r_i,\ \forall i \in [N]\right\}.$$

The (binary) mixing set $\mathcal{M}$ involves N inequalities that share a common continuous variable t, but independent binary variables z_i, $i \in [N]$. The so-called *mixing inequalities* of Günlük and Pochet [83] specialized to binary case, which is known to be equivalent to the so-called star inequalities introduced in [14], are an exponential family of linear inequalities that provide the complete linear description of $\text{conv}(\mathcal{M})$ (see also, Pochet and Wolsey [179, Theorem 18]). Furthermore, this class of inequalities can be separated in polynomial time [10, 83], hence formulation (8) can be strengthened using the mixing inequalities within a branch-and-cut framework. Somewhat surprisingly, Kılınç-Karzan et al. [106] uncover that mixing set $\mathcal{M}$ can be viewed as a polymatroid set corresponding to the epigraph of submodular functions. Indeed, the authors show that mixing inequalities are equivalent to extremal polymatroid inequalities as defined in Lovász [139], Atamtürk and Narayanan [12, Proposition 1].

Luedtke et al. [145] further strengthen formulation (8) by exploiting the cardinality constraint (8d) and by studying the resulting set given by (8c)–(8e) for a fixed j. In this case, an immediate strengthening is that of the big-M. Consider the set

$$\mathcal{M}_C := \left\{ (t, z) \in \mathbb{R}_+ \times \{0,1\}^N : \ t + r_i z_i \geq r_i, \ \forall i \in [N], \sum_{i \in [N]} z_i \leq k \right\}.$$

Sort the values r_i for $i \in [N]$, to obtain a permutation σ such that:

$$r_{\sigma_1} \geq r_{\sigma_2} \geq \cdots \geq r_{\sigma_N}.$$

Now observe that due to the cardinality constraint $\sum_{i \in [N]} z_i \leq k$, we must have $t \geq r_{\sigma_{k+1}}$. Therefore, we deduce that

$$\mathcal{M}_C = \left\{ (t, z) \in \mathbb{R}_+ \times \{0,1\}^N : \ t + (r_i - r_{\sigma_{k+1}}) z_i \geq r_i, \ \forall i \in [N], \sum_{i \in [N]} z_i \leq k \right\}.$$

Note, here, that this is an immediate big-M coefficient strengthening that can be readily incorporated into the MIP formulation. This strengthening uses the quantile information that $t \geq r_{\sigma_{k+1}}$.

Due to their common usage, we give a precise definition of the resulting mixing inequalities that make use of the cardinality-based strengthening next. Then, consider a subset $S = \{s_1, s_2, \ldots, s_\ell\} \subseteq \{\sigma_1, \sigma_2, \ldots, \sigma_k\}$ such that $r_{s_i} \geq r_{s_{i+1}}$ for $i = 1, \ldots, \ell$, where $s_1 = \sigma_1$ and $s_{\ell+1} = \sigma_{k+1}$. Luedtke et al. [145] show that a strong mixing inequality valid for $\mathcal{M}_C$ is given by

$$t + \sum_{i=1}^{\ell} \left(r_{s_i} - r_{s_{i+1}} \right) z_{s_i} \geq r_{s_1}. \tag{9}$$

This idea can be adapted to the non-equiprobable case by redefining k as $k := \arg\min\{j : \sum_{i=1}^{j} p_i \leq \epsilon\}$. Furthermore, inequality (9) can be strengthened by further use of the cardinality relation or for the case where the scenarios are not equiprobable when constraint (8d) is in the form of a knapsack inequality [1, 113, 145, 253].

Next, we illustrate this concept on our numerical example (Example 1). Consider the first inequality inside the chance constraint and note that $k = 3$ with respect to ω_1. Note that the scenarios are already ordered in nonincreasing order with respect to the possible values of $r_1(\omega)$. Therefore, we have $t_1 \geq 0.25 = r_1(\omega_4)$. A possible strengthened mixing inequality is for $S = \{1, 3\}$ given by

$$t_1 + (0.75 - 0.5) z_1 + (0.5 - 0.25) z_3 \geq 0.75.$$

It is easy to see the validity of this inequality. If $z_1 = 0$, then we must have $t_1 \geq 0.75$, which satisfies this inequality. If $z_1 = 1$ and $z_3 = 0$, then we must have $t_1 \geq 0.5$, which is also satisfied. Finally, when $z_1 = z_3 = 0$, the inequality reduces to $t_1 \geq 0.25$, which holds due to the $(1-\epsilon)$-quantile relation.

So far, we reviewed inequalities based on an individual inequality inside the chance constraint. If we consider multiple inequalities inside the chance constraint jointly, the resulting set is an intersection of multiple mixing sets that share a common set of binary variables z, but independent continuous variables $t_j, j \in [m]$. For this case, Atamtürk et al. [14, Theorem 3] show that adding the mixing inequalities written for each set to the LP relaxation of the set defined by (8c) and (8e) is sufficient to obtain the convex hull of solutions. Furthermore, Kılınç-Karzan et al. [106] show how to extend their framework exploiting submodularity to recover this result, as well as extend it to propose the so-called *aggregated mixing inequalities* that incorporate lower bounds on the continuous variables based on the quantile relation. For the special case of two-sided chance constraints, the convex hull description provided in Liu et al. [133] are equivalent to the aggregated mixing inequalities. The aggregated mixing inequalities do not directly use the cardinality information, but use it indirectly through the lower bound on the continuous variables obtained from the quantile. In contrast, Küçükyavuz [113] and Zhao et al. [253] propose valid inequalities for a *joint* chance constraint by directly considering the cardinality/knapsack constraint.

2.2 LHS uncertainty

Now consider the problem with uncertainty data in both LHS and RHS. In this setting, the joint linear CCP (7) with LHS uncertainty is reformulated as a mixed-integer linear program [196]

$$\min_{x,z} \quad c^\top x \tag{10a}$$
$$\text{s.t.} \quad x \in \mathcal{X}, \tag{10b}$$
$$T(\omega_i)x \geq r(\omega_i) - M(\omega_i)(1 - z_i), \quad \forall i \in [N], \tag{10c}$$
$$\frac{1}{N}\sum_{i\in[N]} z_i \leq \epsilon, \tag{10d}$$
$$z \in \{0,1\}^N, \tag{10e}$$

where $M(\omega_i), i \in [N]$ is a vector of big-M coefficients such that when $z_i = 1$, inequality (10c) is redundant.

In Section 2.1 we exploited the mixing structure associated with (8c) and (8e) for a fixed j. In other words, we considered an individual inequality inside the (joint) chance constraint. Furthermore, we considered RHS uncertainty only. In contrast, in this section we will consider LHS as well as RHS uncertainty, and we will jointly consider the inequalities inside the chance constraints for any $m \geq 1$.

The mixing procedure described in Section 2.1 relies on the fact that all scenarios share the same LHS for a given $j \in [m]$, that is $t = T_j x$, where T_j is the jth row of T. Due to this, we arrive at a mixing structure with N constraints that share the same continuous variable t and different binary variables. In contrast, in LHS uncertainty case, we no longer have a common continuous variable. Can we still apply the mixing procedure?

As it turns out, we can indeed extend the mixing procedure to generate other classes of valid inequalities for joint chance-constrained programs with LHS uncertainty. To do so, we solve the following single-scenario optimization problem for all scenarios $\omega \in \Omega$ and for a given $\phi \in \mathbb{R}^n$:

$$q_\omega(\phi) = \min_x \quad \phi^\top x \tag{11a}$$

$$x \in \mathcal{P}(\omega), \tag{11b}$$

$$x \in \mathcal{X}. \tag{11c}$$

We sort the values $q_\omega(\phi)$ for $\omega \in \Omega$, to obtain a permutation σ such that:

$$q_{\sigma_1}(\phi) \geq q_{\sigma_2}(\phi) \geq \cdots \geq q_{\sigma_N}(\phi).$$

Observe that $\phi^\top x \geq q_{\sigma_{k+1}}(\phi)$ is a valid inequality. Furthermore, substituting $t = \phi^\top x$ and $r = q(\phi)$ in inequality (9), we obtain a valid inequality of the desired form. These inequalities are referred to as *quantile cuts*. This and related inequalities based on quantile information have been studied in [6, 131, 143, 189, 208, 235]. These inequalities consider the interaction between the decision variables across multiple inequalities in the chance constraint, which results in improved computational performance. In another line of work, Tanner and Ntaimo [212] propose a class of cuts based on the irreducibly infeasible subsystems (IIS) of an LP that requires that a subset of scenarios are satisfied. The authors demonstrate the efficacy of this approach in a vaccine allocation application.

2.3 Alternative formulations and methods

While we focus on natural big-M formulations that can be easily adopted by practitioners, it is important to note that there are alternative reformulations for this class of problems relying on the concept of $(1-\epsilon)$-efficient points, which are an exponential number of points representing the multivariate value-at-risk associated with the chance constraint (12b) to be specified later.

Definition 3. [184] Let $\nu \in \mathbb{R}^m$ be such that $F(\nu) \geq 1-\epsilon$ and $F(\nu - \varepsilon) < 1-\epsilon$ for $\varepsilon \geq \mathbf{0}$, $\varepsilon \neq \mathbf{0}$. The point ν is called $(1-\epsilon)$-*efficient*. □

In Example 1, observe that $\nu \in \{(0.25, 2), (0.5, 1.5), (0.75, 1.25)\}$ is $(1-\epsilon)$-efficient. The $(1-\epsilon)$-efficient points then prescribe the extreme points of the non-convex feasible region as seen in Figure 1.

There are several methods in the literature that rely on the enumeration of the exponentially many $(1-\epsilon)$-efficient points [61, 111, 112, 119, 184, 198]. Such alternative formulations lead to specialized branch-and-bound algorithms described in [22, 23, 196, 197]. Sen [198] uses the $(1-\epsilon)$-efficient points to give a disjunctive programming reformulation of joint chance constraints with finite discrete distributions. Valid inequalities are proposed based on the extreme points of the reverse polar of the disjunctive program, which can be separated by a cut generation linear program (CGLP) [15]. Küçükyavuz [113] gives a compact and tight extended formulation based on disjunctive programming for $m = 1$. Vielma et al. [217] extend this formulation for varying $m > 1$ to obtain a hierarchy of stronger relaxations. Dentcheva et al. [61] use $(1-\epsilon)$-efficient points to obtain various reformulations of probabilistic programs with discrete random variables, and to derive valid bounds on the optimal objective function value. Ruszczyński [196] uses the concept of $(1-\epsilon)$-efficient points to derive consistent orders on different scenarios representing the discrete distribution. The consistent ordering is represented with precedence constraints, and valid inequalities for the resulting precedence-constrained knapsack set are proposed. Beraldi and Ruszczyński [22] propose a branch-and-bound method for probabilistic integer programs using a partial enumeration of the $(1-\epsilon)$-efficient points.

Alternatively, Ahmed et al. [6] and Jiang and Xie [101] consider a Lagrangian relaxation of the MIP formulation by creating copies of the variables, and relaxing the non-anticipativity constraint that these variables are equal. The authors derive extended formulations whose relaxations achieve the stronger bounds than the basic formulation (without mixing strengthening).

Furthermore, for problems with pure binary variables and special structures, i.e., for *combinatorial CCPs*, stronger formulations have been developed (see, e.g., [21, 95, 130, 206, 208, 228]). For example, Song et al. [208] study chance-constrained bin packing problems, and propose a formulation that does not involve additional indicator variables to represent (7b) based on the so-called lifted probabilistic cover inequalities. Later, Wang et al. [225] consider a closely related formulation with multiple chance constraints and derive lifted cover, clique, and projection inequalities based on a bilinear reformulation. In a related line of work, Wang et al. [224] consider a chance-constrained assignment problem and its distributionally robust variant, and propose lifted cover inequalities based on a bilinear reformulation of the problem. For chance-constrained knapsack problems, Yoda and Prékopa [243] provide sufficient conditions for the convexity of the formulation, Klopfenstein and Nace [110], De [54], Han et al. [85], and Joung and Lee [103] derive approximate but more tractable formulations that can provide near-optimal solutions, and Goyal and Ravi [82] derive a fully polynomial time approximation scheme when the random item sizes are independent and Gaussian. In addition, Nikolova [164] studies approximation algorithms for general chance-constrained combinatorial optimization problems with random parameters following either the Gaussian distribution or a general distribution. Xie and Ahmed [236] provide a bicriteria approximation algorithm for a class of chance-constrained covering problems and their distributionally robust variants that finds a solution within constant factor of the violation probability and a constant factor of the optimal objective.

For chance-constrained set covering models with RHS uncertainty, Beraldi and Ruszczyński [23], Saxena et al. [197] propose a specialized branch-and-bound algorithm based on the enumeration of $(1-\epsilon)$-efficient points. Subsequently, Saxena et al. [197] derive polarity cuts to improve the computational performance of this approach. For individual chance-constrained set-covering problems with LHS uncertainty, [73] developed cutting plane approaches for the case that all components of the Bernoulli random vector ω_i are independent. In addition, Wu and Küçükyavuz [228] propose an exact approach for a partial set covering problem for the case that there exists an oracle to retrieve the probability of any events under $\mathbb{P}^0$. In another line of work, Goyal and Ravi [81] and Swamy [210] propose approximation algorithms for chance-constrained set-covering problems with optimality guarantees.

In addition to the aforementioned combinatorial CCPs, Padberg and Rinaldi [172] and Campbell and Thomas [32] study chance-constrained traveling salesman problems, Song and Shen [207] incorporate a chance constraint into a bi-level shortest path interdiction problem, and Ishii et al. [98] and Geetha and Nair [77] study chance-constraint variants of the spanning tree problem.

The focus of this survey is on mixed-integer conic reformulations of CCPs, which yield provably optimal solutions at termination. However, it bears mentioning that there are recent nonlinear programming-based approaches to address the non-convexity of chance constraints. Cheon et al. [46] give a global optimization algorithm that successively partitions the non-convex feasible region until a global optimal solution is obtained. Tayur et al. [213] give an algebraic geometry algorithm for a scheduling problem with joint chance constraints that solves a series of chance-constrained integer programs with varying reliability levels. Peña-Ordieres et al. [175] derive smooth non-convex reformulations of the chance constrained based on the sampled empirical distribution. Other nonlinear programming approaches, which may result in solutions that are stationary points, include difference-of-convex optimization methods [94], sequential outer and inner approximations [78], and sequential cardinality-constrained quadratic optimization methods [50].

Finally, throughout, we have assumed that the risk level ϵ is fixed. However, in practice, the decision-maker may be interested in the trade-offs between risk level and the optimal objective. One way to assess this would be to solve the problem for multiple values of fixed ϵ. For example, Shen [204] proposes a novel *v*ariable risk threshold model in which the risk tolerance is adjustable with an appropriate penalty function in the objective to prevent high risk. The author proposes a MIP formulation for this problem for individual chance constraints. Xie et al. [237, Theorem 8] show that the corresponding optimization problem is strongly NP-hard. Elçi et al. [70] propose a stronger MIP formulation for this problem under RHS uncertainty. Finally, Lejeune and Shen [121] consider joint chance constraints also with LHS uncertainty and propose a Boolean-based mathematical formulation for this model.

2.4 Two-stage Chance-Constrained Programming

Thus far, we have considered a decision-making problem that is static. In other words, the decisions are made here-and-now before the revelation of the outcome of a random event. However, in most practical situations, there are multiple decision stages—intervened by a probabilistic event—and the decision-maker takes recourse actions in the later epochs based on the observed outcome of the event. In this section, we focus on problems that involve two stages. For example, in a power generation setting, the day-ahead problem determines the on/off status of the conventional generators a day before realizing the demand (load) or supply (in case of renewable generators). Then the second-stage problem ensures that the loss-of-load probability is no more than a pre-specified risk level $\epsilon \in (0, 1]$. Therefore, a two-stage chance-constrained model is called for.

As before, the random outcome ω is defined on a probability space $(\Omega, 2^\Omega, \mathbb{P}_N)$. Let $\mathbb{E}[\cdot]$ denote the expectation operator taken with respect to ω. Liu et al. [131] propose the two-stage chance-constrained mixed-integer program

$$\min_x \quad c^\top x + \mathbb{P}_N\left(x \in \mathcal{P}(\omega)\right) \mathbb{E}[h(x, \omega) | x \in \mathcal{P}(\omega)], \tag{12a}$$

$$\mathbb{P}_N(x \in \mathcal{P}(\omega)) \geq 1 - \epsilon \tag{12b}$$

$$x \in \mathcal{X}, \tag{12c}$$

where $\mathcal{P}(\omega) = \{x : \exists y \text{ satisfying } W(\omega)y \geq r(\omega) - T(\omega)x, y \in \mathcal{Y}\}$ and the second-stage problem is given by

$$h(x,\omega) = \min_{y} \quad g(\omega)^\top y \tag{13a}$$
$$W(\omega)y \geq r(\omega) - T(\omega)x \tag{13b}$$
$$y \in \mathcal{Y}. \tag{13c}$$

Here, $g(\omega)$ is a vector of second-stage objective coefficients, $\mathcal{Y}$ is the domain of the second-stage decision vector y. For a related model that considers only the feasibility of the second-stage problem without an associated second-stage cost function $h(x,\omega)$, we refer the reader to [143].

The two-stage chance-constrained problem can be formulated as a large-scale mixed-integer program by introducing a big-M term for each inequality in the chance constraint and a binary variable for each scenario. In particular, analogous to the static CCP, the deterministic equivalent formulation (DEF) of the two-stage CCP may be stated as

$$\min_{x,y,z} \quad c^\top x + \frac{1}{N}\sum_{i\in[N]} g(\omega_i)^\top y(\omega_i) z_i \tag{14a}$$
$$T(\omega_i)x + W(\omega_i)y(\omega_i) \geq r(\omega_i) - M(\omega_i)z_i, \qquad i \in [N] \tag{14b}$$
$$\frac{1}{N}\sum_{i\in[N]} z_i \leq \epsilon, \tag{14c}$$
$$x \in \mathcal{X}, y(\omega_i) \in \mathcal{Y}, \qquad i \in [N] \tag{14d}$$
$$z_i \in \{0,1\} \qquad i \in [N], \tag{14e}$$

where $z_i, i \in [N]$ is a binary variable that equals 0 only if the second-stage problem for scenario ω_i has a feasible solution, and $M(\omega_i)$ is a vector of large enough constants that makes constraint (14b) redundant if $z_i = 1$, i.e., if the second-stage problem for scenario ω_i need not be feasible. The rest of the constraints are interpreted similarly as before.

This formulation poses multiple challenges in addition to the usual difficulties of a formulation with big-M constraints (14b). First, the objective function (14a) is nonlinear. Second, the problem is large scale due to the copies of the variables $y(\omega_i)$ and the large number of binary variables z_i for $i \in [N]$. Nevertheless, the formulation (14) has a decomposable structure—for a fixed first-stage vector x, the problem decomposes into independent scenario problems. Furthermore, if y is a continuous decision vector and $\mathcal{Y}$ is polyhedral, then the second-stage problems are linear programs. Next we describe a Benders-type decomposition algorithm that not only exploits this decomposable structure, but also replaces the weak big-M constraints (14b) with stronger optimality and feasibility cuts, using the mixing structure.

2.4.1 Benders Decomposition-Based Branch-and-Cut Algorithm

Benders method [20], or its specific use in the classical two-stage stochastic programming (without chance constraints) referred to as the L-shaped method [215], is the method of choice for problems that have a similar structure and the second-stage problems are linear programs. However, these methods are not immediately applicable to (14), since both the feasibility and optimality cuts of the Benders method assume that all second stage problems must be feasible, which is not the case for two-stage CCPs. For general recourse problems, feasibility and optimality cuts different from the traditional Benders cuts must be developed.

Let η_i represent a lower bounding approximation of the optimal objective function value of the second-stage problem under scenario $\omega_i, i \in [N]$. Without loss of generality, we assume that $\eta_i \geq 0, i \in [N]$. At each iteration of a Benders decomposition method, a sequence of relaxed master problems (RMP) are solved:

$$\min_{x,z,\eta} \ c^\top x + \frac{1}{N}\sum_{i\in[N]} \eta_i \tag{15a}$$

$$\frac{1}{N}\sum_{i\in[N]} z_i \leq \epsilon, \tag{15b}$$

$$(x, z) \in \mathcal{F}, \tag{15c}$$

$$(x, z, \eta) \in \mathcal{O}, \tag{15d}$$

$$x \in \mathcal{X} \tag{15e}$$

$$z \in \{0, 1\}^N, \tag{15f}$$

where, $\mathcal{F}$ and $\mathcal{O}$ denote the set of feasibility and optimality cuts—to be specified later,—respectively.

At iteration k, let (x^k, z^k) be the optimal solution to the RMP. Given this first-stage solution, suppose that we solve the LP (13) for outcome ω to obtain $h(x^k, \omega)$. The feasibility cuts in set $\mathcal{F}$ are derived from the solution to this LP. If $z_i^k = 0$ for some $i \in [N]$, then the second-stage problem must be feasible. If it is infeasible for a scenario $j \in [N]$, then there exists an extreme ray ψ_{ω_j} associated with the dual of (13) for scenario ω_j that yields the inconsistent solution. Then, letting $\phi = \psi_{\omega_j}^\top T(\omega_j)$ in (11) and following the mixing procedure gives a violated valid inequality that cuts off this infeasible solution (x^k, z^k). If, on the other hand, for all $\omega \in \Omega$, the second-stage problem associated with scenario ω such that $z^k(\omega) = 0$ is indeed feasible, then the current solution (x^k, z^k) is a feasible solution and no feasibility cuts are necessary. However, optimality cuts may be needed. Next we describe how to obtain valid optimality cuts.

Let ψ_{ω_j} be the dual vector associated with the optimal basis of the second-stage problem (13) for scenario ω_j at this iteration. One possible big-M optimality cut is given by [221, 222]

$$\eta_j + M_j z_j \geq \psi_{\omega_j}^\top (r(\omega_j) - T(\omega_j)x), \tag{16}$$

where $M_j, j \in [N]$ is a big-M coefficient vector.

Next we describe a stronger optimality cut proposed by [131] that leads to faster convergence to an optimal solution. Clearly, the traditional Benders optimality cut, $\eta_j \geq \psi_{\omega_j}^\top (r(\omega_j) - T(\omega_j)x)$ is a valid optimality cut for $x \in \mathcal{X}$ (in fact for $x \in \mathcal{P}(\omega)$) if $z_j = 0$. However, it may not be valid for all $x \in \mathcal{X}$ for solutions with $z_j = 1$. To obtain a valid optimality cut, we solve the following secondary problem with $\phi = \psi_{\omega_j}^\top T(\omega_j)$:

$$\bar{v}_{\omega_j}(\phi) = \min_{x,y} \quad \phi x$$
$$x \in \mathcal{X}, \quad y \in \mathcal{Y}.$$

Then we add the optimality cut of the form

$$\eta_j + \left(\psi_{\omega_j}^\top r(\omega_j) - \bar{v}_{\omega_j}(\phi)\right) z_j \geq \psi_{\omega_j}^\top (r(\omega_j) - T(\omega_j)x). \tag{17}$$

To see the validity of this inequality at $z_j = 1$, note that in this case, the second-stage objective function contribution for scenario ω_j is zero. Furthermore, inequality (17) evaluated at $z_j = 1$ reduces to $\eta_j \geq \bar{v}_\omega(\phi) - \phi x$. Because $\bar{v}_\omega(\phi) - \phi x \leq 0$ for all $x \in \mathcal{X}$ and $\eta_j \geq 0$, this inequality is trivially satisfied. The finite convergence of the resulting algorithm is proven in [131] under certain assumptions.

In Table 2, we summarize a set of computational experiments that appear in [131] to show the effectiveness of the approaches discussed so far. The instances are based on a resource planning problem adapted from [143]. In the first stage, the number of servers among s types of servers to employ is determined. The second-stage problem is to allocate the servers to clients of τ types, so that their demands are met with high probability $(1 - \epsilon)$. Instances with various choices of N, ϵ, τ, s are tested and we report the average statistics for three random instances generated for the combination reported in each row. We compare the proposed "Strong" decomposition algorithm which uses the optimality cuts (17) with DEF (14) and the decomposition approach (referred to as "Basic") which uses the mixing-based feasibility cuts and the big-M optimality cuts (16) with an appropriate choice of big-M as described in [131]. We report the solution times (in seconds) only for Strong decomposition, because for DEF and Basic, all instances tested reach the time limit of one hour. We also report the percentage optimality gap at termination under the Gap column. In most cases, DEF is unable to find a feasible solution to the LP relaxation, as indicated by a '-'. In cases when it is able to find a feasible solution, it ends with a gap ranging from 4% to 8%. On the other hand, Basic is able to find a feasible solution for all instances, but is unable to prove optimality for any of the 36 instances tested. It ends after an hour with optimality gaps ranging from 2% to 7%. In contrast, the Strong decomposition algorithm, based on the proposed strong optimality cuts, is able to solve most of the instances to optimality. For the two unsolved instances (indicated by a superscript [1] under the Gap column), the average optimality gap is less than 0.1%. These results highlight the importance of using strong formulations and decomposition for large-scale instances.

It is important to note that in this model, the undesirable outcomes ω such that $x \notin \mathcal{P}(\omega)$ are simply ignored. Liu et al. [131] propose an extension of the two-stage model (12), where they allow so-called recovery decisions for the undesirable scenarios. They discuss how to resolve a potential time inconsistency in two-stage CCP. Furthermore, the Benders decomposition-based solution method is extended to operate in the case of recovery.

Elçi and Noyan [69] extend this framework to a two-stage chance-constrained optimization model with a mean-risk objective, using the conditional value-at-risk as a risk measure. They apply this framework to a humanitarian relief network design problem and demonstrate its effectiveness on a case study based on hurricane preparedness in Southeastern United States. Lodi et al. [136] extend this two-stage framework to convex second-stage problems, motivated by hydro-power scheduling applications. They build an outer approximation of the nonlinear second-stage formulations to design a Benders-type algorithm that converges to an optimal solution under mild assumptions. They demonstrate the computational benefit of the decomposition algorithm on a case study based on hydroplant data from Greece.

We close this subsection by noting the assumption of continuous second-stage variables can be lifted by leveraging

Table 2: Result for instances with random RHS.

Instances		DEF	Basic	Strong	
(N,ϵ)	(s,τ)	`Gap (%)`	`Gap (%)`	`Time`	`Gap (%)`
(2000, 0.05)	(5,10)	4.60	2.34	166	0
	(10,20)	-	2.93	483	0
	(15,30)	-	2.69	1106	0
(2500, 0.05)	(5,10)	4.64	2.61	279	0
	(10,20)	-	3.08	711	0
	(15,30)	-	2.88	1819	0.09[1]
(2000, 0.1)	(5,10)	7.1	5.46	723	0
	(10,20)	-	5.99	1069	0
	(15,30)	-	6.27	1032	0
(2500, 0.1)	(5,10)	7.63	5.32	641	0
	(10,20)	-	5.79	1198	0
	(15,30)	-	6.03	2112	0.02[1]

the developments for decomposition algorithms for classical two-stage stochastic mixed-integer programs, where the second-stage problems also involve integer decisions [35, 75, 115, 117, 167–169, 187, 199–201, 245]. These methods rely on iteratively convexifying the second-stage problems and updating the feasibility and optimality cuts accordingly. These methods can be combined with the Benders-type algorithm we described to enable the solution of two-stage CCPs with integer variables at the second-stage.

2.5 Approximations

Given the difficulty of solving the exact formulations of CCPs or their SAA reformulations, one line of research has focused on inner and outer approximations of CCPs that are more tractable. This tractability often comes at the price of conservatism in the resulting solutions. Here we briefly review these formulations and refer the reader to [5] for a review of relaxations and approximations for CCPs.

- **Scenario approximation.** Scenario approximation (SA) [e.g., 29, 30, 33, 34, 55] entails sampling to approximate the true distribution $\mathbb{P}^0$ with a finite distribution $\mathbb{P}_N$ with a set of outcomes $\Omega = \{\omega_1, \ldots, \omega_N\}$. However, unlike the SAA model (7), a usual stochastic program (not chance-constrained) is solved enforcing that the relations inside the chance constraint hold for each scenario. Thus, the scenario approximation problem is given by

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & x \in \mathcal{P}(\omega), \qquad \omega \in \Omega, && \text{(18a)} \\ & x \in \mathcal{X}, && \text{(18b)} \end{aligned}$$

As a result, for polyhedral $\mathcal{P}(\omega)$ and continuous x, the resulting SA formulation is a large-scale LP. The authors give a finite sample guarantee that the solution to this problem is feasible to the original CCP with high probability. Interestingly, this sample size does not depend on m, under certain assumptions. Unfortunately, the required sample size is typically large and the resulting solution is overly conservative. The SAA approach [144, 173] is aimed at alleviating the conservatism of the SA approach by enforcing the chance constraint, with a smaller risk level, over the finite distribution $\mathbb{P}_N$, albeit as a MIP as opposed to an LP.

- **CVaR approximation.** From Definitions 2 and 3, it is readily apparent that for a univariate random variable X, $\mathrm{CVaR}_{1-\epsilon}(X) \geq \mathrm{VaR}_{1-\epsilon}(X)$. Therefore, for individual chance constraints $(m = 1)$, one can approximate the constraint $\mathbb{P}(r(\omega) - T(\omega)x \leq 0) \geq 1 - \epsilon$, or in other words, $\mathrm{VaR}_{1-\epsilon}(r(\omega) - T(\omega)x) \leq 0$ with $\mathrm{CVaR}_{1-\epsilon}(r(\omega) - T(\omega)x)) \leq 0$. For the case of finite discrete distributions, this approximation leads to tractable reformulations due to the LP representation of CVaR given in (5). In particular, for individual chance constrained CCP (7), the CVaR approximation LP is

$$
\begin{aligned}
\min_{x} \quad & c^\top x \\
\text{s.t.} \quad & \eta + \frac{1}{\epsilon N} \sum_{i \in [N]} w_i \leq 0, \\
& w_i \geq r(\omega_i) - T(\omega_i)x - \eta, \ \forall\, i \in [N], \\
& x \in \mathcal{X}.
\end{aligned}
$$

In general, though, it is not possible to represent CVaR tractably [162]. Nevertheless, Nemirovski and Shapiro [162] give a family of safe (i.e., feasible with high probability) and, in some cases, tractable approximations—referred to as generator-based approximations—that include the Bernstein approximation [178]. They show that the tightest such approximation is a CVaR approximation. However, CVaR approximation is also conservative in some cases [7]. We refer the reader to [160], and references therein, for a survey on related safe tractable approximations for individual chance constraints.

In the case of joint chance constraints $(m > 1)$, it is worthwhile to note that even for the discrete case, while a vector-valued multivariate VaR definition exists (Definition 3), there is no unified definition of multivariate CVaR [see, 150, and the discussions therein]. This poses challenges in formulating related CVaR-based approximations that are tractable. One approach is to scalarize the multivariate random vector $r(\omega) - T(\omega)x$ and use the corresponding univariate CVaR. Considering the ambiguity of the scalarization weights leads to a multivariate CVaR definition that can be represented as a challenging MIP with big-M constraints [165]. MIP strengthening techniques can be used to improve the computational performance of the resulting multivariate CVaR formulations [114, 132, 166].

- **Bonferroni approximation.** Given that joint chance constraints are significantly harder than individual chance constraints, one approximation scheme that is commonly considered replaces the joint chance constraint with m individual chance constraints. In this case, consider replacing the joint chance constraint $\mathbb{P}(T_j(\omega)x \geq r_j(\omega), j \in [m]) \geq 1 - \epsilon$ with

$$
\mathbb{P}(T_j(\omega)x \geq r_j(\omega)) \geq 1 - \epsilon_j, \tag{19}
$$

$$
\text{where} \sum_{j \in [m]} \epsilon_j \leq \epsilon. \tag{20}
$$

From Bonferroni's inequality, it follows that any solution satisfying constraints (19)–(20) also satisfies the joint chance constraint [42, 162]. Because optimizing over ϵ_j is, in general, difficult, a common choice is $\epsilon_j = \epsilon/m, j \in [m]$. However, this is also known to be a conservative approach [41, 162].

Note that while these approximations provide some statistical guarantees for feasibility, they are known to be conservative and do not come with optimality guarantees. Indeed, Xie and Ahmed [236] show an inapproximability result for CCPs. Ahmed [2] uses a similar idea as [162], this time to obtain a convex (Bernstein) relaxation that yield deterministic lower bounds. Integrated chance constraints proposed by Klein Haneveld [108] replaces the non-convex chance constraints with a quantitative measure of shortfalls that lead to polyhedral representations [109] in the discrete case. In this case, they are equivalent to the LP relaxation of the MIP formulation of CCP. Alternatively, statistical lower bounds can be obtained by using order statistics based on SAA solutions [144, 173]. Such deterministic or statistical bounds are useful in assessing the quality of a solution obtained from an approximation.

The finite sample guarantees of sampling based methods [29, 30, 34, 144, 173] are much too large and conservative in practice. On the other hand, for small N, the out-of-sample performance of the SAA solution may even be infeasible to the true problem. For example, in [228], the authors consider a partial set covering problem when an oracle that can evaluate the true probability of the desired event is available. They observe that for sample sizes that lend themselves to a tractable solution of the resulting MIP, the SAA solution is often infeasible to the true problem. This is related to the over-fitting phenomenon in machine learning when the solution of the problem is highly sensitive to the samples $\{\omega_i\}_{i\in[N]}$ used to obtain it. In the next section, we describe an approach that alleviates this problem.

3 Distributionally Robust Chance-Constrained Programming

Given the unavailability of the exact distribution $\mathbb{P}^0$ and the potential overfitting issues due to SAA-based approaches, there has been growing interest in modeling stochastic optimization problems that are *distributionally robust* [see, 190, and references therein].

Formally, a *distributionally robust chance-constrained program* (DRCCP) is modeled as

$$\min_{x} \quad c^\top x \tag{21a}$$

$$\text{s.t.} \quad \sup_{\mathbb{P}\in\mathcal{F}(\beta)} \mathbb{P}(x \notin \mathcal{P}(\omega)) \le \epsilon \tag{21b}$$

$$x \in \mathcal{X}, \tag{21c}$$

where $\mathcal{F}(\beta)$ is an ambiguity set of distributions and β is a set of parameters that describe the ambiguity set. Accordingly, the distributionally robust chance constraint (21b) ensures that the chance constraint is satisfied with respect to all distributions in $\mathcal{F}(\beta)$, even the worst possible one.

Several types of ambiguity sets have been studied in the literature based on various characteristics of the distribution, including moments, shape information (e.g., symmetry and unimodality), support, mixture models, and discrepancy measures (e.g., Wasserstein and ϕ-divergence) [3, 31, 43, 68, 71, 87, 102, 118, 124, 162, 216, 232, 238, 240, 254].

These ambiguity sets lead to different computational tractability and conservatism of the corresponding DRCCP. In this survey, we will focus on moment-based ambiguity sets (Section 3.1) and Wasserstein ambiguity sets (Section 3.2).

3.1 Moment-based ambiguity

There are many successful developments on the tractability of single and joint chance constraints with moment ambiguity sets, which characterize $\mathcal{P}$ based on moment information of $\mathbb{P}^0$ [31, 87, 88, 124, 233, 241, 254].

For known mean value μ and covariance matrix Σ, El Ghaoui et al. [68] characterize a moment ambiguity set

$$\mathcal{F}(\mu, \Sigma) := \{\mathbb{P} : \mathbb{E}[\omega] = \mu, \mathbb{E}[(\omega - \mu)(\omega - \mu)^\top] = \Sigma\}.$$

All probability distributions in $\mathcal{F}(\mu, \Sigma)$ need to have the designated first two moments, and are otherwise allowed to have different distribution types (e.g., Gaussian, Gaussian mixture, etc.) or different support (e.g., discrete or continuous). Perhaps surprisingly, El Ghaoui et al. [68, Theorem 1] show that DRCCP is second-order conic representable for individual chance constraints (i.e., $m = 1$). Specifically, if $T(\omega) := \omega^\top A + T_0$ for some data matrix $A \in \mathbb{R}^{d\times n}$ and vector $T_0 \in \mathbb{R}^{1\times n}$ and $r(\omega) := b^\top \omega + r_0$ for some data vector $b \in \mathbb{R}^d$ and constant $r_0 \in \mathbb{R}$, then constraint (21b) is equivalent to

$$\mu^\top (b - Ax) + \sqrt{\frac{1-\epsilon}{\epsilon}} \|\Sigma^{1/2}(b - Ax)\|_2 \le T_0 x - r_0. \tag{22}$$

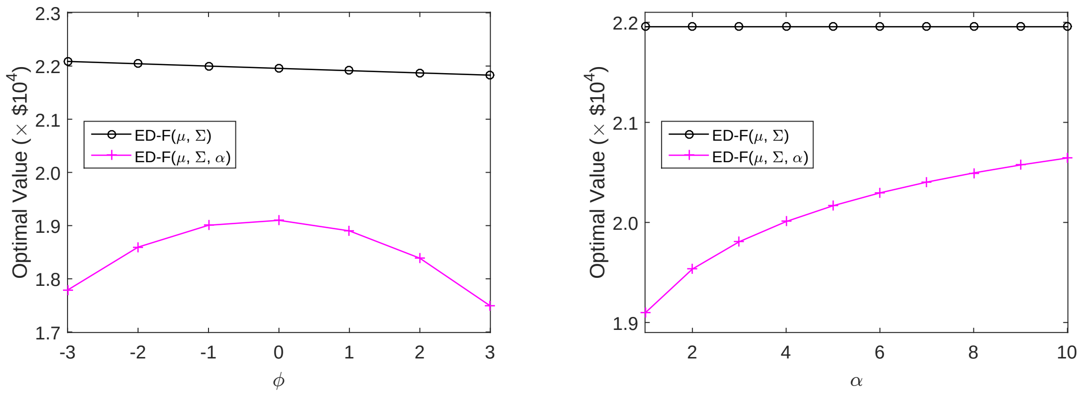
This indicates that DRCCP may improve not only the out-of-sample performance of CCP when the sample size N is small but also the computational tractability. The same result is also discovered by Calafiore and El Ghaoui [31] and Wagner [219]. In addition, Zymler et al. [254] point out an interesting fact that, for $m = 1$ and ambiguity set $\mathcal{F}(\mu, \Sigma)$, constraint (21b) is equivalent to its conservative approximation that replaces the chance constraint with CVaR, i.e., $\sup_{\mathbb{P}\in\mathcal{F}(\mu,\Sigma)} \mathrm{CVaR}_{1-\epsilon}(r(\omega) - T(\omega)x) \le 0$.

For individual chance constraints, the result of El Ghaoui et al. [68] can be extended in multiple directions while maintaining both *exactness* and *computational tractability*. For example, Cheng et al. [45] incorporate support information into $\mathcal{F}(\mu, \Sigma)$ (e.g., specifying that $\mathbb{P}$ is supported on a convex set) and derive an exact reformulation of (21b) based on linear matrix inequalities. Zhang et al. [248] consider potential errors of estimating the mean value μ and covariance matrix Σ, e.g., when this is done based on inadequate historical data. To address this, they adopt an alternative ambiguity set proposed by Delage and Ye [56] to allow the true mean value of ω to be within an ellipsoid centered at μ and the true covariance matrix to be bounded from above by Σ. For this extended ambiguity set, Zhang et al. [248] show that constraint (21b) is still second-order conic representable. For ambiguity set $\mathcal{F}(\mu, \Sigma)$, Xu et al. [238] study a distributionally robust variant of the stochastic dominance constraint (see, e.g., Dentcheva and Ruszczyński [60]), which requires different risk tolerances for violating a chance constraint with different magnitudes. More precisely, they study constraints $\sup_{\mathbb{P}\in\mathcal{F}(\mu,\Sigma)} \mathbb{P}[T(\omega)x \ge r(\omega) - s] \le \epsilon - \beta(s)$ for all $s \ge 0$, where $\beta(s)$ is a pre-specified non-decreasing function of s, and show that these constraints are conic representable for various $\beta(s)$ functions. Furthermore, Yang and Xu [241] and Xie and Ahmed [233] consider an extension that allows the event $x \in \mathcal{P}(\omega)$ to depend non-linearly on x and ω, e.g., $x \in \mathcal{P}(\omega)$ if and only if $f(x, \omega) \ge 0$, where function $f(x, \omega)$ is concave in x and quasiconvex in ω. For example, Yang and Xu [241, Corollary 2] recast (21b) as a linear matrix inequality if $r(\omega)$, as well as each entry of $T(\omega)$, is either convex quadratic or linear in ω.

It is also possible to extend El Ghaoui et al. [68] by incorporating shape information into the ambiguity set $\mathcal{F}(\mu, \Sigma)$. For example, Calafiore and El Ghaoui [31, Lemma 3.1] strengthens $\mathcal{F}(\mu, \Sigma)$ by additionally requiring $\mathbb{P}$ to be *centrally symmetric* (that is, $\mathbb{P}[A] = \mathbb{P}[-A]$ for any Borel set $A \subseteq \mathbb{R}^d$) and derives a conservative approximation of constraint (21b). Hanasusanto [86] considers a similar ambiguity set and allows the true covariance matrix to be bounded from above by Σ (instead of matching it exactly as in $\mathcal{F}(\mu, \Sigma)$). Consequently, Hanasusanto [86, Theorem 3.4.3] recasts (21b) as a set of conic constraints. Different from [31], Li et al. [124, Theorem 1] strengthens $\mathcal{F}(\mu, \Sigma)$ by requiring that $\mathbb{P}$ is *α-unimodal* (a generalized notion of unimodality; see Dharmadhikari and Joag-Dev [62] for definition). They show that constraint (21b) is equivalent to a set of second-order conic constraints. Hanasusanto [86, Example 3.4.4] considers a similar ambiguity set, which bound the true covariance matrix from above by Σ, and recasts (21b) as linear matrix inequalities. Stellato [209] also considers a similar ambiguity set as in Li et al. [124] but requires $\mathbb{P}$ to be centered around μ. In that case, Stellato [209, Section 4.1.1] recasts (21b) as a single second-order conic constraint. There are works that consider other shape information and provide *tractable conservative approximations* of (21b) (i.e., maintaining computational tractability at a potential cost of exactness). For example, Chen et al. [42] replace the covariance information in $\mathcal{F}(\mu, \sigma)$ with bounds on forward and backward deviations, which capture the *asymmetry* of $\mathbb{P}$, and derive a conservative approximation of (21b) via second-order conic constraints. Li et al. [123] drop the covariance restriction from $\mathcal{F}(\mu, \Sigma)$ while adding in that $\mathbb{P}$ is *log-concave* and supported on an ellipsoid centered at μ. For this case, Li et al. [123] derive conservative and relaxing approximations of (21b), all via second-order conic constraints. Postek et al. [180] replace the covariance information in $\mathcal{F}(\mu, \Sigma)$ with the mean absolute deviation (MAD) from the mean and further require that ω is componentwise *independent*. For that case, Postek et al. [180] derive a conservative approximation of (21b) based on second-order conic constraints.

The special case of *combinatorial DRCCPs* with individual chance constraints is in general *intractable* because of the binary decision variables. Nevertheless, various formulation strengthening and algorithmic techniques can be applied to solve these problems more effectively. For example, Ahmed and Papageorgiou [3] exploit supermodularity of their distributionally robust set covering problem to derive a stronger and compact reformulation. Zhang et al. [248] derive a submodular relaxation of their DRCCP reformulation for a general binary packing problem and apply extended polymatroid inequalities. Zhang et al. [252] integrate various algorithmic techniques, including coefficient strengthening and structure-aware reformulation, into a branch-and-price algorithm to solve a bin packing problem.

Tractable reformulations for distributionally robust *joint* chance constraints, i.e., constraint (21b) with $m \geq 2$, are much scarcer than for individual chance constraints. Indeed, Hanasusanto et al. [88, Section 2.3] show that DRCCP becomes NP-hard if the ambiguity set involves any non-homogeneous dispersion measure (e.g., covariance as in $\mathcal{F}(\mu, \Sigma)$) or any non-conic support (e.g., a hyperrectangle), or if $T(\omega)$ involves any uncertainty (i.e., if $T(\omega) \neq T_0$ for some data matrix $T_0 \in \mathbb{R}^{m \times n}$). Nevertheless, tractable reformulations do exist for ambiguity sets different from $\mathcal{F}(\mu, \Sigma)$ or for chance constraints less general than (21b). For example, Hanasusanto et al. [88, Theorem 2] characterize an ambiguity set by the mean value, a positively homogeneous dispersion measure (e.g., MAD), and a conic support of ω, and derive a second-order conic reformulation of constraint (21b), in which $T(\omega) = T_0$. Xie



(a) Optimal Value vs. ϕ

(b) Optimal Value vs. α

Figure 2: Optimal values of ED-$\mathcal{F}(\mu, \Sigma)$ and ED-$\mathcal{F}(\mu, \Sigma, \alpha)$ with various ϕ and α

and Ahmed [234, Theorem 2] consider a two-sided variant of (21b) with $m = 2$ and $T_1(\omega) = -T_2(\omega)$ and derive a second-order conic reformulation of constraint (21b) with regard to ambiguity set $\mathcal{F}(\mu, \Sigma)$. Xie and Ahmed [233] derive exact and tractable reformulations of (21b) with regard to multiple types ambiguity sets, e.g., when $\mathcal{F}(\beta)$ involves linear moment constraints only (i.e., on the mean value of ω) or when $\mathcal{F}(\beta)$ consists of a single (possibly nonlinear) moment constraint. Xie et al. [237] consider a subclass of constraints (21b) with separable uncertainties across individual inequalities, i.e., each row of $[T(\omega); r(\omega)]$ involves a different set of uncertain parameters and, correspondingly, a different ambiguity set. They show that, if either $T(\omega)$ or $r(\omega)$ involves no uncertainty, then (21b) admits an exact and tractable reformulation by applying the Bonferroni approximation (or union bound; see Bonferroni [28]).

Various conservative approximations for distributionally robust joint chance constraints have been proposed. Chen et al. [41] propose to approximate the chance constraint in (21b) by using CVaR and subsequently approximate the resulting distributionally robust CVaR (DR-CVaR) constraint via a classical inequality of order statistics. These two layers of approximation lead to a set of second-order conic constraints. Later, Zymler et al. [254] show that the second-layer approximation can be circumvented by deriving an exact reformulation of the DR-CVaR constraint, yielding a linear matrix inequality approximation of (21b). The approximations of [41] and [254] can both be further improved by tuning certain scaling parameters. Unfortunately, it appears to be difficult to simultaneously optimize such scaling parameters and the decision x in DRCCP. Cheng et al. [45] obtain a different approximation from that of [254] when different rows of $T(\omega)$ are independent.

In Figs. 2a–2b, we summarize a case study of a distributionally robust chance-constrained economic dispatch (ED) problem that appears in Li et al. [124] to demonstrate the difference between $\mathcal{F}(\mu, \Sigma)$ and an alternative ambiguity set that incorporates α-unimodality into $\mathcal{F}(\mu, \Sigma)$, denoted by $\mathcal{F}(\mu, \Sigma, \alpha)$. Their case study uses the IEEE 30-bus system and incorporates two uncertain parameters, representing prediction errors of the forecast power outputs at two wind farms. The formulation and parameters of this problem can be found in [124, Section 5.1]. In particular, we assume that the uncertainties are α-unimodal with a mode at $[0, 0]^\top$ and have a mean value $\mu = \phi[1, 1]^\top$ with

$\phi \in \{-3, -2, \ldots, 3\}$. In Fig. 2a, we compare the optimal values of ED with regard to $\mathcal{F}(\mu, \Sigma)$ and that of ED with regard to $\mathcal{F}(\mu, \Sigma, \alpha)$ with $\alpha = 1$ and various ϕ values. From this figure, we observe that the optimal value of ED-$\mathcal{F}(\mu, \Sigma)$ is consistently higher than that of ED-$\mathcal{F}(\mu, \Sigma, \alpha)$. This confirms that incorporating unimodality into the ambiguity set makes DRCCP less conservative. In Fig. 2b, we compare the optimal values of ED-$\mathcal{F}(\mu, \Sigma)$ and ED-$\mathcal{F}(\mu, \Sigma, \alpha)$ with $\phi = 0$ and various α values. From this figure, we observe that, although the discrepancy between ED-$\mathcal{F}(\mu, \Sigma)$ and ED-$\mathcal{F}(\mu, \Sigma, \alpha)$ declines as α increases, the convergence is sub-linear (in fact, it takes place when α exceeds 10^4). This demonstrates the significant influence of unimodality upon the ambiguity set and the corresponding DRCCP.

The case study just described highlights the utility of available distribution information in reducing the degree of conservatism. In this regard, moment ambiguity sets are known to be more conservative than their counterparts based on discrepancy measures (e.g., a Wasserstein ambiguity set) when more data samples are available. On the other hand, there is a trade-off between conservatism and tractability—unlike with moment-based ambiguity sets, DRCCP with a Wasserstein ambiguity set is not polynomially solvable in general [236]. However, there have been recent developments in MIP formulations for DRCCP under Wasserstein ambiguity, which we describe in the next section.

3.2 Wasserstein ambiguity

Due to its desirable statistical properties, the so-called *Wasserstein* ambiguity set has witnessed an explosion of interest. Wasserstein ambiguity set $\mathcal{F}(N, \theta)$ is defined as the θ-radius Wasserstein ball of distributions on $\mathbb{R}^d$ around the empirical distribution $\mathbb{P}_N$. This is defined as

$$d_W(\mathbb{P}, \mathbb{P}') := \inf_{\Pi} \left\{ \mathbb{E}_{(\omega, \omega')\sim\Pi}[\|\omega - \omega'\|] : \Pi \text{ has marginal distributions } \mathbb{P}, \mathbb{P}' \right\},$$

where the *1-Wasserstein distance*, based on a norm $\|\cdot\|$, between two distributions $\mathbb{P}$ and $\mathbb{P}'$ is used. The Wasserstein ambiguity set is then defined as $\mathcal{F}(\mathbb{P}_N, \theta) := \{\mathbb{P} : d_W(\mathbb{P}_N, \mathbb{P}) \leq \theta\}$. Given a decision $x \in \mathcal{X}$ and random realization $\omega \in \mathbb{R}^d$, we first define a safety set, $\mathcal{S}(x)$, of outcomes such that $\mathcal{S}(x) = \{\omega \in \Omega : x \in \mathcal{P}(w)\}$. The distance from ω to the unsafe set is

$$\text{dist}(\omega, \mathcal{S}(x)) := \inf_{\omega' \in \mathbb{R}^d} \{\|\omega - \omega'\| : \omega' \notin \mathcal{S}(x)\}. \tag{23}$$

Chen et al. [43, Theorem 3] and Xie [232, Proposition 1] show that the formulation

$$\min_{x,v,u} \quad c^\top x$$

$$x \in \mathcal{X},\ v \geq 0, u_i \geq 0,\ i \in [N], \tag{24a}$$

$$\text{dist}(\omega_i, \mathcal{S}(x)) \geq v - u_i,\ i \in [N], \tag{24b}$$

$$\epsilon\, v \geq \theta + \frac{1}{N} \sum_{i \in [N]} u_i \tag{24c}$$

is an equivalent formulation of (21), by using the dual representation for the worst-case probability $\mathbb{P}[x \notin \mathcal{P}(\omega)]$ under the Wasserstein ambiguity set $\mathbb{P} \in \mathcal{F}(\mathbb{P}_N, \theta)$ provided in [27, 76, 153]. (See also Hota et al. [96] for a deterministic non-convex reformulation of (21) and CVaR-based inner approximation of (21) for certain safety sets.)

Note that formulation (21) is non-convex due to constraint (24b). However, for certain safety sets $\mathcal{S}(\cdot)$, MIP reformulations are possible [43, 99, 232]. Therefore, we can once again formulate a deterministic equivalent model and solve it using off-the-shelf optimization software, thereby enabling the usage of these models by practitioners.

3.2.1 RHS Uncertainty

In this section, we consider joint chance constraints with RHS uncertainty under certain common form of a safety set. In particular, let

$$\mathcal{S}(x) := \{\omega : \, Tx \geq r(\omega)\}, \tag{25}$$

where $r(\omega) := B\omega + e$, for a given an $m \times d$ data matrix B, $e \in \mathbb{R}^m$, and T is a given $m \times n$ data matrix. For $m = 1$ (resp. $m > 1$), we say that the problem is an individual (resp. joint) chance-constrained problem with RHS uncertainty. Let T_j and B_j be a row vector of appropriate dimension corresponding to the jth row of T and B, respectively. In this case, the distance function is evaluated as [43]

$$\operatorname{dist}(\omega, \mathcal{S}(x)) = \max\left\{0, \, \min_{j \in [m]} \frac{T_j x - B_j \omega - e_j}{\|B_j\|_*}\right\}, \tag{26}$$

where $\|\cdot\|_*$ is the dual norm. We can then introduce binary variables, z, to capture the non-convex constraint (24b) to arrive at the mixed-integer *linear* program [43, Proposition 2]

$$\min_{z,u,v,x} \quad c^\top x \tag{27a}$$

$$\text{s.t.} \quad z \in \{0,1\}^N, \; v \geq 0, \; u_i \geq 0, \; i \in [N], \; x \in \mathcal{X}, \tag{27b}$$

$$\epsilon\, v \geq \theta + \frac{1}{N} \sum_{i \in [N]} u_i, \tag{27c}$$

$$M(1 - z_i) \geq v - u_i, \quad i \in [N], \tag{27d}$$

$$\frac{T_j x - B_j \omega_i - e_j}{\|B_j\|_*} + M_i z_i \geq v - u_i, \quad i \in [N], \; j \in [m], \tag{27e}$$

where $M_i, i \in [N]$ is a sufficiently large Big-M coefficient.

A few remarks are in order. The computational studies of [43, 232] indicate that this MIP reformulation is difficult to solve in certain cases—state-of-the-art solvers terminate with large optimality gaps after an hour time limit. To address this challenge, Ho-Nguyen et al. [91] propose a number of results that make an order of magnitude improvement in the solution times. Note that formulation (30) is not immediately amenable to the improvements we described for the SAA counterpart. For example, constraints (30e) do not have the mixing structure that the SAA counterpart benefited greatly from. In particular, the continuous variables u_i are not shared across scenarios, whereas the mixing set requires common continuous variables. On the other hand, as argued in [91], the SAA counterpart is a relaxation of (30). By making a key observation that relates the nominal SAA problem for $\mathbb{P}_N$ to formulation (30), Ho-Nguyen et al. [91] give a stronger formulation and valid inequalities based on the same set of binary variables z. Furthermore, this strengthening does have the mixing structure. They also use pre-processing techniques to reduce the formulation size drastically. On a related note, Ji and Lejeune [99] give a different MIP formulation of (21) under Wasserstein ambiguity under additional assumptions on the support of ω.

3.2.2 LHS uncertainty

In this section, we consider joint chance constraints with RHS uncertainty under certain common form of a safety set. In particular, let

$$\mathcal{S}(x) := \{\omega : \ T(\omega)x \geq r(\omega)\}, \tag{28}$$

where $r_j(\omega) := b^\top \omega^j + e_j, j \in [m]$, for a given vector $b \in \mathbb{R}^\kappa$, $\omega^j, j \in [m]$ is a projection of ω to a κ-dimensional vector, and $e \in \mathbb{R}^m$. Also, let the jth row of $T(\omega)$ be given by $T_j(\omega) := \omega^\top A + T_j$ for some $n \times \kappa$ data matrix $A^\top$ and $T \in \mathbb{R}^{m \times n}$. In this case, the distance function is measured by

$$\operatorname{dist}(\omega, \mathcal{S}(\mathbf{x})) = \max\left\{0, \ \min_{p \in [P]} \frac{T_j(\omega)x - r_j(\omega)}{\|A^\top x - b\|_*}\right\}, \tag{29}$$

We can then introduce binary variables, z to represent the non-convex constraint (24b) and make a transformation of variables to arrive at the mixed-integer *conic* program ([232, Theorem 2] and [44, Proposition 1 (for $m = 1$)]

$$\begin{aligned}
\min_{z,u,v,x} \quad & c^\top x && \text{(30a)}\\
\text{s.t.} \quad & z \in \{0,1\}^N, \ v \geq 0, \ u_i \geq 0, \ i \in [N], \ x \in \mathcal{X}, && \text{(30b)}\\
& \epsilon\, v \geq \theta \|A^\top x - b\|_* + \frac{1}{N} \sum_{i \in [N]} u_i, && \text{(30c)}\\
& M_i(1 - z_i) \geq v - u_i, \quad i \in [N], && \text{(30d)}\\
& T_j(\omega_i)x - r_j(\omega_i) + M_i z_i \geq v - u_i, \quad i \in [N], \ j \in [m], && \text{(30e)}
\end{aligned}$$

where $M_i, i \in [N]$ is a sufficiently large Big-M coefficient, under the assumption that $A^\top x \neq b$ for any $x \in \mathcal{X}$. This assumption can be relaxed with appropriate safeguards as described in [44, 92, 232].

As in the case of SAA, the computational studies show that the LHS uncertainty case is a more challenging case than the RHS uncertainty only. First, the resulting formulation is no longer linear, but conic. Furthermore, the coefficients of the common variables x are scenario-dependent unlike the RHS uncertainty case. So it is not clear if similar enhancements that Ho-Nguyen et al. [91] performed for the RHS uncertainty case can be done here. To this end, Ho-Nguyen et al. [92] establish the link between the DRCCP and its SAA counterpart for the LHS case to identify mixing-type valid inequalities and strengthen the formulation. This results in significant improvements in the performance of the resulting MIP formulation. Distributionally robust variants of the resource planning problem (described in Section 2.4) with $N = 100$ that are unsolvable or terminate with high end gaps (40-80%) with the original formulation are now solvable or have much small end gaps (<15%) with the enhancements proposed in [92].

For *combinatorial DRCCPs*, for which the decision variables are pure binary, further strengthening is possible. Xie [232] observe the submodularity of the norm and the terms in the distance operator, and propose the use of polymatroid inequalities to strengthen the formulation. They report significant improvements in the performance of the resulting algorithm. Kılınç-Karzan et al. [107] show how the polymatroid inequalities derived from the conic constraint can be generalized to the case of mixed-binary decisions. In addition, Shen and Jiang [203] derive polymatroid inequalities when the random parameters are binary-valued and show how these inequalities can be further strengthened via mixing and lifting schemes. In a related line of work, Wang et al. [224] consider an assignment problem and derive lifted cover inequalities based on a bilinear reformulation of their DRCCP.

Conservative approximations for DRCCP with Wasserstein ambiguity are related to their SAA counterparts described in Section 2.5. The approach of Erdoğan and Iyengar [71] may be seen as a (robust) scenario approximation counterpart of [29, 33] with similar sample complexity results when the uncertainty set is defined by a Prohorov metric, which is related to a Wasserstein metric. Furthermore, for distributionally robust CCPs under Wasserstein ambiguity [96] give an approximation based on a CVaR interpretation of the reformulation [see, also, 232, for this and two other approximations based on the scenario approximation and VaR approximation].

4 Applications

CCP is used to model risk-averse decision-making problems in a plethora of applications, ranging from chemical processes [89, 90] to water quality management [211]. In this section, we review a few recent and active application domains—this is not meant to be an exhaustive list.

Finance. Chance constraints (or equivalently, VaR as defined in (3)) have been applied in finance to control risks. Linsmeier and Pearson [129] provide motivation of using VaR as a risk measure in significant volatile financial markets. VaR has been widely adopted (e.g., by the US Securities and Exchange Commission) as a method of quantifying risks. Lemus Rodriguez [122], El Ghaoui et al. [68], Natarajan et al. [159], Zymler et al. [255], Huang and Zhao [97], Yao et al. [242], Çetinkaya and Thiele [37], Barrieu and Scandolo [18], Lotfi and Zenios [138], Li et al. [126], and Ji and Lejeune [100] apply VaR and worst-case VaR (analogous to the distributionally robust chance constraints) in finance via mathematical optimization. In addition, Rujeerapaiboon et al. [195] and Choi et al. [47] apply chance constraints in multi-period portfolio optimization.

Healthcare. Chance constraints find applications in appointment scheduling (e.g., Deng and Shen [57]), surgery planning (e.g., Deng et al. [58], Wang et al. [223], and Zhang et al. [249]), operating room planning (e.g., Wang et al. [225], Wang et al. [224], and Najjarbashi and Lim [158]), vaccine allocation (e.g., Tanner and Ntaimo [212]), and social distancing during a pandemic (e.g., Duque et al. [67]), among others.

Power Systems. Zhang and Li [244], Bienstock et al. [24], Zhang et al. [247], Duan et al. [66], Lubin et al. [141, 142] Dall'Anese et al. [52], Xie and Ahmed [234], Li et al. [123], and Li et al. [125] study chance-constrained variants of the optimal power flow problem. Ozturk et al. [171], Pozo and Contreras [181], and Wang et al. [222] consider chance constraints in the unit commitment problem. Vrakopoulou et al. [218], Pozo and Contreras [181], and Wu et al. [229] apply chance constraints to schedule electricity systems in face of random outages and contingencies. Liu et al. [134], Liu et al. [135], Ravichandran et al. [191], and Zhang et al. [251] employ chance constraints to model an integrated system of power grid and electric vehicles. Other power system applications include coordinated load control (e.g., Zhang et al. [247] and Zhang et al. [250]), power grid topology control (e.g., Qiu and Wang [188] and Mazadi et al. [149]), and hydro power plant scheduling (e.g., Wu et al. [230] and Lodi et al. [137]). We refer the reader to a recent survey [214] and references therein for a more detailed review of CCP in energy management.

Transportation and Routing. Dinh et al. [63], Moser et al. [155], Pelletier et al. [174], Du et al. [64], Wu et al. [231], Muraleedharan et al. [156], Ghosal and Wiesemann [79], and Florio et al. [74] study chance constraints in the optimal route design for vehicles (also see Cordeau et al. [49]). Blackmore et al. [26], Farrokhsiar and Najjaran [72], Banerjee et al. [16], Du Toit and Burdick [65], d. S. Arantes et al. [51], Castillo-Lopez et al. [36], and Oh et al. [170] study chance constraints to find paths for robots while avoiding obstacles.

Supply Chain, Logistics, and Scheduling. Wang [220], Song and Luedtke [206], Hong et al. [95], Elçi and Noyan [69], Elçi et al. [70], and Noyan et al. [166] employ chance constraints in the design of networks for logistics and humanitarian relief. Lejeune and Ruszczyński [120], Murr and Prékopa [157], Zhang et al. [246], and Liu and Küçükyavuz [130] apply chance constraints in logistics. Gurvich et al. [84] study chance constraints in the staffing of call centers. Cohen et al. [48] apply chance constraints to cloud computing. Lu et al. [140] apply chance constraints in non-profit resource allocation.

Wireless Communication. Li et al. [128], Soltani et al. [205], Mokari et al. [154], and Xu and Nallanathan [239] apply chance-constrained programming to accommodate the data rate requirement in orthogonal frequency division multiple access (OFDMA) systems. Ma and Sun [147] and Li et al. [127] apply chance constraints on the beamforming problem in communication networks.

5 Concluding Remarks

In this survey, we reviewed mixed-integer conic formulations of CCPs under various distributional assumptions. We described the trade-offs between tractability and conservatism of the corresponding optimization models, as well as the trade-offs between the amount of distributional information used and over-fitting. There is some theoretical guidance on selecting sample sizes or other design parameters, such as the Wasserstein ball radius. However, this guidance is conservative, and instead the parameter choices are made and statistically verified using out-of-sample tests and cross-validation, in practice. There are many opportunities that arise from the recent developments in CCP models. As we outlined, these models often lead to mixed-integer conic formulations, which optimization software is now able to handle in modest sizes. The novel mixed-integer conic CCP models when coupled with parallel developments in strengthening mixed-integer conic formulations [11–13, 107, 232, 248] are likely to enable the solution of large-scale problems before resorting to conservative approximations. Such strengthening approaches often exploit hidden submodularity—a recurring structure in many reformulations we discussed. Approximations continue to play an important role in applications where faster solution times are needed. In such cases, it is of interest to be able to provide some performance guarantees. In this regard, recent research in deriving strong relaxations and approximation algorithms for structured problems is promising.

We have primarily discussed single- or two-stage problems in this survey. Conceptually, one can also envision CCPs with multiple decision epochs. Zhang et al. [246] consider multi-stage CCPs and give valid inequalities for the SAA reformulation. Lulli and Sen [146] consider a multi-stage problem under a finite discrete demand distribution, and propose a model wherein non-anticipativity is enforced only for the scenarios that meet the desired service constraint. The authors propose a branch-and-price algorithm, for the resulting formulation. Andrieu et al. [8], González Grandón et al. [80], and references therein, consider problems with dynamic chance constraints, and propose solution methods under certain continuous distributions. Meraklı and Küçükyavuz [151] consider the risk associated with parameter uncertainty in infinite-horizon Markov decision processes, and formulate this problem using a chance-constrained optimization framework. Models and methods for multi-stage CCPs are sparser due to their inherent difficulty not only in modeling, by taking into account the time consistency of solutions, but also in designing scalable solution methods. This is an area of further research.

In closing, we believe that the developments in easy-to-implement reformulations will usher in new and exciting applications of CCPs, given the increasingly uncertain conditions of operations in various sectors (extreme weather, autonomous devices, renewable power, pandemics, political unrest, etc.).

Acknowledgments

Simge Küçükyavuz is supported, in part, by ONR grant N00014-19-1-2321 and NSF grant 2007814. Ruiwei Jiang is supported, in part, by NSF grant ECCS-1845980.

References

[1] A. Abdi and R. Fukasawa. On the mixing set with a knapsack constraint. *Mathematical Programming*, 157: 191–217, 2016.

[2] S. Ahmed. Convex relaxations of chance constrained optimization problems. *Optimization Letters*, 8(1): 1–12, 2014.

[3] S. Ahmed and D. J. Papageorgiou. Probabilistic set covering with correlations. *Operations Research*, 61(2): 438–452, 2013.

[4] S. Ahmed and A. Shapiro. Solving chance-constrained stochastic programs via sampling and integer programming, 2008. TutORials in Operations Research, INFORMS 2008.

[5] S. Ahmed and W. Xie. Relaxations and approximations of chance constraints under finite distributions. *Mathematical Programming*, 170(1):43–65, 2018.

[6] S. Ahmed, J. Luedtke, Y. Song, and W. Xie. Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. *Mathematical Programming*, 162:51–81, 2017.

[7] G. J. Alexander and A. M. Baptista. A comparison of var and cvar constraints on portfolio selection with the mean-variance model. *Management Science*, 50(9):1261–1273, 2004.

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

[9] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3): 203–228, 1999.

[10] A. Atamtürk. On the facets of the mixed-integer knapsack polyhedron. *Mathematical Programming*, 98(1-3): 145–175, 2003.

[11] A. Atamtürk and A. Gómez. Submodularity in conic quadratic mixed 0-1 optimization. *Operations Research*, 68:609–630, 2020.

[12] A. Atamtürk and V. Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008.

[13] A. Atamtürk and V. Narayanan. Conic mixed-integer rounding cuts. *Mathematical Programming*, 122:1–20, 2010.

[14] A. Atamtürk, G. L. Nemhauser, and M. W. Savelsbergh. The mixed vertex packing problem. *Mathematical Programming*, 89(1):35–53, 2000.

[15] E. Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.

[16] A. G. Banerjee, M. Ono, N. Roy, and B. C. Williams. Regression-based lp solver for chance-constrained finite horizon optimal control with nonconvex constraints. In *2011 American Control Conference (ACC)*, pages 131–138, 2011.

[17] J. Barrera, T. Homem-de Mello, E. Moreno, B. K. Pagnoncelli, and G. Canessa. Chance-constrained problems and rare events: an importance sampling approach. *Mathematical Programming*, 157(1):153–189, 2016.

[18] P. Barrieu and G. Scandolo. Assessing financial model risk. *European Journal of Operational Research*, 242 (2):546 – 556, 2015. ISSN 0377-2217.

[19] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4): 769–805, 1998.

[20] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, 1962.

[21] P. Beraldi and M. E. Bruni. An exact approach for solving integer problems under probabilistic constraints with random technology matrix. *Annals of Operations Research*, 177(1):127–137, 2009.

[22] P. Beraldi and A. Ruszczyński. A branch and bound method for stochastic integer programs under probabilistic constraints. *Optimization Methods and Software*, 17(3):359–382, 2002.

[23] P. Beraldi and A. Ruszczyński. The probabilistic set-covering problem. *Operations Research*, 50(6):956–967, 2002.

[24] D. Bienstock, M. Chertkov, and S. Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495, 2014.

[25] J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Verlag, New York, 1997.

[26] L. Blackmore, M. Ono, and B. C. Williams. Chance-constrained optimal path planning with obstacles. *IEEE Transactions on Robotics*, 27(6):1080–1094, 2011.

[27] J. Blanchet and K. Murthy. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research*, 44(2):565–600, 2019.

[28] C. E. Bonferroni. *Teoria statistica delle classi e calcolo delle probabilità*. Libreria internazionale Seeber, 1936.

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

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

[31] G. C. Calafiore and L. El Ghaoui. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130(1):1–22, 2006.

[32] A. M. Campbell and B. W. Thomas. Probabilistic traveling salesman problem with deadlines. *Transportation Science*, 42(1):1–21, 2008.

[33] M. C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

[34] M. C. Campi and S. Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of Optimization Theory and Applications*, 148:257–280, 2011.

[35] C. C. Carøe and J. Tind. A cutting-plane approach to mixed 0-1 stochastic integer programs. *European Journal of Operational Research*, 101(2):306–316, 1997.

[36] M. Castillo-Lopez, P. Ludivig, S. A. Sajadi-Alamdari, J. L. Sanchez-Lopez, M. A. Olivares-Mendez, and H. Voos. A real-time approach for chance-constrained motion planning with dynamic obstacles. *IEEE Robotics and Automation Letters*, 5(2):3620–3625, 2020.

[37] E. Çetinkaya and A. Thiele. Data-driven portfolio management with quantile constraints. *OR Spectrum*, 37 (3):761–786, 2015.

[38] A. Charnes and W. W. Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11(1):18–39, 1963.

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

[40] F. Chen and D. Krass. Inventory models with minimal service level constraints. *European Journal of Operational Research*, 134:120–140, 2001.

[41] W. Chen, M. Sim, J. Sun, and C.-P. Teo. From CVaR to uncertainty set: Implications in joint chance-constrained optimization. *Operations Research*, 58(2):470–485, 2010.

[42] X. Chen, M. Sim, and P. Sun. A robust optimization perspective on stochastic programming. *Operations Research*, 55(6):1058–1071, 2007.

[43] Z. Chen, D. Kuhn, and W. Wiesemann. Data-driven chance constrained programs over Wasserstein balls. *arXiv:1809.00210*, 2018.

[44] Z. Chen, M. Sim, and H. Xu. Distributionally robust optimization with infinitely constrained ambiguity sets. *Operations Research*, 67(5):1328–1344, 2019.

[45] J. Cheng, E. Delage, and A. Lisser. Distributionally robust stochastic knapsack problem. *SIAM Journal on Optimization*, 24(3):1485–1506, 2014.

[46] M. Cheon, S. Ahmed, and F. Al-Khayyal. A branch-reduce-cut algorithm for the global optimization of probabilistically constrained linear programs. *Mathematical Programming*, 108(2–3):617–634, 2006.

[47] B.-G. Choi, N. Rujeerapaiboon, and R. Jiang. Multi-period portfolio optimization: Translation of autocorrelation risk to excess variance. *Operations Research Letters*, 44(6):801–807, 2016.

[48] M. C. Cohen, P. W. Keller, V. Mirrokni, and M. Zadimoghaddam. Overcommitment in cloud services: Bin packing with chance constraints. *Management Science*, 65(7):3255–3271, 2019.

[49] J.-F. Cordeau, G. Laporte, M. W. Savelsbergh, and D. Vigo. Vehicle routing. In C. Barnhart and G. Laporte, editors, *Handbooks in operations research and management science*, volume 14, chapter 6, pages 367–428. Elsevier, 2007.

[50] F. E. Curtis, A. Wächter, and V. M. Zavala. A sequential algorithm for solving nonlinear optimization problems with chance constraints. *SIAM Journal on Optimization*, 28(1):930–958, 2018.

[51] M. d. S. Arantes, C. F. M. Toledo, B. C. Williams, and M. Ono. Collision-free encoding for chance-constrained nonconvex path planning. *IEEE Transactions on Robotics*, 35(2):433–448, 2019.

[52] E. Dall'Anese, K. Baker, and T. Summers. Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems*, 32(5):3427–3438, 2017.

[53] J. Danielsson, B. N. Jorgensen, C. G. de Vries, and X. Yang. Optimal portfolio allocation under the probabilistic VaR constraint and incentives for financial innovation. *Annals of Finance*, 4(3):1614–2446, 2008.

[54] A. De. Boolean function analysis meets stochastic optimization: An approximation scheme for stochastic knapsack. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1286–1305. SIAM, 2018.

[55] D. P. de Farias and B. Van Roy. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 29(3):462–478, 2004.

[56] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.

[57] Y. Deng and S. Shen. Decomposition algorithms for optimizing multi-server appointment scheduling with chance constraints. *Mathematical Programming*, 157(1):245–276, 2016.

[58] Y. Deng, S. Shen, and B. Denton. Chance-constrained surgery planning under conditions of limited and ambiguous data. *INFORMS Journal on Computing*, 31(3):559–575, 2019.

[59] D. Dentcheva. Optimization models with probabilistic constraints. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*. London, 2006.

[60] D. Dentcheva and A. Ruszczyński. Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566, 2003.

[61] D. Dentcheva, A. Prékopa, and A. Ruszczyński. Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming*, 89(1):55–77, 2000.

[62] S. Dharmadhikari and K. Joag-Dev. *Unimodality, Convexity, and Applications*. Elsevier, 1988.

[63] T. Dinh, R. Fukasawa, and J. Luedtke. Exact algorithms for the chance-constrained vehicle routing problem. *Mathematical Programming*, 172(1-2):105–138, 2018.

[64] B. Du, D. Sun, S. G. Manyam, and D. W. Casbeer. Cooperative air-ground vehicle routing using chance-constrained optimization. In *2020 American Control Conference (ACC)*, pages 392–397, 2020.

[65] N. E. Du Toit and J. W. Burdick. Probabilistic collision checking with chance constraints. *IEEE Transactions on Robotics*, 27(4):809–815, 2011.

[66] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu. Distributionally robust chance-constrained approximate AC-OPF with wasserstein metric. *IEEE Transactions on Power Systems*, 33(5):4924–4936, 2018.

[67] D. Duque, D. P. Morton, B. Singh, Z. Du, R. Pasco, and L. A. Meyers. Timing social distancing to avert unmanageable covid-19 hospital surges. *Proceedings of the National Academy of Sciences*, 117(33): 19873–19878, 2020.

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

[69] Ö. Elçi and N. Noyan. A chance-constrained two-stage stochastic programming model for humanitarian relief network design. *Transportation Research Part B: Methodological*, 108:55 – 83, 2018.

[70] Ö. Elçi, N. Noyan, and K. Bülbül. Chance-constrained stochastic programming under variable reliability levels with an application to humanitarian relief network design. *Computers & Operations Research*, 96:91 – 107, 2018.

[71] E. Erdoğan and G. Iyengar. Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107(1-2):37–61, 2005.

[72] M. Farrokhsiar and H. Najjaran. Unscented predictive motion planning of a nonholonomic system. In *2011 IEEE International Conference on Robotics and Automation*, pages 4480–4485, 2011.

[73] M. Fischetti and M. Monaci. Cutting plane versus compact formulations for uncertain (integer) linear programs. *Mathematical Programming Computation*, 4(3):239–273, 2012.

[74] A. M. Florio, R. F. Hartl, S. Minner, and J.-J. Salazar-González. A branch-and-price algorithm for the vehicle routing problem with stochastic demands and probabilistic duration constraints. *Transportation Science*, 55 (1):122–138, 2021.

[75] D. Gade, S. Küçükyavuz, and S. Sen. Decomposition algorithms with parametric Gomory cuts for two-stage stochastic integer programs. *Mathematical Programming*, 144(1-2):39–64, 2014.

[76] R. Gao and A. J. Kleywegt. Distributionally robust stochastic optimization with Wasserstein distance. *arXiv:1604.02199*, 2016.

[77] S. Geetha and K. P. K. Nair. On stochastic spanning tree problem. *Networks*, 23(8):675–679, 1993.

[78] A. Geletu, A. Hoffmann, M. Klöppel, and P. Li. An inner-outer approximation approach to chance constrained optimization. *SIAM Journal on Optimization*, 27(3):1834–1857, 2017.

[79] S. Ghosal and W. Wiesemann. The distributionally robust chance-constrained vehicle routing problem. *Operations Research*, 68(3):716–732, 2020.

[80] T. González Grandón, R. Henrion, and P. Pérlulliez-Aros. Dynamic probabilistic constraints under continuous random distributions. *Mathematical Programming*, 2020. doi: 10.1007/s10107-020-01593-z. Article in advance.

[81] V. Goyal and R. Ravi. Approximation algorithms for robust covering problems with chance constraints. Technical report, 2008. `https://kilthub.cmu.edu/ndownloader/files/12232889`.

[82] V. Goyal and R. Ravi. A PTAS for the chance-constrained knapsack problem with random item sizes. *Operations Research Letters*, 38(3):161–164, 2010.

[83] O. Günlük and Y. Pochet. Mixing mixed-integer inequalities. *Mathematical Programming*, 90(3):429–457, 2001.

[84] I. Gurvich, J. Luedtke, and T. Tezcan. Staffing Call Centers with Uncertain Demand Forecasts: A Chance-Constrained Optimization Approach. *Management Science*, 56(7):1093–1115, 2010.

[85] J. Han, K. Lee, C. Lee, K.-S. Choi, and S. Park. Robust optimization approach for a chance-constrained binary knapsack problem. *Mathematical Programming*, 157(1):277–296, 2016.

[86] G. A. Hanasusanto. *Decision Making under Uncertainty: Robust and Data-Driven Approaches*. PhD thesis, Imperial College London, 2015.

[87] G. A. Hanasusanto, V. Roitch, D. Kuhn, and W. Wiesemann. A distributionally robust perspective on uncertainty quantification and chance constrained programming. *Mathematical Programming*, 151(1):35–62, 2015.

[88] G. A. Hanasusanto, V. Roitch, D. Kuhn, and W. Wiesemann. Ambiguous joint chance constraints under mean and dispersion information. *Operations Research*, 65(3):751–767, 2017.

[89] R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computers & Mathematics with Applications*, 45(1):247 – 262, 2003.

[90] R. Henrion, P. Li, A. Möller, M. C. Steinbach, M. Wendt, and G. Wozny. Stochastic optimization for operating chemical processes under uncertainty. In M. Grötschel, S. O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems*, pages 457–478. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[91] N. Ho-Nguyen, F. Kılınç-Karzan, S. Küçükyavuz, and D. Lee. Distributionally robust chance-constrained programs with right-hand side uncertainty under Wasserstein ambiguity. *arXiv preprint arXiv:2003.12685*, 2020. To appear in *Mathematical Programming.*

[92] N. Ho-Nguyen, F. Kılınç-Karzan, S. Küçükyavuz, and D. Lee. Strong formulations for distributionally robust chance-constrained programs with left-hand side uncertainty under Wasserstein ambiguity. *arXiv preprint arXiv:2007.06750*, 2020.

[93] B. Hobbs. Optimization methods for electric utility resource planning. *European Journal of Operational Research*, 83(1):1–20, 1995.

[94] L. J. Hong, Y. Yang, and L. Zhang. Sequential convex approximations to joint chance constrained programs: A monte carlo approach. *Operations Research*, 59(3):617–630, 2011.

[95] X. Hong, M. A. Lejeune, and N. Noyan. Stochastic network design for disaster preparedness. *IIE Transactions*, 47(4):329–357, 2015.

[96] A. R. Hota, A. Cherukuri, and J. Lygeros. Data-driven chance constrained optimization under Wasserstein ambiguity sets. In *2019 American Control Conference (ACC)*, pages 1501–1506, July 2019. doi: 10.23919/ACC.2019.8814677.

[97] X. Huang and T. Zhao. Mean-chance model for portfolio selection based on uncertain measure. *Insurance: Mathematics and Economics*, 59:243 – 250, 2014.

[98] H. Ishii, S. Shiode, T. Nishida, and Y. Namasuya. Stochastic spanning tree problem. *Discrete Applied Mathematics*, 3(4):263–273, 1981.

[99] R. Ji and M. Lejeune. Data-driven distributionally robust chance-constrained optimization with Wasserstein metric. *SSRN Electronic Journal*, 2018. ISSN 1556-5068. doi: 10.2139/ssrn.3201356.

[100] R. Ji and M. A. Lejeune. Risk-budgeting multi-portfolio optimization with portfolio and marginal risk constraints. *Annals of Operations Research*, 262(2):547–578, 2018.

[101] N. Jiang and W. Xie. ALSO-X is better than CVaR: Convex approximations for chance constrained programs revisited, 2020.

[102] R. Jiang and Y. Guan. Data-driven chance constrained stochastic program. *Mathematical Programming*, 158: 291–327, 2016.

[103] S. Joung and K. Lee. Robust optimization-based heuristic algorithm for the chance-constrained knapsack problem using submodularity. *Optimization Letters*, 14(1):101–113, 2020.

[104] P. Kall and S. W. Wallace. *Stochastic Programming*. Wiley John & Sons, Chichester, 1994.

[105] S. Kataoka. A stochastic programming model. *Econometrica*, 31(1/2):181–196, 1963.

[106] F. Kılınç-Karzan, S. Küçükyavuz, and D. Lee. Joint chance-constrained programs and the intersection of mixing sets through a submodularity lens. *arXiv:1910.01353*, 2019.

[107] F. Kılınç-Karzan, S. Küçükyavuz, and D. Lee. Conic mixed-binary sets: Convex hull characterizations and applications. *arXiv:2012.14698*, 2020.

[108] W. K. Klein Haneveld. *On Integrated Chance Constraints*, pages 113–138. Springer Berlin Heidelberg, Berlin, Heidelberg, 1986.

[109] W. K. Klein Haneveld and M. H. van der Vlerk. Integrated chance constraints: Reduced forms and an algorithm. *Computational Management Science*, 3(4):245–269, 2006.

[110] O. Klopfenstein and D. Nace. A robust approach to the chance-constrained knapsack problem. *Operations Research Letters*, 36(5):628–632, 2008.

[111] A. Kogan and M. A. Lejeune. Threshold boolean form for joint probabilistic constraints with random technology matrix. *Mathematical Programming*, 147(1):391–427, 2014.

[112] A. Kogan, M. A. Lejeune, and J. Luedtke. Erratum to: Threshold boolean form for joint probabilistic constraints with random technology matrix. *Mathematical Programming*, 155(1):617–620, 2016.

[113] S. Küçükyavuz. On mixing sets arising in chance-constrained programming. *Mathematical Programming*, 132(1-2):31–56, 2012.

[114] S. Küçükyavuz and N. Noyan. Cut generation for optimization problems with multivariate risk constraints. *Mathematical Programming*, 159(1-2):165–199, 2016.

[115] S. Küçükyavuz and S. Sen. An introduction to two-stage stochastic mixed-integer programming. In R. Batta and J. Peng, editors, *TutORials in Operations Research: Leading Developments from INFORMS Communities*, chapter 1, pages 1–27. INFORMS, 2017.

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

[117] G. Laporte and F. V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133–142, 1993.

[118] J. B. Lasserre and T. Weisser. Distributionally robust polynomial chance-constraints under mixture ambiguity sets. *Mathematical Programming*, 2019. doi: 10.1007/s10107-019-01434-8. Article in advance.

[119] M. Lejeune. Pattern-based modeling and solution of probabilistically constrained optimization problems. *Operations Research*, 60(6):1356–1372, 2012.

[120] M. A. Lejeune and A. Ruszczyński. An efficient trajectory method for probabilistic inventory-production-distribution problems. *Operations Research*, 55(2):378–394, 2007.

[121] M. A. Lejeune and S. Shen. Multi-objective probabilistically constrained programs with variable risk: Models for multi-portfolio financial optimization. *European Journal of Operational Research*, 252(2):522 – 539.

[122] G. J. Lemus Rodriguez. *Portfolio optimization with quantile-based risk measures*. PhD thesis, Massachusetts Institute of Technology, 1999.

[123] B. Li, R. Jiang, and J. L. Mathieu. Distributionally robust chance constrained optimal power flow assuming log-concave distributions. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[124] B. Li, R. Jiang, and J. L. Mathieu. Ambiguous risk constraints with moment and unimodality information. *Mathematical Programming*, 173(1):151–192, Jan 2019.

[125] B. Li, R. Jiang, and J. L. Mathieu. Distributionally robust chance-constrained optimal power flow assuming unimodal distributions with misspecified modes. *IEEE Transactions on Control of Network Systems*, 6(3): 1223–1234, 2019.

[126] L. Li, H. Shao, R. Wang, and J. Yang. Worst-case range value-at-risk with partial information. *SIAM Journal on Financial Mathematics*, 9(1):190–218, 2018.

[127] Q. Li, A. M.-C. So, and W.-K. Ma. Distributionally robust chance-constrained transmit beamforming for multiuser miso downlink. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3479–3483. IEEE, 2014.

[128] W. W. Li, Y. J. Zhang, A. M. So, and M. Z. Win. Slow adaptive ofdma systems through chance constrained programming. *IEEE Transactions on Signal Processing*, 58(7):3858–3869, 2010.

[129] T. J. Linsmeier and N. D. Pearson. Value at risk. *Financial Analysts Journal*, 56(2):47–67, 2000.

[130] X. Liu and S. Küçükyavuz. A polyhedral study of the static probabilistic lot-sizing problem. *Annals of Operations Research*, 261(1-2):233–254, 2018.

[131] X. Liu, S. Küçükyavuz, and J. Luedtke. Decomposition algorithms for two-stage chance-constrained programs. *Mathematical Programming*, 157(1):219–243, 2016.

[132] X. Liu, S. Küçükyavuz, and N. Noyan. Robust multicriteria risk-averse stochastic programming models. *Annals of Operations Research*, 259(1):259–294, 2017.

[133] X. Liu, F. Kılınç-Karzan, and S. Küçükyavuz. On intersection of two mixing sets with applications to joint chance-constrained programs. *Mathematical Programming*, 175:29–68, 2019.

[134] Z. Liu, F. Wen, and G. Ledwich. Optimal siting and sizing of distributed generators in distribution systems considering uncertainties. *IEEE Transactions on Power Delivery*, 26(4):2541–2551, 2011.

[135] Z. Liu, Q. Wu, S. S. Oren, S. Huang, R. Li, and L. Cheng. Distribution locational marginal pricing for optimal electric vehicle charging through chance constrained mixed-integer programming. *IEEE Transactions on Smart Grid*, 9(2):644–654, 2018.

[136] A. Lodi, E. Malaguti, G. Nannicini, and D. Thomopulos. Nonlinear chance-constrained problems with applications to hydro scheduling. *Mathematical Programming*, 2019. doi: 10.1007/s10107-019-01447-3. Article in advance.

[137] A. Lodi, E. Malaguti, G. Nannicini, and D. Thomopulos. Nonlinear chance-constrained problems with applications to hydro scheduling. *Mathematical Programming*, pages 1–40, 2019. doi: 10.1007/s10107-019-01447-3. Article in advance.

[138] S. Lotfi and S. A. Zenios. Robust VaR and CVaR optimization under joint ambiguity in distributions, means, and covariances. *European Journal of Operational Research*, 269(2):556 – 576, 2018.

[139] L. Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art: Bonn 1982*, pages 235–257. Springer, Berlin, Heidelberg, 1983.

[140] M. Lu, H. Nakao, S. Shen, and L. Zhao. Non-profit resource allocation and service scheduling with cross-subsidization and uncertain resource consumptions. *Omega*, pages 102–191, 2020.

[141] M. Lubin, Y. Dvorkin, and S. Backhaus. A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Transactions on Power Systems*, 31(5):3840–3849, 2016.

[142] M. Lubin, Y. Dvorkin, and L. Roald. Chance constraints for improving the security of ac optimal power flow. *IEEE Transactions on Power Systems*, 34(3):1908–1917, 2019.

[143] J. Luedtke. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*, 146:219–244, 2014.

[144] J. Luedtke and S. Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

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

[146] G. Lulli and S. Sen. Branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. 50(6):786–796, 2004.

[147] S. Ma and D. Sun. Chance constrained robust beamforming in cognitive radio networks. *IEEE Communications Letters*, 17(1):67–70, 2013.

[148] V. Marianov and M. Rios. A probabilistic quality of service constraint for a location model of switches in ATM communication networks. *Annals of Operations Research*, 96:237–243, 2000.

[149] M. Mazadi, W. D. Rosehart, O. P. Malik, and J. A. Aguado. Modified chance-constrained optimization applied to the generation expansion problem. *IEEE Transactions on Power Systems*, 24(3):1635–1636, 2009.

[150] M. Meraklı and S. Küçükyavuz. Vector-valued multivariate conditional value-at-risk. *Operations Research Letters*, 46(3):300 – 305, 2018.

[151] M. Meraklı and S. Küçükyavuz. Risk aversion to parameter uncertainty in Markov decision processes with an application to slow-onset disaster relief. *IISE Transactions*, 52:811–831, 2020.

[152] B. L. Miller and H. M. Wagner. Chance constrained programming with joint constraints. *Operations Research*, 13(6):930–965, 1965.

[153] P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, Sep 2018.

[154] N. Mokari, S. Parsaeefard, P. Azmi, H. Saeedi, and E. Hossain. Robust ergodic uplink resource allocation in underlay ofdma cognitive radio networks. *IEEE Transactions on Mobile Computing*, 15(2):419–431, 2016.

[155] D. Moser, R. Schmied, H. Waschl, and L. del Re. Flexible spacing adaptive cruise control using stochastic model predictive control. *IEEE Transactions on Control Systems Technology*, 26(1):114–127, 2018.

[156] A. Muraleedharan, A. T. Tran, H. Okuda, and T. Suzuki. Scenario-based model predictive speed controller considering probabilistic constraint for driving scene with pedestrian. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–7, 2020.

[157] M. R. Murr and A. Prékopa. Solution of a product substitution problem using stochastic programming. In S. P. Uryasev, editor, *Probabilistic Constrained Optimization: Methodology and Applications*, pages 252–271. Springer US, Boston, MA, 2000.

[158] A. Najjarbashi and G. J. Lim. A decomposition algorithm for the two-stage chance-constrained operating room scheduling problem. *IEEE Access*, 8:80160–80172, 2020.

[159] K. Natarajan, D. Pachamanova, and M. Sim. Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science*, 54(3):573–585, 2008.

[160] A. Nemirovski. On safe tractable approximations of chance constraints. *European Journal of Operational Research*, 219(3):707 – 718, 2012.

[161] A. Nemirovski and A. 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, 2005.

[162] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2007.

[163] T. Niimura, B. S. Kermanshahi, and R. Yokoyama. Multi-stage optimization of generation planning including power system reliability constraints. *International Journal of Energy Systems*, 10(3):144–148, 1990.

[164] E. Nikolova. Approximation algorithms for reliable stochastic combinatorial optimization. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 338–351. Springer, 2010.

[165] N. Noyan and G. Rudolf. Optimization with multivariate conditional value-at-risk constraints. *Operations Research*, 61(4):990–1013, 2013.

[166] N. Noyan, M. Meraklı, and S. Küçükyavuz. Two-stage stochastic programming under multivariate risk constraints with an application to humanitarian relief network design. *Mathematical Programming*, pages 1–39, 2019. doi: 10.1007/s10107-019-01373-4. Article in advance.

[167] L. Ntaimo. Fenchel decomposition for stochastic mixed-integer programming. *Journal of Global Optimization*, 55(1):141–163, 2013.

[168] L. Ntaimo and S. Sen. The million variable “march” for stochastic combinatorial optimization. *Journal of Global Optimization*, 32:385–400, 2005.

[169] L. Ntaimo and S. Sen. A comparative study of decomposition algorithms for stochastic combinatorial optimization. *Computational Optimization and Applications*, 40:299–319, 2008.

[170] Y. Oh, K. Cho, Y. Choi, and S. Oh. Chance-constrained multi-layered sampling-based path planning for temporal logic-based missions. *IEEE Transactions on Automatic Control*, pages 1–15, 2020. doi: 10.1109/TAC.2020.3044273. Article in advance.

[171] U. A. Ozturk, M. Mazumdar, and B. A. Norman. A solution to the stochastic unit commitment problem using chance constrained programming. *IEEE Transactions on Power Systems*, 19(3):1589–1598, 2004.

[172] M. Padberg and G. Rinaldi. A branch-and-cut approach to a traveling salesman problem with side constraints. *Management Science*, 35(11):1393–1412, 1989.

[173] B. K. Pagnoncelli, S. 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.

[174] S. Pelletier, O. Jabali, and G. Laporte. The electric vehicle routing problem with energy consumption uncertainty. *Transportation Research Part B: Methodological*, 126:225–255, 2019.

[175] A. Peña-Ordieres, J. R. Luedtke, and A. Wächter. Solving chance-constrained problems via a smooth sample-based nonlinear approximation. *SIAM Journal on Optimization*, 30(3):2221–2250, 2020.

[176] G. C. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In S. Uryasev, editor, *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Dordrecht, 2000.

[177] G. C. Pflug and W. Römisch. *Modelling, managing and measuring risk*. World Scientific publishing, Singapore, 2007.

[178] J. Pintér. Deterministic approximations of probability inequalities. *Zeitschrift för Operations-Research*, 33 (4):219–239, 1989.

[179] Y. Pochet and L. A. Wolsey. Polyhedra for lot-sizing with Wagner-Whitin costs. *Mathematical Programming*, 67:297–323, 1994.

[180] K. Postek, A. Ben-Tal, D. den Hertog, and B. Melenberg. Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research*, 66(3):814–833, 2018.

[181] D. Pozo and J. Contreras. A chance-constrained unit commitment with an $n - k$ security criterion and significant wind generation. *IEEE Transactions on Power Systems*, 28(3):2842–2851, 2013.

[182] A. Prékopa. On probabilistic constrained programming. In *Proceedings of the Princeton Symposium on Mathematical Programming*, volume 113, page 138. Princeton, NJ, 1970.

[183] A. Prékopa. Contributions to the theory of stochastic programming. *Mathematical Programming*, 4(1): 202–221, 1973.

[184] A. Prékopa. Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution. *ZOR Zeitschrift für Operations Research Methods and Models of Operations Research*, 34(6):441–461, 11 1990.

[185] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[186] A. Prékopa. Probabilistic programming. In A. Ruszczyński and A. Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 267–351. Elsevier, 2003.

[187] Y. Qi and S. Sen. The ancestral Benders' cutting plane algorithm with multi-term disjunctions for mixed-integer recourse decisions in stochastic programming. *Mathematical Programming*, 161:193–235, 2017.

[188] F. Qiu and J. Wang. Chance-constrained transmission switching with guaranteed wind power utilization. *IEEE Transactions on Power Systems*, 30(3):1270–1278, 2015.

[189] F. Qiu, S. Ahmed, S. S. Dey, and L. A. Wolsey. Covering linear programming with violations. *INFORMS Journal on Computing*, 26(3):531–546, 2014.

[190] H. Rahimian and S. Mehrotra. Distributionally robust optimization: A review. *arXiv:1908.05659*, 2019.

[191] A. Ravichandran, S. Sirouspour, P. Malysz, and A. Emadi. A chance-constraints-based control strategy for microgrids with energy storage and integrated electric vehicles. *IEEE Transactions on Smart Grid*, 9(1): 346–359, 2018.

[192] R. Rockafellar. Coherent approaches to risk in optimization under uncertainty. In *TutORrials in Operations Research*, pages 38–61. INFORMS, 2007.

[193] R. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3):21–41, 2000.

[194] R. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002.

[195] N. Rujeerapaiboon, D. Kuhn, and W. Wiesemann. Robust growth-optimal portfolios. *Management Science*, 62(7):2090–2109, 2016.

[196] A. Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93(2):195–215, 2002.

[197] A. Saxena, V. Goyal, and M. A. Lejeune. MIP reformulations of the probabilistic set covering problem. *Mathematical Programming*, 121(1):1–31, 2010.

[198] S. Sen. Relaxations for probabilistically constrained programs with discrete random variables. *Operations Research Letters*, 11(2):81–86, 1992.

[199] S. Sen. *Stochastic Integer Programming Algorithms: Beyond Benders' Decomposition*. Wiley Handbook in OR/MS, World Wide Web, 2010.

[200] S. Sen and J. L. Higle. The C^3 theorem and a D^2 algorithm for large scale stochastic mixed-integer programming: set convexification. *Mathematical Programming*, 104(1):1–20, 2005.

[201] S. Sen and H. D. Sherali. On the convergence of cutting plane algorithms for a class of nonconvex mathematical programs. *Mathematical Programming*, 106(2):203–223, 2006.

[202] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[203] H. Shen and R. Jiang. Chance-constrained set covering with Wasserstein ambiguity. Technical report, 2020. https://arxiv.org/abs/2010.05671.

[204] S. Shen. Using integer programming for balancing return and risk in problems with individual chance constraints. *Computers & Operations Research*, 49:59 – 70, 2014.

[205] N. Y. Soltani, S. Kim, and G. B. Giannakis. Chance-constrained optimization of ofdma cognitive radio uplinks. *IEEE Transactions on Wireless Communications*, 12(3):1098–1107, 2013.

[206] Y. Song and J. R. Luedtke. Branch-and-cut approaches for chance-constrained formulations of reliable network design problems. *Mathematical Programming Computation*, 5(4):397–432, 2013.

[207] Y. Song and S. Shen. Risk-averse shortest path interdiction. *INFORMS Journal on Computing*, 28(3): 527–539, 2016.

[208] Y. Song, J. R. Luedtke, and S. Küçükyavuz. Chance-constrained binary packing problems. *INFORMS Journal on Computing*, 26(4):735–747, 2014.

[209] B. Stellato. Data-driven chance constrained optimization. Master's thesis, ETH Zürich, 2014.

[210] C. 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.

[211] A. K. Takyi and B. J. Lence. Surface water quality management using a multiple-realization chance constraint method. *Water Resources Research*, 35(5):1657–1670, 1999.

[212] M. W. Tanner and L. Ntaimo. IIS branch-and-cut for joint chance-constrained stochastic programs and application to optimal vaccine allocation. *European Journal of Operational Research*, 207(1):290 – 296, 2010.

[213] S. R. Tayur, R. R. Thomas, and N. R. Natraj. An algebraic geometry algorithm for scheduling in presence of setups and correlated demands. *Mathematical Programming*, 69(1–3):369–401, 1995.

[214] W. Van Ackooij, R. Henrion, A. Moller, and R. Zorgati. Chance constrained programming and its applications to energy management. In I. Dritsas, editor, *Stochastic Optimization - Seeing the Optimal for the Uncertain*. IntechOpen, 2011.

[215] R. M. Van Slyke and R. Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.

[216] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Review*, 49(1):52–64, 2007.

[217] J. Vielma, S. Ahmed, and G. Nemhauser. Mixed integer linear programming formulations for probabilistic constraints. *Operations Research Letters*, 40(3):153 – 158, 2012.

[218] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. A probabilistic framework for reserve scheduling and $\mathrm{n}-1$ security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems*, 28(4):3885–3896, 2013.

[219] M. R. Wagner. Stochastic 0–1 linear programming under limited distributional information. *Operations Research Letters*, 36(2):150–156, 2008.

[220] J. Wang. The β-reliable median on a network with discrete probabilistic demand weights. *Operations research*, 55(5):966–975, 2007.

[221] J. Wang and S. Shen. Risk and energy consumption tradeoffs in cloud computing service via stochastic optimization models. In *Proceedings of the 5th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2012)*, Chicago, IL, 2012.

[222] Q. Wang, Y. Guan, and J. Wang. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Transactions on Power Systems*, 27(1):206–215, 2012.

[223] S. Wang, J. Li, and C. Peng. Distributionally robust chance-constrained program surgery planning with downstream resource. In *2017 International Conference on Service Systems and Service Management*, pages 1–6. IEEE, 2017.

[224] S. Wang, J. Li, and S. Mehrotra. A solution approach to distributionally robust chance-constrained assignment problems. Technical report, 2019. http://www.optimization-online.org/DB_HTML/2019/05/7207.html.

[225] S. Wang, J. Li, and S. Mehrotra. Chance-constrained multiple bin packing problem with an application to operating room planning. Technical report, 2019. http://www.optimization-online.org/DB_HTML/2019/02/7053.html.

[226] R. J.-B. Wets. Stochastic programming: Solution techniques and approximation schemes. In B. A., K. B., and G. M., editors, *Mathematical Programming The State of the Art*, Handbooks in Operations Research and Management Science, pages 566–603. Springer, 1983.

[227] R. J.-B. Wets. Stochastic programming. In *Optimization*, volume 1 of *Handbooks in Operations Research and Management Science*, chapter VIII, pages 573 – 629. Elsevier, 1989.

[228] H. Wu and S. Küçükyavuz. Probabilistic partial set covering with an oracle for chance constraints. *SIAM Journal on Optimization*, 29(1):690–718, 2019.

[229] H. Wu, M. Shahidehpour, Z. Li, and W. Tian. Chance-constrained day-ahead scheduling in stochastic power system operation. *IEEE Transactions on Power Systems*, 29(4):1583–1591, 2014.

[230] J. Wu, J. Zhu, G. Chen, and H. Zhang. A hybrid method for optimal scheduling of short-term electric power generation of cascaded hydroelectric plants based on particle swarm optimization and chance-constrained programming. *IEEE Transactions on Power Systems*, 23(4):1570–1579, 2008.

[231] P. Wu, J. Xie, and J. Chen. Safe path planning for unmanned aerial vehicle under location uncertainty. In *2020 IEEE 16th International Conference on Control Automation (ICCA)*, pages 342–347, 2020.

[232] W. Xie. On distributionally robust chance constrained programs with Wasserstein distance. *Mathematical Programming*, 2019. doi: 10.1007/s10107-019-01445-5. Article in advance.

[233] W. Xie and S. Ahmed. On deterministic reformulations of distributionally robust joint chance constrained optimization problems. *SIAM Journal on Optimization*, 28(2):1151–1182, 2016.

[234] W. Xie and S. Ahmed. Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Transactions on Power Systems*, 33(2):1860–1867, 2018.

[235] W. Xie and S. Ahmed. On quantile cuts and their closure for chance constrained optimization problems. *Mathematical Programming*, 172:621–646, 2018.

[236] W. Xie and S. Ahmed. Bicriteria approximation of chance-constrained covering problems. *Operations Research*, 68(2):516–533, 2020.

[237] W. Xie, S. Ahmed, and R. Jiang. Optimized Bonferroni approximations of distributionally robust joint chance constraints. *Mathematical Programming*, 2019. doi: 10.1007/s10107-019-01442-8. Article in advance.

[238] H. Xu, C. Caramanis, and S. Mannor. Optimization under probabilistic envelope constraints. *Operations Research*, 60(3):682–699, 2012.

[239] L. Xu and A. Nallanathan. Energy-efficient chance-constrained resource allocation for multicast cognitive ofdm network. *IEEE Journal on Selected Areas in Communications*, 34(5):1298–1306, 2016.

[240] I. Yang. Wasserstein distributionally robust stochastic control: A data-driven approach. *IEEE Transactions on Automatic Control*, pages 1–8, 2020. doi: 10.1109/TAC.2020.3030884. Article in advance.

[241] W. Yang and H. Xu. Distributionally robust chance constraints for non-linear uncertainties. *Mathematical Programming*, 155(1-2):231–265, 2016.

[242] H. Yao, Y. Li, and K. Benson. A smooth non-parametric estimation framework for safety-first portfolio optimization. *Quantitative Finance*, 15(11):1865–1884, 2015.

[243] K. Yoda and A. Prékopa. Convexity and solutions of stochastic multidimensional 0-1 knapsack problems with probabilistic constraints. *Mathematics of Operations Research*, 41(2):715–731, 2016.

[244] H. Zhang and P. Li. Chance constrained programming for optimal power flow under uncertainty. *IEEE Transactions on Power Systems*, 26(4):2417–2424, 2011.

[245] M. Zhang and S. Küçükyavuz. Finitely convergent decomposition algorithms for two-stage stochastic pure integer programs. *SIAM Journal on Optimization*, 24(4):1933–1951, 2014.

[246] M. Zhang, S. Küçükyavuz, and S. Goel. A branch-and-cut method for dynamic decision making under joint chance constraints. *Management Science*, 60(5):1317–1333, 2014.

[247] Y. Zhang, S. Shen, and J. L. Mathieu. Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. *IEEE Transactions on Power Systems*, 32(2): 1378–1388, 2017.

[248] Y. Zhang, R. Jiang, and S. Shen. Ambiguous chance-constrained binary programs under mean-covariance information. *SIAM Journal on Optimization*, 28(4):2922–2944, 2018.

[249] Y. Zhang, S. Shen, and S. A. Erdogan. Solving 0–1 semidefinite programs for distributionally robust allocation of surgery blocks. *Optimization Letters*, 12(7):1503–1521, 2018.

[250] Y. Zhang, J. Dong, T. Kuruganti, S. Shen, and Y. Xue. Distributionally robust building load control to compensate fluctuations in solar power generation. In *2019 American Control Conference (ACC)*, pages 5857–5863. IEEE, 2019.

[251] Y. Zhang, M. Lu, and S. Shen. On the values of vehicle-to-grid electricity selling in electric vehicle sharing. *Manufacturing & Service Operations Management*, 2020. doi: 10.1287/msom.2019.0855. Article in advance.

[252] Z. Zhang, B. T. Denton, and X. Xie. Branch and price for chance-constrained bin packing. *INFORMS Journal on Computing*, 32(3):547–564, 2020.

[253] M. Zhao, K. Huang, and B. Zeng. A polyhedral study on chance constrained program with random right-hand side. *Mathematical Programming*, 166:19–64, 2017.

[254] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2011.

[255] S. Zymler, D. Kuhn, and B. Rustem. Worst-case value at risk of nonlinear portfolios. *Management Science*, 59(1):172–188, 2013.