On Multidimensional Disjunctive Inequalities for Chance-Constrained Stochastic Problems with Finite Support

Diego Cattaruzza, Martine Labbé, Matteo Petris, Marius Roland, Martin Schmidt

Abstract. We consider mixed-integer linear chance-constrained problems for which the random vector that parameterizes the feasible region has finite support. Our key objective is to improve branch-and-bound or -cut approaches by introducing new types of valid inequalities that improve the dual bounds and, by this, the overall performance of such methods. We introduce so-called primal-dual as well as covering valid inequalities. By re-scaling the latter inequalities, we obtain so-called multi-disjunctive valid inequalities, which generalize known inequalities from the literature. We provide theoretical results regarding dominance relations, closure properties, and hardness of the separation problems. Given these insights, we propose heuristic separation procedures and present extensive numerical results showing the effectiveness of our method in comparison to state-of-the-art inequalities from the literature.

1. Introduction

Chance-constrained mathematical programs (CCPs) are a particular class of mathematical programs, which can be stated, in their general form, as

$$\min_{x} \quad f(x) \tag{1a}$$
$$\text{s.t.} \quad \mathbb{P}(x \in P(\omega) \subset \mathbb{R}^n) \geq 1 - \tau, \tag{1b}$$
$$x \in X \subset \mathbb{R}^n, \tag{1c}$$

where $f(x)$ is the objective function to optimize defined over a vector of variables $x \in \mathbb{R}^n$. Moreover, X is a set of deterministic constraints and $P(\omega)$ is another feasible region parameterized by a random vector ω. Expression (1b) translates the fact that the outcome $x \notin P(\omega)$ is an undesirable event, i.e., it may be accepted only if the probability that it happens is lower than a confidence value $\tau \in (0, 1)$, which is usually chosen close to 0. The study of CCPs dates back to the 1950s and the seminal works [10, 11, 31].

In this paper, we consider the case of a linear objective function and the parameterized feasible region being given as a parameterized polyhedron. The latter is expressed as a set of linear inequalities with a matrix of random constraint coefficients and a vector of random right-hand sides, hence both are expressions of the random vector ω. We assume the latter has finite support.

Date: May 26, 2025.

2020 *Mathematics Subject Classification.* 90C11, 90C15, 90C57.

Key words and phrases. Chance constraints, Stochastic optimization, Mixed-integer linear optimization, Valid inequalities, Closures.

Formally, we consider the (mixed-integer) linear chance-constrained problem

$$\min_{x} \quad c^\top x \tag{2a}$$
$$\text{s.t.} \quad x \in X, \tag{2b}$$
$$\mathbb{P}(A^\omega x - b^\omega \geq 0) \geq 1 - \tau, \tag{2c}$$

where $X \subset \mathbb{R}^n$ is a non-empty and compact set defined by deterministic constraints, which possibly include integrality restrictions on some or all of the variables. The cost vector is given by $c \in \mathbb{R}^n$. Let $(\Omega, 2^\Omega, \mathbb{P})$ be a discrete and finite probability space such that $\Omega = \{\omega^s : s \in \mathcal{S}\}$ and $p^s = \mathbb{P}(\omega = \omega^s) \geq 0$ for $s \in \mathcal{S}$ and $\sum_{s\in\mathcal{S}} p^s = 1$ holds. We consider the cases of finitely many scenarios, i.e., we have $|\mathcal{S}| < +\infty$. Moreover, let $A^\omega \in \mathbb{R}^{m\times n}$ be a matrix of random constraint coefficients and let $b^\omega \in \mathbb{R}^m$ be a vector of random right-hand sides. Finally, τ is the risk tolerance.

Formulation (2) can be rewritten to obtain a scenario-based formulation by introducing a vector $y \in \{0,1\}^{|\mathcal{S}|}$ of binary variables and, for each scenario $s \in \mathcal{S}$, a vector $\bar{b}^s \in \mathbb{R}^m$ of suitable values. With this at hand, we can re-write (2) as

$$\min_{x,y} \quad c^\top x \tag{3a}$$
$$\text{s.t.} \quad x \in X, \tag{3b}$$
$$A^s x \geq \bar{b}^s + (b^s - \bar{b}^s) y^s, \quad s \in \mathcal{S}, \tag{3c}$$
$$\sum_{s\in\mathcal{S}} p^s y^s \geq 1 - \tau, \tag{3d}$$
$$y^s \in \{0,1\}, \quad s \in \mathcal{S}. \tag{3e}$$

Given a scenario $s \in \mathcal{S}$, the corresponding constraint (3c) is active if and only if the respective binary variable y_s is equal to 1. Otherwise, the constraint is inactive as the choice of $\bar{b}^s$ is such that the corresponding constraint is satisfied for all $x \in X$. Constraint (3d) imposes that the sum of the probabilities over the selected scenarios is at least $1 - \tau$. Said differently, only constraints associated to subsets of scenarios so that their probability sums up to a value lower than τ are allowed to be violated. Finally, constraints (3e) impose binary conditions on the y variables.

In what follows, for $a < b$, we will use the notation $[a : b] := \{a, \ldots, b\} \subset \mathbb{Z}$. Moreover, given $\bar{\mathcal{S}} \subset \mathcal{S}$, we will abbreviate $p(\bar{\mathcal{S}}) = \sum_{s\in\bar{\mathcal{S}}} p^s$.

Outline and Contributions. The organization and the contributions of the paper are illustrated in the following. Section 2 reviews the related and relevant literature. Section 3 is dedicated to the introduction of two families of valid inequalities for the special case of CCPs that we consider in this paper. One family of inequalities is inspired from the observation that the chance constraint in the CCPs is nothing but the $(1-\tau)$-quantile of a probability distribution. Taking inspiration from the proof of validity of such inequalities, we coin them as *primal-dual* valid inequalities. The second family of inequalities is called *covering* valid inequalities, from which we derive, when re-scaling is possible, what we call *multi-disjunctive* valid inequalities. The latter are a generalization of the simple-disjunctive valid inequalities proposed in [34]. We provide dominance relations among the different families of valid inequalities and show that covering valid inequalities dominate primal-dual valid inequalities. In Section 4, we show that the problem of separating covering valid inequalities is NP-hard and we devise heuristic procedures to find violated cuts. Section 5 presents results on the closure of the multi-disjunctive valid inequalities with respect to the closure of the so-called quantile cuts valid for CCPs; see also [28, 34, 42]. Finally, we conduct numerical experiments to compare the impact of the valid inequalities and also compare our inequalities with those from the literature. These numerical results are presented in Section 6. Afterward, Section 7 concludes the paper.

2. Literature Review

Problem (2) is shown to be NP-hard even in the special case in which all scenarios are equiprobable, $X = \mathbb{R}^n_+$, A^ω equals the identity matrix (and it thus does not depend on the parameter ω) and $c = (1, \dots, 1)^\top \in \mathbb{R}^m$; see [28].

The problem formulation (2) is characterized by the chance constraint (2c), which enforce a probabilistic requirement over all constraints defined by the random rows of the constraint matrix A^ω and the random values of the b^ω vector of right-hand sides. This setting is referred to in the literature as joint chance constraints. This is in contrast with single CCPs in which the chance constraint requires that each constraint implied by A^ω and b^ω is satisfied with a high probability. Both single and joint CCPs have been extensively studied in the literature.

One may note that the chance-constraint (2c) is nothing but the value-at-risk $\mathrm{VaR}_{1-\tau}[b^\omega - A^\omega]$, which is known of being a nonconvex function in the general case. Note that the VaR is a risk measure that is used in various domains. One prominent example is portfolio optimization. For instance, [14, 15] consider the problem of maximizing the VaR subject to a lower bound on the expected return while [3–5, 26] maximize the expected return given a lower bound on the VaR. Due to the VaR being nonconvex, CCPs have a nonconvex feasible region in general, which makes their resolution particularly challenging. Convex approximations of the chance constraint (2c) have been studied in, e.g., [21, 32, 35]. Note, in particular, that a conservative convex approximation of the chance constraint (2c) is the conditional value-at-risk $\mathrm{CVaR}_{1-\tau}[b^\omega - A^\omega]$.

One common assumption is to consider finite discrete probability distributions, which are represented by a finite set of scenarios, each associated with a given probability. The reader is referred to [24] for a recent survey and a taxonomy of CCPs that arise from finite discrete distributions, as it is the case for the problem tackled in this paper. While also other approaches exist, see, e.g., [13, 25], in the case of finite distributions, Problem (2) can be re-formulated to obtain a mixed-integer formulation as the one in (3). Such reformulations are obtained by introducing as many binary variables as scenarios, having the role to (de-)activate scenarios in order to satisfy the chance constraint (2c); see, e.g., [2]. Such formulations are known to suffer from poor continuous relaxations as they need the introduction of big-M values. To mitigate this effect, techniques to strengthen the big-M values have been considered in [33, 34, 40]. However, getting the tightest possible value may be very time consuming as it results in solving several chance-constrained problems. To circumvent this drawback, [28] proposed a decomposition-based method that avoids introducing big-Ms.

To overcome the poor quality of the linear relaxation of the mixed-integer reformulation (3) with big-M values, several authors have proposed valid inequalities with the aim to obtain stronger formulations. A first family of valid inequalities are the so-called mixing-set inequalities. The authors of [23] consider a CCP as in (2) but with a deterministic constraint matrix. In [23] it is noted that such CCPs contain the mixing set [20] as a sub-structure. This is used to derive valid CCP inequalities. Moreover, in [34], the authors study mixing-set inequalities that arise in the special case of a covering-type constraint matrix of dimension $1 \times n$ for each scenario. They also present results on the closure of such inequalities. A second family of valid inequalities are the so-called quantile cuts, which are based on the relation between the chance-constrained Problem (2c) and the definition of a quantile. They are studied, among others, in [28] and [42]. The latter papers also provide results about the closure of these valid inequalities.

Recently, [36] proposed an exact method based on scenario reduction for CCPs with finite support.

Finally, one may note that CCPs have been applied in different contexts such as generic set-covering problems [7, 38], bin-packing problems [40], transportation problems [29], multi-dimensional knapsack problems [1], and lot-sizing problems [6, 23, 43] as well as in areas such as reliable network design problems [39], multi-area reserve sizing for the Nordic countries [12], call-center staffing [28], or optimal vaccine allocation [41].

3. Valid Inequalities

In this section, we present two families of inequalities that are valid for Problem (3). In Section 3.1, we introduce what we call the primal-dual valid inequalities (PD-VIs), while in Section 3.2, we present what we coin the multi-disjunctive valid inequalities (MD-VIs). The latter are a generalization of the simple disjunctive cuts proposed in [34].

3.1. **Primal-Dual Valid Inequalities.** Given the formulation (3), one can take a convex combination of all the constraints in each scenario $s \in \mathcal{S}$. In particular, let us consider the m-dimensional unit simplex $\Lambda^m = \{\lambda \in \mathbb{R}^m_{\geq 0} \colon \mathbb{1}^\top \lambda = 1\}$ and coefficient vectors $\lambda^s \in \Lambda^m$ for all $s \in \mathcal{S}$. The requirement $\mathbb{1}^\top \lambda^s = 1$ is used to prevent from unboundedness issues, which we may encounter otherwise; see Appendix B. With this, we can obtain the relaxed set of constraints

$$x \in X, \tag{4a}$$

$$(\lambda^s)^\top A^s x \geq (\lambda^s)^\top \bar{b}^s + (\lambda^s)^\top (b^s - \bar{b}^s) y^s, \quad s \in \mathcal{S}, \tag{4b}$$

$$\sum_{s \in \mathcal{S}} p^s y^s \geq 1 - \tau, \tag{4c}$$

$$y^s \in \{0, 1\}, \quad s \in \mathcal{S}. \tag{4d}$$

Each convex combination in (4b) can be seen as a realization of a random variable for which we can calculate the $(1-\tau)$-quantile as presented next. Let us now define

$$q = \mathbb{Q}_{1-\tau} \left[(\lambda^\omega)^\top (b^\omega - A^\omega x) \right]$$

to be the $(1-\tau)$-quantile of $\{(\lambda^s)^\top (b^s - A^s x) \colon s \in \mathcal{S}\}$. We have that

$$q = \min \left\{ \bar{q} \in \mathbb{R} \colon \sum_{s \in \mathcal{N}(\bar{q}, x)} p^s \geq 1 - \tau \right\}$$

holds with

$$\mathcal{N}(\bar{q}, x) = \left\{ s \in \mathcal{S} \colon \bar{q} \geq (\lambda^s)^\top (b^s - A^s x) \right\}.$$

This can be used to prove the correctness of the following valid inequalities for formulation (3). Before, we introduce the following lemma.

Lemma 1. *It holds $q \leq 0$ if and only if (4) is feasible.*

Proof. Let us first suppose that $q \leq 0$ holds. This means that there exists $\bar{q} \leq 0$ such that $\sum_{s \in \mathcal{N}(\bar{q}, x)} p^s \geq 1 - \tau$ and $\bar{q} \geq (\lambda^s)^\top b^s - (\lambda^s)^\top A^s x$ holds for all $s \in \mathcal{N}(\bar{q}, x)$. Thus, $0 \leq (\lambda^s)^\top A^s x - (\lambda^s)^\top b^s$ holds for all $s \in \mathcal{N}(\bar{q}, x)$. The claim follows by setting $y^s = 1$ for all scenarios $s \in \mathcal{N}(\bar{q}, x)$.

For the other direction, let us suppose that (4) is feasible and let $(\bar{x}, \bar{y})$ be a feasible point. Now, w.l.o.g., we assume that $\bar{y}^1 = \cdots = \bar{y}^{\bar{s}} = 1$, $\bar{y}^{\bar{s}+1} = \cdots = \bar{y}^{|\mathcal{S}|} = 0$ holds. By setting $\mathcal{N}(\bar{q}, x) = \{1, 2, \ldots, \bar{s}\}$, the claim follows. □

Proposition 1. *Consider $\bar{\mathcal{S}} \subset \mathcal{S}$ with $p(\bar{\mathcal{S}}) \leq 1-\tau$ and vectors $\lambda^s \in \Lambda^m$, $s \in \mathcal{S}$. Then, the inequality*

$$\sum_{s\in\bar{\mathcal{S}}} p^s (\lambda^s)^\top (A^s x - b^s) - \sum_{i=1}^{n} U_i(\lambda) x_i + L(\lambda) \leq 0 \tag{5}$$

is a valid for Problem (3), where $U_i(\lambda)$, $i = 1,\dots,n$, and $L(\lambda)$ are solutions to the optimization problems

$$\begin{aligned} U_i(\lambda) = \max_{z} \quad & \sum_{s\in\mathcal{S}} (\lambda^s)^\top A^s_{.i} z^s \\ s.t. \quad & \sum_{s\in\mathcal{S}} z^s = 1-\tau, && (\alpha_i) \\ & 0 \leq z^s \leq p^s, \quad s \in \mathcal{S}, && (\beta_i^s) \end{aligned}$$

and

$$\begin{aligned} L(\lambda) = \min_{z} \quad & \sum_{s\in\mathcal{S}} (\lambda^s)^\top b^s z^s \\ s.t. \quad & \sum_{s\in\mathcal{S}} z^s = 1-\tau, && (\gamma) \\ & 0 \leq z^s \leq p^s, \quad s \in \mathcal{S}. && (\delta^s) \end{aligned}$$

Proof. Let Δ be a discrete random variable with realizations d^s, $s \in \mathcal{S}$. It is well known, see, e.g., [30], that the $(1-\tau)$-quantile $-q$ of the random variable $-\Delta$ is the optimal solution to the linear problem

$$\begin{aligned} \min_{q,w} \quad & (1-\tau)q + \sum_{s\in\mathcal{S}} w^s p^s \\ \text{s.t.} \quad & q + w^s \geq d^s, \quad s \in \mathcal{S}, \\ & w^s \geq 0, \quad s \in \mathcal{S}, \\ & q \in \mathbb{R}. \end{aligned}$$

The dual of this problem reads

$$\begin{aligned} \max_{z} \quad & \sum_{s\in\mathcal{S}} d^s z^s \\ \text{s.t.} \quad & \sum_{s\in\mathcal{S}} z^s = 1-\tau, \\ & 0 \leq z^s \leq p^s, \quad s \in \mathcal{S}. \end{aligned}$$

By strong duality, we have

$$(1-\tau)q + \sum_{s\in\mathcal{S}} w^s p^s \leq \sum_{s\in\mathcal{S}} d^s z^s.$$

Moreover, we have that $w^s \geq \max\{0, d^s - q\}$ holds for all $s \in \mathcal{S}$. It then follows that for all $\bar{\mathcal{S}} \subset \mathcal{S}$ with $p(\bar{\mathcal{S}}) \leq 1-\tau$, we have

$$(1-\tau)q + \sum_{s\in\bar{\mathcal{S}}} p^s (d^s - q) \leq \sum_{s\in\mathcal{S}} d^s z^s,$$

which is equivalent to

$$\left(1-\tau-p(\bar{\mathcal{S}})\right) q + \sum_{s\in\bar{\mathcal{S}}} p^s d^s \leq \sum_{s\in\mathcal{S}} d^s z^s.$$

In our framework, we have that $d^s = (\lambda^s)^\top (A^s x - b^s)$ holds for all $s \in \mathcal{S}$ and we thus obtain

$$(1 - \tau - p(\bar{\mathcal{S}}))q + \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top (A^s x - b^s) \leq \sum_{s \in \mathcal{S}} (\lambda^s)^\top (A^s x - b^s) z^s.$$

By Lemma 1, we can suppose $q \geq 0$ and we thus get

$$\begin{aligned} &\sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top (A^s x - b^s) \\ \leq\ &\sum_{s \in \mathcal{S}} (\lambda^s)^\top (A^s x - b^s) z^s \\ \leq\ &\sum_{s \in \mathcal{S}} (\lambda^s)^\top (A^s x) z^s - \sum_{s \in \mathcal{S}} (\lambda^s)^\top b^s z^s, \end{aligned}$$

which implies

$$\sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top (A^s x - b^s) \leq \sum_{i=1}^n U_i(\lambda) x_i - L(\lambda)$$

with

$$\begin{aligned} U_i(\lambda) = \max_z \quad & \sum_{s \in \mathcal{S}} (\lambda^s)^\top A^s_{\cdot i} z^s \\ \text{s.t.} \quad & \sum_{s \in \mathcal{S}} z^s = 1 - \tau, && (\alpha) \\ & 0 \leq z^s \leq p^s, \quad s \in \mathcal{S}, && (\beta^s) \end{aligned}$$

and

$$\begin{aligned} L(\lambda) = \min_z \quad & \sum_{s \in \mathcal{S}} (\lambda^s)^\top b^s z^s && && \text{(6a)} \\ \text{s.t.} \quad & \sum_{s \in \mathcal{S}} z^s = 1 - \tau, && (\gamma) && \text{(6b)} \\ & 0 \leq z^s \leq p^s, \quad s \in \mathcal{S}. && (\delta^s) && \text{(6c)} \end{aligned}$$

□

In the following, we will call the inequalities (5) primal-dual valid inequalities (PD-VIs).

3.2. **Covering and Multi-Disjunctive Valid Inequalities.** In this section we first introduce what we call covering valid inequalities. Then, by scaling (if possible), the coefficients A^s by the corresponding b^s, we introduce the so-called multi-disjunctive valid inequalities. The latter are a generalization of the valid inequalities introduced in [34] under the name simple-disjunctive cuts.

For what follows, let us introduce $\Lambda^{m,\mathcal{S}} = \{\lambda = (\lambda^s)_{s \in \mathcal{S}} \colon \lambda^s \in \Lambda^m, s \in \mathcal{S}\}$.

Proposition 2. *The inequalities*

$$\sum_{i=1}^n U_i(\lambda, \bar{\mathcal{S}}) x_i \geq L(\lambda, \bar{\mathcal{S}}), \quad \bar{\mathcal{S}} \subseteq \mathcal{S},\ p(\bar{\mathcal{S}}) > \tau,\ \lambda \in \Lambda^{m,\mathcal{S}} \tag{7}$$

with

$$U_i(\lambda, \bar{\mathcal{S}}) = \max_w \left\{ \sum_{s \in \bar{\mathcal{S}}} (\lambda^s)^\top A^s_{\cdot i} w^s \colon \sum_{s \in \bar{\mathcal{S}}} w^s = p(\bar{\mathcal{S}}) - \tau,\ 0 \leq w^s \leq p^s,\ s \in \bar{\mathcal{S}} \right\}$$

and

$$L(\lambda,\bar{\mathcal{S}}) = \min_{w}\left\{\sum_{s\in\bar{\mathcal{S}}}(\lambda^s)^\top b^s w^s \colon \sum_{s\in\bar{\mathcal{S}}} w^s = p(\bar{\mathcal{S}})-\tau,\ 0\le w^s\le p^s,\ s\in\bar{\mathcal{S}}\right\}$$

are valid for Problem (3).

Proof. Since $p(\bar{\mathcal{S}}) > \tau$ holds, there exists a subset $\mathcal{S}^c \subset \bar{\mathcal{S}}$ such that $p(\mathcal{S}^c) \ge p(\bar{\mathcal{S}}) - \tau$ and $(\lambda^s)^\top(A^s x - b^s) \ge 0$ holds for all $s \in \mathcal{S}^c$. Taking a weighted sum of these inequalities with coefficients w^s for $s \in \mathcal{S}^c$ with $\sum_{s\in\mathcal{S}^c} w^s = p(\bar{\mathcal{S}}) - \tau$ and $0 \le w^s \le p^s$ yields

$$\sum_{s\in\mathcal{S}^c} w^s(\lambda^s)^\top(A^s x - b^s) \ge 0.$$

Inequality (7) is obtained by additionally using the definitions of $L(\lambda,\bar{\mathcal{S}})$ and $U_i(\lambda,\bar{\mathcal{S}})$. □

Moreover, if all right-hand side coefficients b_i^s have the same sign, scaling the covering VIs (7) leads to the following valid inequalities.

Corollary 1. *Suppose that all right-hand side coefficients b_i^s are non-zero and have the same sign. Given $\bar{\mathcal{S}} \subseteq \mathcal{S}, p(\bar{\mathcal{S}}) > \tau$, let us define*

$$\chi(\bar{\mathcal{S}}) = \begin{cases} p(\bar{\mathcal{S}})-\tau, & \text{if all } b_i^s, s\in\mathcal{S}, i\in[1:m] \text{ are strictly positive,}\\ \tau - p(\bar{\mathcal{S}}), & \text{if all } b_i^s, s\in\mathcal{S}, i\in[1:m] \text{ are strictly negative.}\end{cases}$$

Then, the inequalities

$$\sum_{i=1}^n V_i(\lambda,\bar{\mathcal{S}})x_i \ge \chi(\bar{\mathcal{S}}),\quad \bar{\mathcal{S}}\subseteq\mathcal{S},\ p(\bar{\mathcal{S}})>\tau,\ \lambda\in\Lambda^{m,\mathcal{S}} \tag{8}$$

with

$$V_i(\lambda,\bar{\mathcal{S}}) = \max_{w}\left\{\sum_{s\in\bar{\mathcal{S}}}(\lambda^s)^\top \frac{A^s_{\cdot i}}{|b_i^s|} w^s \colon \sum_{s\in\bar{\mathcal{S}}} w^s = p(\bar{\mathcal{S}})-\tau,\ 0\le w^s\le p^s,\ s\in\bar{\mathcal{S}}\right\}$$

are valid for Problem (3).

In what follows, we call the inequalities in (7) covering valid inequalities (C-VIs), while the inequalities in (8) are referred to as multi-disjunctive valid inequalities (MD-VIs). Note that these inequalities include, as a special case, the so-called simple-disjunctive cuts introduced in [34].

3.3. **Dominance Relations.** We now discuss dominance relations between the previously proposed valid inequalities.

Proposition 3. *There is no dominance relation between the C-VIs* (7) *and the MD-VIs* (8).

Proof. See the specific example given in Appendix A. □

Proposition 4. *For any given $\bar{\mathcal{S}} \subset \mathcal{S}$ and $\lambda \in \Lambda^{m,\mathcal{S}}$, the C-VIs* (7) *dominate the PD-VIs* (5).

Proof. Let us take $\bar{\mathcal{S}} \subset \mathcal{S}$ such that $p(\bar{\mathcal{S}}) \le 1-\tau$ and $\lambda \in \Lambda^{m,\mathcal{S}}$. The terms of the inequalities (5) can be re-arranged to obtain

$$\sum_{i=1}^n U_i(\lambda)x_i - \sum_{s\in\bar{\mathcal{S}}} p^s(\lambda^s)^\top A^s x \ge L(\lambda) - \sum_{s\in\bar{\mathcal{S}}} p^s(\lambda^s)^\top b^s.$$

To conclude the proof, one needs to prove that both

$$L(\lambda, \mathcal{S} \setminus \bar{\mathcal{S}}) \geq L(\lambda) - \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top b^s \tag{9}$$

and

$$\sum_{i=1}^{n} U_i(\lambda) x_i - \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top A^s x \geq \sum_{i=1}^{n} U_i(\lambda, \mathcal{S} \setminus \bar{\mathcal{S}}) x_i \tag{10}$$

hold as $p(\mathcal{S} \setminus \bar{\mathcal{S}}) > \tau$.

Let us first note that Inequality (9) can be written as

$$L(\lambda, \mathcal{S} \setminus \bar{\mathcal{S}}) + \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top b^s \geq L(\lambda). \tag{11}$$

Since the term $\sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top b^s$ is constant, the left-hand side of Inequality (11) is equivalent to

$$\begin{aligned} \min_{w} \quad & \sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} (\lambda^s)^\top b^s w^s + \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top b^s \\ \text{s.t.} \quad & \sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} w^s = p(\mathcal{S} \setminus \bar{\mathcal{S}}) - \tau, \\ & 0 \leq w^s \leq p^s, \quad s \in \mathcal{S} \setminus \bar{\mathcal{S}}, \end{aligned}$$

which is in turn equivalent to

$$\min_{w,z} \quad \sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} (\lambda^s)^\top b^s w^s + \sum_{s \in \bar{\mathcal{S}}} (\lambda^s)^\top b^s z^s \tag{12a}$$

$$\sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} w^s = p(\mathcal{S} \setminus \bar{\mathcal{S}}) - \tau, \tag{12b}$$

$$\sum_{s \in \bar{\mathcal{S}}} z^s = p(\bar{\mathcal{S}}), \tag{12c}$$

$$0 \leq w^s \leq p^s, \quad s \in \mathcal{S} \setminus \bar{\mathcal{S}}, \tag{12d}$$

$$0 \leq z^s \leq p^s, \quad s \in \bar{\mathcal{S}}. \tag{12e}$$

Summing up the equations in (12b) and (12c), one obtains Formulation (6), having the optimal value $L(\lambda)$. Since Formulation (6) is a relaxation of Formulation (12), it follows that Inequality (11) holds. To prove that Inequality (10) holds, we note that it can be rewritten as

$$\sum_{i=1}^{n} \left(U_i(\lambda) - \sum_{i=1}^{n} \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top A^s_{.i} \right) x_i \geq \sum_{i=1}^{n} U_i \left(\lambda, \mathcal{S} \setminus \bar{\mathcal{S}}\right) x_i,$$

and by arguments similar to the previous ones, one can prove that

$$U_i(\lambda) - \sum_{s \in \bar{\mathcal{S}}} p^s (\lambda^s)^\top A^s_{.i} \geq U_i(\lambda, \mathcal{S} \setminus \bar{\mathcal{S}}).$$

holds for each $i \in [1 : n]$. □

4. Separation Routines

We now discuss separation routines for the MD-VIs presented in Section 3.2. Note that we do not discuss separation routines for the PD-VIs presented in Section 3.1 because they are dominated by the C-VIs as stated in Proposition 4.

We start by showing in Section 4.1 that separating C-VIs and MD-VIs are NP-hard problems. Preliminary results showed that considering MD-VIs provides better results than considering C-VIs. As such, in this section we discuss routines to separate MD-VIs. We provide in Section 4.2 a mathematical formulation for the separation problem for MD-VIs. This serves as a preliminary step toward presenting the heuristic routines we developed to detect violated inequalities. The latter are presented in Section 4.3.

4.1. Complexity of the Separation Problem for C-VIs and MD-VIs.

Proposition 5. *For a fixed value of* $p(\bar{\mathcal{S}})$*, the separation problem for the C-VIs is NP-hard even in the special case of* $p^s = 1/|\mathcal{S}|$ *for all* $s \in \mathcal{S}$*,* $\tau = k/|\mathcal{S}|$ *for a* $k \in [1 : |\mathcal{S}| - 1]$*,* $(\lambda)^\top A^s_{.i} \in \{0, 1\}$ *for all* $s \in \mathcal{S}$*,* $i \in [1 : m]$*, and* $(\lambda)^\top b^s = \Delta$ *for all* $s \in \mathcal{S}$*.*

Proof. Under the previous conditions, the decision version (D-SEP) of the separation problem of the C-VIs for a point x^* consists in determining whether there exists a subset $\bar{\mathcal{S}} \subseteq \mathcal{S}$ such that $|\bar{\mathcal{S}}| = B$ and

$$\sum_{i=1}^{n} U_i(\lambda, \bar{\mathcal{S}}) x_i^* < L(\lambda, \bar{\mathcal{S}})$$

holds. This problem clearly belongs to NP. Furthermore, we show that CLIQUE reduces to it; see, e.g., problem GT19 in [16]. To this end, for an instance of CLIQUE given by a graph $G = (V, E)$ and an integer B, we define an instance of (D-SEP) as follows. We set $\mathcal{S} = V$, $E = [1 : m]$, $k = B - 1$, $x_i^* = \chi$ for all $i \in [1 : m]$,

$$\Delta = \chi \left(|E| - \left(\frac{(|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)}{2} - 1 \right) \right).$$

We also set $\lambda^\top A^s_{.i} = 1$ if the edge i is incident to node s and 0 otherwise. Then, we have

$$\begin{aligned} U_i(\lambda, \bar{\mathcal{S}}) &= \max_{s \in \bar{\mathcal{S}}} \{(\lambda)^\top A^s_{.i}\}, \\ L(\lambda, \bar{\mathcal{S}}) &= \Delta = \chi \left(|E| - \left(\frac{(|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)}{2} - 1 \right) \right). \end{aligned}$$

It holds $U_i(\lambda, \bar{\mathcal{S}}) = 0$ if and only if edge i has both end nodes in $\mathcal{S} \setminus \bar{\mathcal{S}}$. Moreover, $|\mathcal{S} \setminus \bar{\mathcal{S}}| = |\mathcal{S}| - B$. Then,

$$\sum_{i=1}^{n} U_i(\lambda, \bar{\mathcal{S}}) x_i^* = \chi(|E| - |E(\mathcal{S} \setminus \bar{\mathcal{S}})|) < \chi \left(|E| - \left(\frac{(|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)}{2} - 1 \right) \right),$$

which is satisfied if and only if

$$|E(\mathcal{S} \setminus \bar{\mathcal{S}})| > \frac{(|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)}{2} - 1,$$

which itself is satisfied if and only if $\mathcal{S} \setminus \bar{\mathcal{S}}$ is a clique of size $|\mathcal{S}| - B$ of $G = (V, E)$. □

Proposition 6. *For a fixed value of* $p(\bar{\mathcal{S}})$*, the separation problem for the MD-VIs is NP-hard even in the special case of* $p^s = 1/|\mathcal{S}|$ *for all* $s \in \mathcal{S}$*,* $\tau = k/|\mathcal{S}|$ *for a* $k \in [1 : |\mathcal{S}| - 1]$*,* $(\lambda)^\top \frac{A^s_{.i}}{|b^s_i|} \in \{0, 1\}$ *for all* $s \in \mathcal{S}$*,* $i \in [1 : m]$*.*

Proof. The proof is very similar to the last one. The decision version (D-SEP) of the separation problem of the MD-VIs for a point x^* consists in determining if there exists a subset $\bar{\mathcal{S}} \subseteq \mathcal{S}$ with $|\bar{\mathcal{S}}| = B$ and

$$\sum_{i=1}^{n} V_i(\lambda, \bar{\mathcal{S}}) x_i^* < \frac{1}{|\mathcal{S}|}.$$

This problem clearly belongs to NP. Furthermore, we again show that CLIQUE reduces to it. To this end, for an instance of CLIQUE given by a graph $G = (V, E)$

and an integer B, we define an instance of (D-SEP) as follows. We set $\mathcal{S} = V$, $E = [1:m]$, $k = B - 1$, $x_i^* = \frac{1}{\Delta}$ for all $i \in [1:m]$,

$$\Delta = |\mathcal{S}| \left(|E| - \left(\frac{(|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)}{2} - 1 \right) \right).$$

We also set $\lambda^\top \frac{A^s_{.i}}{b^s_i} = 1$ if edge i is incident to node s and 0 otherwise. Then, we again have

$$V_i(\lambda, \bar{\mathcal{S}}) = \max_{s \in \bar{\mathcal{S}}} \left\{ (\lambda)^\top \frac{A^s_{.i}}{b^s_i} \right\}.$$

It holds $V_i(\lambda, \bar{\mathcal{S}}) = 0$ if and only if edge i has both end nodes in $\mathcal{S} \setminus \bar{\mathcal{S}}$. Moreover, $|\mathcal{S} \setminus \bar{\mathcal{S}}| = |\mathcal{S}| - B$. Then,

$$\sum_{i=1}^n V_i(\lambda, \bar{\mathcal{S}}) x_i^* = \frac{1}{\Delta}(|E| - |E(\mathcal{S} \setminus \bar{\mathcal{S}})|) < \frac{1}{|\mathcal{S}|},$$

which itself is satisfied if and only if $|E(\mathcal{S} \setminus \bar{\mathcal{S}})| > (|\mathcal{S}| - B)(|\mathcal{S}| - B - 1)/2 - 1$, which itself is satisfied if and only if $\mathcal{S} \setminus \bar{\mathcal{S}}$ is a clique of size $|\mathcal{S}| - B$ of $G = (V, E)$. □

4.2. **Mathematical Formulation for the Separation Problem of Multi-Disjunctive Valid Inequalities.** Let us consider

$$\begin{aligned} V_i(\lambda, \bar{\mathcal{S}}) = \max_w \quad & \sum_{s \in \bar{\mathcal{S}}} (\lambda^s)^\top \frac{A^s_{.i}}{|b^s_i|} w^s \\ \text{s.t.} \quad & \sum_{s \in \bar{\mathcal{S}}} w^s = p(\bar{\mathcal{S}}) - \tau, && (\alpha) \\ & 0 \le w^s \le p^s, \quad s \in \bar{\mathcal{S}}. && (\beta^s) \end{aligned}$$

The dual problem is given by

$$\begin{aligned} \min_{\alpha, \beta} \quad & (p(\bar{\mathcal{S}}) - \tau)\alpha + \sum_{s \in \bar{\mathcal{S}}} p^s \beta^s \\ \text{s.t.} \quad & \alpha + \beta^s \ge (\lambda^s)^\top \frac{A^s_{.i}}{|b^s_i|}, \quad s \in \bar{\mathcal{S}}, \\ & \beta^s \ge 0, \quad s \in \bar{\mathcal{S}}. \end{aligned}$$

Hence, the separation problem for the MD-VIs at a given point $\bar{x}$ can be formulated as the following mixed integer nonlinear problem:

$$\min \quad \sum_{i=1}^n \left(\left(\sum_{s \in \mathcal{S}} p^s v^s - \tau \right) \alpha_i + \sum_{s \in \mathcal{S}} \beta^s_i p^s v^s \right) \bar{x}_i + \varphi \left(\sum_{s \in \mathcal{S}} p^s v^s - \tau \right) \tag{13a}$$

$$\text{s.t.} \quad \alpha_i + \beta^s_i \ge (\lambda^s)^\top \frac{A^s_{.i}}{|b^s_i|}, \quad i \in [1:n], s \in \mathcal{S}, \tag{13b}$$

$$\sum_{i=1}^m \lambda^s_i = 1, \quad s \in \mathcal{S}, \tag{13c}$$

$$\sum_{s \in \mathcal{S}} p^s v^s \ge \tau, \tag{13d}$$

$$v^s \in \{0, 1\}, \quad s \in \mathcal{S}, \tag{13e}$$

$$\beta^s_i \ge 0, \quad i \in [1:n], s \in \mathcal{S}, \tag{13f}$$

$$\lambda^s_i \ge 0, \quad i \in [1:n], s \in \mathcal{S}, \tag{13g}$$

with

$$\varphi = \begin{cases} -1, & \text{if all } b^s_i, s \in \mathcal{S}, i \in [1:m], \text{ are strictly positive,} \\ 1, & \text{if all } b^s_i, s \in \mathcal{S}, i \in [1:m], \text{ are strictly negative.} \end{cases}$$

Given the complexity result on the separation of MD-VIs, one may expect that the resolution of the previous problem with a commercial solver requires a considerable amount of computation time if the size of the instance grows. Hence, in the next section, we propose two heuristic routines to separate MD-VIs.

4.3. **Heuristic Separation Procedures for MD-VIs.**

4.3.1. *An ADM-Like Approach.* Alternating direction methods (ADMs) have been initially proposed in [19] as extensions of Lagrangian methods. They are iterative procedures typically used to tackle problems defined by means of two vectors of decision variables, which are subject to some coupling constraints. Instead of solving the monolithic original problem, in every iteration of an ADM, one sequentially solves two smaller subproblems, each of which determines a new value for one of the variable vectors, having fixed the value of the other one. In recent years, ADMs have been exploited to solve large-scale optimization problems in the field of gas transport [17, 18], machine learning [8, 27], bilevel problems [22], supply chain problems [37], or maintenance planning and portfolio optimization problems [9].

In what follows, let $(\bar{x}, \bar{y})$ denote the point to be separated. The ADM-like separation routine leverages the decomposable structure of Problem (13) w.r.t. the variables λ and v. Iteratively, one set of variables is fixed while the other is optimized using Model (13).

If λ is fixed, (13) reduces to a combinatorial problem that, according to Proposition 6, is NP-hard. We thus heuristically set $v^s = 1$ for all $s \in \mathcal{S}$ satisfying $\lambda^s(A^s\bar{x} - b^s) < 0$. This choice is motivated by the fact that the most violated PD-VI (5) is obtained in this way. Although, they are weaker than the MD-VIs, choosing the scenarios that provide the most violated PD-VI is a good proxy.

If v is fixed, (13) simplifies to a linear problem with continuous variables. The optimal solution can be computed efficiently.

To initialize the ADM, we set $v^s = 1$ if $\min_{j\in[1:m]}(A^s_j\bar{x} - b^s_j) < 0$ and $v^s = 0$ otherwise. This ensures that all scenarios with at least one violated constraint for the current solution $\bar{x}$ are selected. By construction, this leads to $\sum_{s\in\mathcal{S}} p^s v^s \geq \tau$, as $\bar{x}$ would otherwise be a feasible solution to the original problem.

4.3.2. *A Constraint-Wise Approach.* The constraint-wise approach is a straightforward heuristic that is not directly linked to Problem (13). As indicated by its name, the constraint-wise approach computes the most violated inequality (MD-VI) for each constraint individually. Let $(\bar{x}, \bar{y})$ denote the fractional solution to be separated. For each constraint $j \in [1 : m]$, we proceed as follows.

(i) For each scenario $s \in \mathcal{S}$, we include the scenario in the set $\bar{\mathcal{S}}$ and set $\lambda^s_j = 1$ if $A^s_j\bar{x} - b^s_j < 0$. Otherwise, we exclude the scenario from $\bar{\mathcal{S}}$, i.e., $s \notin \bar{\mathcal{S}}$. If $s \notin \bar{\mathcal{S}}$, the value of λ^s_j can be chosen arbitrarily as it does not appear in the inequality.
(ii) Compute the parameters $U_i(\lambda, \bar{\mathcal{S}})$ and $L(\lambda, \bar{\mathcal{S}})$ as described in Proposition 2.

Once all inequalities have been computed, we identify and select the inequality with the largest violation and add it to the formulation.

5. Closures for CCCOP

In this section, we restrict ourselves to the case in which all the linear systems $A^s x \geq b^s$ involved in Problem (3) have all the coefficients in A^s and b^s being nonnegative and $X = \mathbb{R}^n_{\geq 0}$. This means, we consider covering-type constraints. We call this special case problem a *chance-constrained covering type linear optimization problem* (CCCOP). We assume w.l.o.g. that $b^s = \mathbb{1}$ holds for all $s \in \mathcal{S}$. Indeed, if there would exist $i \in [1 : m]$ and $s \in \mathcal{S}$ with $b^s_i = 0$, the corresponding constraint can be eliminated as it is trivially satisfied. (Note that we also assume $x \in X = \mathbb{R}^n_{\geq 0}$.) Otherwise, if $b_i^s > 0$ holds, re-scaling the coefficients of the corresponding constraint is possible. Thus, the CCCOP is formulated as

$$\begin{aligned} \min_{x,y} \quad & c^\top x \\ \text{s.t.} \quad & x \in \mathbb{R}^n_{\geq 0}, \\ & A^s x \geq y^s \mathbb{1}, \quad s \in \mathcal{S}, \\ & \sum_{s \in \mathcal{S}} p^s y^s \geq 1 - \tau, \\ & y^s \in \{0,1\}, \quad s \in \mathcal{S}. \end{aligned}$$

In [34], the so-called *covering-type k-violation linear problem* (CKVLP) is considered. The CKVLP is a special case of the CCCOP in which all scenarios are equiprobable, i.e., $p^s = 1/|\mathcal{S}|, s \in \mathcal{S}$ and $m = 1$, leading to $A^s \in \mathbb{R}^{1 \times n}$ and $b^s = 1$ for all $s \in \mathcal{S}$. The authors of [34] propose a family of inequalities valid for the CKVLP called simple-disjunctive cuts. The simple-disjunctive cuts are special cases of our MD-VIs (8) where $p^s = 1/|\mathcal{S}|$, $m = 1$, and $\bar{\mathcal{S}} \subseteq \mathcal{S}$ is chosen such that $p(\bar{\mathcal{S}}) = \tau + 1/|\mathcal{S}|$.

Recall that $\Lambda^m = \{\lambda \in \mathbb{R}^m_{\geq 0} \colon \mathbb{1}^\top \lambda = 1\}$ denotes the m-dimensional unit simplex and that we set $\Lambda^{m,\mathcal{S}} = \{\lambda = (\lambda^s)_{s \in \mathcal{S}} \colon \lambda^s \in \Lambda^m, s \in \mathcal{S}\}$. We now define the closure multi-$\mathcal{D}$ of the MD-VIs (8) as

$$\text{multi-}\mathcal{D} = \bigcap_{\bar{\mathcal{S}} \subset \mathcal{S}, p(\bar{\mathcal{S}}) > \tau} \; \bigcap_{\{\lambda \in \Lambda^{m,\mathcal{S}}\}} \left\{ x \in \mathbb{R}^n_{\geq 0} \colon \sum_{i=1}^n V_i(\lambda, \bar{\mathcal{S}}) x_i \geq p(\bar{\mathcal{S}}) - \tau \right\}$$

with

$$V_i(\lambda, \bar{\mathcal{S}}) = \max_w \left\{ \sum_{s \in \bar{\mathcal{S}}} (\lambda^s)^\top A^s_{\cdot i} w^s \colon \sum_{s \in \bar{\mathcal{S}}} w^s = p(\bar{\mathcal{S}}) - \tau, \; 0 \leq w^s \leq p^s, \; s \in \bar{\mathcal{S}} \right\}.$$

Let us also define the set of critical scenario subsets as

$$\mathcal{C} = \left\{ \bar{\mathcal{S}} \subseteq \mathcal{S} \colon p(\bar{\mathcal{S}}) > \tau, \, p(\bar{\mathcal{S}} \setminus \{s\}) \leq \tau, \, s \in \bar{\mathcal{S}} \right\}.$$

Let simple-$\mathcal{D}$ denote the closure of all MD-VIs (8) corresponding to a (simple) disjunction in which at least one constraint set must be satisfied:

$$\text{simple-}\mathcal{D} = \bigcap_{\bar{\mathcal{S}} \in \mathcal{C}} \; \bigcap_{\{\lambda \in \Lambda^{m,\mathcal{S}}\}} \left\{ x \in \mathbb{R}^n_{\geq 0} \colon \sum_{i=1}^n V_i(\lambda, \bar{\mathcal{S}}) x_i \geq 1 \right\}$$

with

$$V_i(\lambda, \bar{\mathcal{S}}) = \max_{s \in \bar{\mathcal{S}}} \left\{ (\lambda^s)^\top A^s_{\cdot i} \colon s \in \bar{\mathcal{S}} \right\}.$$

Finally, following [28, 34, 42], we define a quantile cut as follows. For any $\alpha \in \mathbb{R}^n_{\geq 0}$, let $\beta^s_\alpha = \min_x \{\alpha^\top x \colon A^s x \geq \mathbb{1}, x \geq 0\}$ and β_α is the $(1 - \tau)$-quantile $\mathbb{Q}_{1-\tau}[\beta^s_\alpha]$ of the set $\{\beta^s_\alpha, s \in \mathcal{S}\}$, i.e.,

$$\beta_\alpha = \min \left\{ q \in \mathbb{R} \colon \sum_{s \in \mathcal{S} : q \geq \beta^s_\alpha} p^s \geq 1 - \tau \right\}.$$

Then, $\alpha^\top x \geq \beta_\alpha$ is valid for CCCOP and is called a quantile cut. The closure of all quantile cuts, denoted by $\mathcal{M}$, is defined as

$$\mathcal{M} = \bigcap_{\alpha \in \mathbb{R}^n_{\geq 0}} \left\{ x \in \mathbb{R}^n_{\geq 0} \colon \alpha^\top x \geq \beta_\alpha \right\}.$$

The identity simple-$\mathcal{D} = \mathcal{M}$ has been shown in [34] for the CKVLP. This result extends to the CCCOP and the closure multi-$\mathcal{D}$ of our MD-VIs (8) is contained in the two previous ones. In order to prove this, we first need the following lemma.

Lemma 2. *Let $\mathcal{P}$ denote the problem $\min_x\{\alpha^\top x\colon Ax \geq \mathbb{1}, x \geq 0\}$. That means, it is covering-type linear problem. Moreover, let x^* be an optimal solution to $\mathcal{P}$. Then, x^* is also optimal for $\min_x\{\alpha^\top x\colon (\lambda^*)^\top Ax \geq \mathbb{1}\}$, where λ^* is a normalized optimal solution to the dual of $\mathcal{P}$. This means $\lambda^* = \varepsilon^*/(\mathbb{1}^\top\varepsilon^*)$ holds for $\varepsilon^* \in \arg\max_\varepsilon\{\mathbb{1}^\top\varepsilon\colon A^\top\varepsilon \leq \alpha, \varepsilon \geq 0\}$.*

Proof. Let x^* denote an optimal solution to $\min_x\{\alpha^\top x\colon Ax \geq \mathbb{1}, x \geq 0\}$. Then, it is also feasible for the surrogate problem $\min_x\{\alpha^\top x\colon (\lambda^*)^\top Ax \geq \mathbb{1}, x \geq 0\}$. Furthermore, for any feasible point x to this surrogate problem, we have

$$\alpha^\top x \geq (\varepsilon^*)^\top Ax \geq \mathbb{1}^\top\varepsilon^* = \alpha^\top x^*,$$

where the first inequality comes from the feasibility of the dual solution, the second inequality comes from the feasibility of x, and the last equality follows from strong duality. □

Proposition 7. *For the CCCOP, multi-$\mathcal{D} \subseteq$ simple-$\mathcal{D} = \mathcal{M}$ holds. In addition, there exist instances for which the inclusion is strict.*

Proof. The first inclusion comes from that fact that the MD-VIs defining simple-$\mathcal{D}$ are contained in the set of those defining multi-$\mathcal{D}$.

To prove that simple-$\mathcal{D} \subseteq \mathcal{M}$ holds, we consider a valid quantile cut $\alpha^\top x \geq \beta_\alpha$ and assume that the scenarios $s \in \mathcal{S}$ are ranked so that $\beta_\alpha^s \leq \beta_\alpha^{s+1}$ holds for $s = 1, \ldots, |\mathcal{S}| - 1$ and that $\beta_\alpha = \beta_\alpha^k$ holds, where $k = \lceil|\mathcal{S}|(1-\tau)\rceil$. From Lemma 2, it follows that there exists $\lambda^s \in \Lambda^m$, $s \in \mathcal{S}$, such that $\alpha^\top x \geq \beta_\alpha$ is valid for the sets

$$\{x \in \mathbb{R}^n_{\geq 0}\colon (\lambda^s)^\top A^s x \geq 1\}$$

for $k \leq s \leq |\mathcal{S}|$. Hence, it is also valid for

$$\bigcup_{s=k}^{|\mathcal{S}|} \left\{x \in \mathbb{R}^n_{\geq 0}\colon (\lambda^s)^\top A^s x \geq 1\right\}.$$

The convex hull of this set is given by

$$\left\{x \in \mathbb{R}^n_{\geq 0}\colon \sum_{i=1}^n V_i(\lambda, \bar{\mathcal{S}})x_i \geq 1\right\}$$

with $\bar{\mathcal{S}} = [k : |\mathcal{S}|] \in \mathcal{C}$. Hence, every quantile cut is dominated by some simple-disjunctive cut.

To prove the converse, let $\sum_{i=1}^n V_i(\lambda, \bar{\mathcal{S}})x_i \geq 1$ be a simple-disjunctive cut. Then,

$$V_i(\lambda, \bar{\mathcal{S}}) = \max_{s\in\bar{\mathcal{S}}}\left\{(\lambda^s)^\top A^s_{.i}\colon s \in \bar{\mathcal{S}}\right\}$$

and for each $s \in \bar{\mathcal{S}}$, we have

$$\begin{aligned}\beta_V^s &= \min\left\{V^\top x\colon A^s x \geq 1\right\} \geq \min\left\{V^\top x\colon (\lambda^s)^\top A^s x \geq 1\right\}\\ &= \min\left\{V_i(\lambda, \bar{\mathcal{S}})/(\lambda^s)^\top A^s_{.i})\colon i \in [1:n]\right\} \geq 1.\end{aligned}$$

Given that $p(\bar{\mathcal{S}}) > \tau$, the quantile cut associated to V has a right-hand side not smaller than 1. Hence, any simple-disjunctive cut is dominated by a quantile cut.

Finally, consider the following instance of CKLVP with four variables $x_i \geq 0$, $i = 1, \ldots, 4$, four scenarios $\mathcal{S} = \{1, 2, 3, 4\}$ with probability equal to $1/4$ each, and the unique constraint $x_i \geq 1$ of scenario i. The list of inequalities (8) consists of

- $x_i + x_j \geq 1$ for $i, j \in \{1, 2, 3, 4\}$ with $i \neq j$,
- $x_i + x_j + x_k \geq 2$ for $i, j, k \in \{1, 2, 3, 4\}$ with $i \neq j \neq k$,

- $x_1 + x_2 + x_3 + x_4 \geq 3$.

The simple disjunctive cuts are the inequalities involving two variables. In addition, all the above constraints are facet defining and together with the nonnegativity constraints they define the convex hull F of the feasible solutions of this particular CKVLP that is given by

$$F = \text{conv}(\{(1,1,1,0),(1,1,0,1),(1,0,1,1),(0,1,1,1),(1,1,1,1)\}) + \mathbb{R}^4_{\geq 0}. \qquad \square$$

6. Numerical Results

In this section, we assess the impact of the MD-VIs (8) with two types of experiments. The first type aims at evaluating the potential of the MD-VIs in improving the dual bound at the root node. To do so, we implemented a pure cutting-plane method in which we solve the linear relaxation of Formulation (3) and then separate MD-VIs. The procedure stops when no violated valid inequality is found or the time limit is reached. We refer to these experiments as the CP-tests. The second type of experiments aims at evaluating the performance of a MIP solver when Formulation (3) is solved and the MD-VIs are added in the root node of the branch-and-bound tree. To do so, first, we solve the linear relaxation of the formulation and separate the inequalities as in the CP-tests, i.e., via a cutting-plane method. Then, we impose the integrality constraints for the variables and we let the MIP solver explore the branch-and-bound tree. We chose this strategy because we realized that the MIP solver was not including all the detected violated inequalities in the formulation when inserting them via a callback. We refer to these experiments as MIP-tests.

In both CP- and MIP-tests, we compare the MD-VIs against two well-known families of valid inequalities for chance-constrained problems from the literature, namely the mixing-set inequalities considered in [34] and the quantile cuts in [1]. We build instances of two classes of problems which have the structure of problem (2), namely the *covering-type k-violation linear program* (CKVLP) [34] and the *chance-constrained multidimensional knapsack problem* (CCMKP) [40], to run the CP- and MIP-tests.

Finally, in our experimental analysis, we do not assess the PD-VIs (5) as they are dominated by the MD-VIs (see Proposition 4), nor the C-VIs (7) as the MD-VIs showed better results in preliminary numerical tests.

6.1. Instances.

We build two sets of instances for the CKVLP and the CCMKP problems. Specifically, one set is generated using a reference method from the literature, referred to as lit, and the other set is generated using a new method we devised, referred to as new. The new instance generation method is designed to allow for a comparison of the robustness of different inequality families and cut generation methods across various instance types. Here, by robustness, we refer to the ability of the inequalities to maintain strong performance under different structural variations in the problem data.

6.1.1. *Instances for the CKVLP.*

We generate instances for the CKVLP problem with $n \in \{20, 80\}$ decision variables and $|\mathcal{S}| \in \{100, 500, 1000, 3000\}$ equiprobable scenarios, i.e., we have $p^s = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$. For each combination of n and $|\mathcal{S}|$, we create five instances using both the lit and the new method.

The lit method follows the approach proposed in [34]. Specifically, for each scenario $s \in \mathcal{S}$, the entries of the $1 \times n$ matrix A^s are drawn independently from a uniform distribution over the interval $[0.8, 1.5]$. In the new method, we first generate a base row A, where each entry is sampled as in the lit method. Then, for each scenario $s \in \mathcal{S}$, the entries A_i^s of the row A^s are sampled from a normal distribution with mean A_i and standard deviation $0.1|A_i|$. For all instances, the right-hand side

b^s is set to 1.1, and the coefficients of the objective function c are drawn as integer values from a uniform discrete distribution over the interval $[1, 100]$.

6.1.2. *Instances for the CCMKP.* We generate instances for the CCMKP problem with two configurations: (i) $n = 20$ items and $m = 10$ dimensions, and (ii) $n = 40$ items and $m = 30$ dimensions.

In the lit method, we follow the standard approach as described in [1, 40]. A single base system of inequalities $Ax \geq b$ is generated, and the entries A^s_{ij} of each scenario matrix A^s are sampled from a normal distribution with mean A_{ij} and standard deviation $0.1|A_{ij}|$. The right-hand side b^s is kept unchanged and is thus equal to b. In the new method, we generate four independent base inequality systems. Then each base matrix A is used to sample 25 % of the scenarios using the same procedure as in the lit method. This design aims to introduce greater heterogeneity across the different scenario matrices, allowing us to analyze how different families of valid inequalities influence the linear relaxation when entirely distinct systems of inequalities are present.

Finally we remark that, in all the instances of the CKVLP and of the CCMKP, we set $\tau = 0.2$.

6.2. **Experimental Setup.** In this section, we first describe the separation strategy that we adopt for the MD-VIs and for the families of inequalities used for the sake of comparison, i.e., the mixing-set inequalities and the quantile cuts. Then, we present the implementation details for the algorithms.

The MD-VIs are separated with the heuristic procedures presented in Section 4, namely the ADM-like approach and the constraint-wise approach. We denote these two procedures by ADM and C-wise, respectively. Precisely, in the experiments involving instances of the CKVLP, we separate the MD-VIs by means of the C-wise procedure, only. Indeed, as these instances only have a single constraint per scenario, the two procedures lead to the same violated inequality (if it exists). In the experiments involving instances of the CCMKP, we evaluate both the ADM and the C-wise procedure. When the ADM procedure is considered, it is run for a single iteration only. Indeed, optimizing in the λ-direction is computationally heavy due to the size of the linear optimization model to be solved; see Problem (13). In addition, preliminary experiments showed that further iterations of the procedure do not improve significantly the violation of the inequalities.

Given $\eta \in \mathbb{R}^n$ and a scenario $s \in \mathcal{S}$, a mixing-set inequality as given in [34] is given by

$$\eta^\top x + (\beta^q_\eta - \beta^s_\eta) y^s \geq \beta^q_\eta$$

where the coefficients β^s_η are defined in different manners for the CKVLP and the CCMKP and β^q_η is the τ-quantile of the values $\{\beta^s_\eta\}_{s \in \mathcal{S}}$. It should be noted that a generic separation approach, applicable to any form of linear CCP with finite support, currently does not exist in the literature. This is likely due to the computational challenges posed by separating these inequalities. Specifically, two key challenges arise: (i) selecting an appropriate direction η, and (ii) computing the β^s_η coefficients. Even if η is already selected, computing the β coefficients generally requires solving $|\mathcal{S}|$ linear problems, which can be computationally prohibitive in practice. In [34], the authors focus on the CKVLP case. They select $\eta = A^s_{i\cdot}$, where $A_{i\cdot}\bar{x} \leq b_i$ is the most violated scenario inequality with respect to the current solution $\bar{x}$. For CKVLP, the β coefficients can be efficiently computed using a greedy algorithm. Following their approach, we adopt the same method for separating mixing-set inequalities in the CKVLP case. For the more "general" CCMKP, however, the computation of β values is significantly more challenging. Solving $|\mathcal{S}|$ linear programs is too costly for large-scale problems. Since no prior work directly addresses this issue, we propose a heuristic inspired by the approach in [34]. While the direction η is selected in the same way, we compute the β values using the quantile estimation method described in [40]. This method provides efficient lower bounds for β values and has been shown to perform well on CCMKP instances. Moreover, this approach is regarded as one of the best performing ones in the literature. Finally, we remark that addressing the lack of efficient heuristics for constructing reasonable mixing sets for general CCPs might lead to promising directions for future research. Progress in this area could have substantial value for practical applications, particularly when large-scale scenario sets are present.

Given $\eta \in \mathbb{R}^n$, a quantile cut is given by

$$\eta^\top x \geq \beta^q_\eta$$

where β^q_η is defined as for the mixing-set inequalities. Although [1] proposed a heuristic procedure to separate the quantile cuts, early experiments revealed that it is rather ineffective for the instances of the CKVLP and CCMKP that we consider here. To address this limitation, we adopt a separation procedure similar to the one used for mixing-set inequalities. In particular, we define the coefficients η and compute the right-hand side β^q_η using the same methodology as for the mixing-set inequalities. We refer the reader to [34] and [1] for further details on the mixing-set inequalities and quantile cuts, respectively.

Further observe that the separation procedures for the MD-VIs, the mixing-set inequalities, and the quantile cuts may provide several violated inequalities, except when the MD-VIs are separated via the ADM procedure where at most one violated inequality is provided. Early experiments showed that the best separation strategy for the MD-VIs and the quantile cuts is to only include the most violated inequality. In contrast to that, the same experiments suggested that the best strategy for the mixing-set inequalities is to include at most 100 violated inequalities among those with the highest violation. Values $\bar{b}^s \in \mathbb{R}^m$ in Formulation (3) are set naively. For the CKVLP, they are set to $\bar{b}^s := b^s$ as suggested in [34]. On the other hand, for the CCMKP instances, they are set as suggested in [40], i.e., $\bar{b}^s := b^s - A^s \mathbb{1}$, where $\mathbb{1}$ is the vector of all ones in appropriate dimension. By setting these naive values we aim to make the experiments as simple as possible and to allow for a clear comparison between the different approaches.

The algorithms are implemented using Python. In addition, Cython is employed to speed up the solution of the greedy algorithms to compute the coefficients of the mixing-set inequalities and the quantile cuts. The linear and integer problems are solved using Gurobi 10.0.3 with a single thread. A one-hour time limit is imposed to the solver in the MIP-tests and the CP-tests. We run the experiments on a machine equipped with an Intel Gold 618 Skylake processor with 2.4 GHz and a 16 GB RAM.

6.3. **CP-Tests for the CKVLP and CCMKP Instances.** In this section, we report the results of the CP-tests with MD-VIs, the mixing-set inequalities, and the quantile cuts on both the CKVLP and CCMKP instances.

Table 1 reports the results for the instances of the CKVLP. Each row of the table represents a subset of five instances of the CKVLP. The characteristics of these subsets are reported in the first three columns of the table whose headings are the following.

n-m: Number n of x-variables and number m of constraints in each scenario.
Gen.: Instance generation method, i.e., either lit or new.
$|\mathcal{S}|$: Number of scenarios.

The next nine columns report the results of the CP-tests with the MD-VIs, mixing-set inequalities and quantile cuts. The headings used in those columns are as follows.

T: Average computational time of the solution of one separation problem.

TABLE 1. Results of the CP-tests with MD-VIs, mixing-set inequalities and quantile cuts on the instances of the CKVLP.

Instances			MD-VIs			Mixing-set			Quantile		
n-m	Gen.	$\lvert\mathcal{S}\rvert$	T	it.	v (%)	T	it.	v(%)	T	it.	v (%)
20-1	lit	100	0.00	3.80	**58.90**	0.00	2.40	17.32	0.00	2.40	15.82
		500	0.01	3.40	**50.97**	0.01	2.00	23.10	0.01	2.00	22.30
		1000	0.02	3.40	**34.43**	0.01	3.60	19.87	0.01	2.40	19.03
		3000	0.08	3.20	**36.53**	0.05	6.40	20.92	0.04	3.40	17.88
	new	100	0.00	3.00	**32.61**	0.00	2.40	30.57	0.00	2.20	30.11
		500	0.01	3.00	37.24	0.01	2.20	**41.82**	0.01	2.00	41.72
		1000	0.02	3.00	34.01	0.01	3.20	**48.76**	0.01	2.00	48.76
		3000	0.09	3.00	31.72	0.05	6.20	**37.67**	0.04	2.20	37.46
80-1	lit	100	0.00	4.40	**35.99**	0.00	3.00	6.14	0.00	2.40	3.51
		500	0.01	4.00	**33.03**	0.01	3.60	7.81	0.01	2.60	5.88
		1000	0.03	4.40	**22.91**	0.02	3.80	6.38	0.02	3.20	3.77
		3000	0.10	3.20	**42.22**	0.07	4.00	8.36	0.07	2.20	7.21
	new	100	0.00	3.20	**27.04**	0.00	3.40	17.37	0.00	3.20	15.39
		500	0.01	3.00	**31.13**	0.01	3.00	26.55	0.01	2.60	24.79
		1000	0.03	3.00	**30.64**	0.02	4.40	28.29	0.02	3.40	27.46
		3000	0.12	3.00	**33.62**	0.08	6.00	30.99	0.07	2.60	30.80

it.: Average number of iterations of the CP-tests;

v **(%):** Average improvement of the optimal value of the linear relaxation of Model (3) expressed as a percentage, i.e., $100(v_{NI} - v_I)/v_{NI}$, where v_{NI} is the optimal value of the linear relaxation of Formulation (3) with no inequalities and v_I is the optimal value of the linear relaxation of Model (3) after no further cutting planes can be added.

Regardless of the considered family of inequalities, we observe that the algorithm runs only for a few iterations before the separation problem does not detect any further violated inequality. In addition, the time to solve one separation problem associated with any of the three families of inequalities is negligible. Indeed, it is always less than 0.12 seconds on average. This is due to the fact that these instances are characterized by one constraint per scenario. The MD-VIs yield the best average improvements of the optimal value of the linear relaxation of Model (3) for 13 out of the 16 subsets of instances. The improvements are remarkable and span from 22.91 % to 58.90 %. In the remaining three subsets, the best improvements are attained with the mixing-set inequalities. The instances of these three subsets are characterized by n-m =20-1 and are generated with method new. However, we report that the improvements of the MD-VIs are at least equal to 31.72 % on average for these three instance sets.

Table 2 reports the results of the CP-tests with MD-VIs, mixing-set inequalities, and quantile cuts on the instances of the CCMKP. The rows of the table group five instances of the CCMKP. In addition to the column headings in Table 1, we include a new heading "$s(\lambda)$" for the results, where the MD-VIs are separated using the ADM approach. Here, $s(\lambda)$ is the average cosine similarity computed between each pair of multipliers λ^{s_1} and λ^{s_2} for $s_1, s_2 \in \mathcal{S}$. Cosine similarity is the cosine of the angle between the vectors, calculated as the scalar product of the vectors divided by the product of their norms. Since cosine similarity depends only on the angle and not on the magnitudes of the vectors, values range from -1 (exactly opposite vectors) to 1 (identical vectors). Comparing these cosine similarity values allows us to analyze how different instance generation methods influence the underlying MD-VIs obtained through the ADM approach.

Table 2. Results of the CP with MD-VIs, Mixing-set inequalities and quantile cuts on the instances of the CCMKP.

Instances			MD-VIs										Quantile		
			ADM				C-wise			Mixing-set					
n-m	Gen.	$\lvert\mathcal{S}\rvert$	$s(\lambda)$	T	it.	v (%)	T	it.	v (%)	T	it.	v(%)	T	it.	v (%)
20-10	lit	100	0.87	0.18	18.80	**30.21**	0.01	4.00	28.05	0.11	4.80	25.07	0.11	4.20	21.18
		500	0.73	0.42	13.80	**25.50**	0.04	4.00	23.47	0.52	6.20	22.10	0.51	3.80	19.84
		1000	0.89	0.81	15.00	**27.00**	0.07	3.80	25.14	1.09	9.00	24.47	1.04	5.00	21.89
		3000	0.88	3.27	11.80	**25.80**	0.26	3.40	24.77	3.88	17.80	22.99	3.23	4.80	20.82
	new	100	0.33	0.23	21.60	**24.03**	0.01	3.60	4.30	0.12	5.20	16.30	0.12	2.40	0.92
		500	0.33	0.43	30.80	**27.27**	0.04	4.20	10.97	0.60	8.00	18.60	0.58	4.00	1.55
		1000	0.29	0.79	15.80	**25.10**	0.09	3.00	11.31	1.17	8.40	15.92	1.09	4.20	2.65
		3000	0.31	3.04	22.20	**25.20**	0.37	3.80	5.21	3.72	18.20	16.67	3.40	3.80	2.11
40-30	lit	100	0.77	0.62	39.80	**27.65**	0.05	4.20	25.05	1.36	6.40	22.78	1.37	5.60	15.81
		500	0.79	1.27	21.00	**26.13**	0.19	3.80	23.66	6.72	7.40	19.72	6.67	4.60	16.04
		1000	0.73	3.17	47.40	**28.19**	0.50	4.20	25.46	13.55	12.40	22.74	11.83	6.20	18.94
		3000	0.79	8.96	20.80	**25.59**	1.66	3.60	23.25	41.06	17.20	19.05	38.16	5.20	15.85
	new	100	0.23	0.76	63.80	**26.08**	0.07	2.60	4.18	1.25	6.80	17.78	1.37	3.00	0.99
		500	0.26	1.15	31.80	**23.30**	0.28	3.20	6.42	6.89	6.80	14.27	6.74	4.00	1.67
		1000	0.22	2.64	39.80	**24.47**	0.67	3.00	4.40	13.18	10.80	15.73	12.66	4.20	2.14
		3000	0.25	8.75	32.20	**25.08**	2.80	3.80	7.13	40.25	19.80	14.93	36.87	4.00	1.77

Overall, the CP-tests where the MD-VIs are separated via the ADM approach lead to the best results. Although this method requires several iterations, the computational time to solve a single separation problem is acceptable, averaging less than 9 seconds across all subsets of instances. This approach yields the best average improvements in the optimal value of the linear relaxation of Model (3) on all subsets of instances. We note that the performance of the ADM approach remains consistent across different instance families. This can be attributed to the fact that, for each scenario, a set of multipliers λ^s is computed. Consequently, more effective inequalities of the form $\lambda^s A^s x \geq \lambda^s b^s$ are constructed for each scenario $s \in \mathcal{S}$. This observation is supported by the values of the average cosine similarity $s(\lambda)$. For the lit instances, the similarity between multipliers is close to one, indicating that the multipliers used to construct MD-VIs are highly similar. In contrast, for the new instances, the similarity drops to approximately 0.3, suggesting that separation of λ produces different multipliers for each scenario. The other valid inequality generation methods are unable to take advantage of this aspect effectively, which explains their weaker performance. These results highlight the strength of the MD-VIs. By adapting to the feasible regions $A^s x \leq b^s$, the ADM approach demonstrates its robustness in generating effective inequalities across diverse problem instances.

The CP-tests in which the MD-VIs are separated via the C-wise approach yields the shortest computational times. Indeed, it requires less than 5 iterations on average before stopping and the average computational time to solve a separation problem is less than 0.5 seconds for 13 out of the 16 subsets of instances and it is at most 2.8 seconds for the remaining three subsets. Despite being the fastest approach, the average improvements of the optimal value of the linear relaxation are competitive on the instances generated with the method from the literature (lit). Indeed, for such instances, the MD-VIs are separated via the C-wise approach yield the second-best of such improvements. However, we observe that the quality of the produced inequalities drops if the new instances are considered. This is because, in this case, the λ multipliers are the same for all scenarios. The CP-tests with the mixing-set inequalities stop after several iterations and the average computational time to solve a single separation problem is in general higher than the one of the MD-VIs. In particular, for the instances with n-m being 40-30, the computation times reach 40 seconds if the number of scenarios is 3000. The explanation is that computing the coefficients of such inequalities entails solving as many knapsack problems as the number of scenarios times the number of constraints in the scenarios, i.e., $|\mathcal{S}|\, m$. We remark that, as highlighted earlier, these parts of the code have been implemented using the Cython package, ensuring an efficient implementation of these separation heuristics. Despite being time consuming, we report that separating the mixing-set inequalities yields good improvements in terms of the optimal value of the linear relaxation over all the subset of instances—in contrast to the CP-tests where the MD-VIs are separated with method C-wise. The CP-tests with the quantile cuts stop after few iterations (less than 7 on average). However, the average computational time to separate them follows the same trend as the one to separate the mixing-set inequalities. Indeed, the same comments about the computation of the coefficients of the mixing-set inequalities applies for the right hand-side of quantile cuts. The average improvements in terms of the optimal value of the linear relaxation are the worst ones, in particular on the instances generated via the new method.

6.4. **MIP-Tests for the CKVLP and CCMKP Instances.** We now turn to the results of the MIP-tests for the CKVLP and CCMKP instances. We test four configurations of the solver:

MIP: MIP solver without inequality separation;

MIP-MD: MIP solver with MD-VIs separated using the cutting-plane method;

MIP-MS: MIP solver with the mixing-set inequalities using the cutting-plane method;

MIP-Q: MIP solver with the quantile cuts separated using the cutting-plane method.

In all configurations, we leave active the standard inequalities automatically separated by Gurobi.

Table 3 reports the results of the four configurations on the CKVLP instances. The meaning of the rows and the first three columns reporting the characteristics of the instances are the same as in Table 1. For each configuration, we report the following statistics:

gap/time: Average computational time if all the instances in the group are solved to optimality. Otherwise average optimality gap expressed as a percentage returned by the solver with the number of instances solved to optimality in parenthesis.

cuts: Average number of violated inequalities inserted in the formulation for the three configurations in which inequalities are separated.

There is no clear winner between the configurations on the instances with n-m equal to 20-1. First, MIP-MD and MIP-Q detect only few violated inequalities (less than 3 on average for all the subset of instances) which is due to the structure of the instances (one constraint per scenario). On the contrary, MIP-MS detects many violated inequalities. All configurations are able to solve to optimality the instances generated with method lit and with $|\mathcal{S}| \leq 500$ as well as those generated with method new and with $|\mathcal{S}| \leq 1000$. The average computational times to obtain the optimal solutions are comparable among all the configurations. We observe that MIP-Q solves to optimality all the instances generated with method lit and with $|\mathcal{S}| = 1000$, whereas all the other configurations fail to solve one of the instances in the subset. When the number of scenarios is equal to 3000, MIP-MD provides the largest number of optimal solutions (three out of five) and the best average optimality gap (1.61 %) on the instances generated with method lit. However, even if

Table 3. Results of the MIP with MD-VIs, the mixing-set inequalities, and the quantile cuts for CKVLP instances.

Instances			MIP	MIP-MD		MIP-MS		MIP-Q	
n-m	Gen.	$\|\mathcal{S}\|$	gap/time	cuts	gap/time	cuts	gap/time	cuts	gap/time
20-1	lit	100	**0.24s**	2.80	0.73s	24.80	0.48s	1.40	0.37s
		500	3.35s	2.40	**1.01s**	100.00	4.74s	1.00	2.49s
		1000	0.03%(4)	2.40	0.10%(4)	199.40	0.06%(4)	1.40	**272.61s**
		3000	3.18%(2)	2.20	**1.61%(3)**	501.00	4.43%(1)	2.40	4.75%(0)
	new	100	**0.13s**	2.00	0.28s	24.40	0.34s	1.20	0.28s
		500	10.01s	2.00	1.18s	104.00	2.07s	1.00	**0.91s**
		1000	20.57s	2.00	1.67s	151.60	1.65s	1.00	**1.22s**
		3000	**0.72%(4)**	2.00	0.97%(4)	466.00	0.97%(3)	1.20	1.25%(4)
80-1	lit	100	0.98s	3.40	**0.91s**	30.60	1.31s	1.40	0.97s
		500	0.58%(4)	3.00	**0.03%(4)**	213.60	0.61%(4)	1.60	0.58%(4)
		1000	3.01%(1)	3.40	**1.94%(1)**	252.80	2.62%(2)	2.20	3.29%(0)
		3000	0.90%(1)	2.20	**0.13%(4)**	271.20	2.14%(2)	1.20	0.52%(3)
	new	100	**0.59s**	2.20	**0.59s**	44.00	1.19s	2.20	0.68s
		500	12.42s	2.00	**4.95s**	161.60	22.85s	1.60	6.06s
		1000	0.01%(4)	2.00	**472.05s**	281.80	0.21%(4)	2.40	0.34%(4)
		3000	1.02%(3)	2.00	0.76%(4)	459.60	1.55%(2)	1.60	**0.53%(4)**

it provides the same number of optimal solutions of MIP on the instances generated with method new (four out of five), the latter returns the smallest average optimality gap (0.72 %).

MIP-MD yields the most consistent results on the instances with n-m equal to 80-1. Again all configurations are able to solve to optimality with comparable computational times the instances generated with method lit and with $|\mathcal{S}| = 100$ as well as those generated with method new and with $|\mathcal{S}| \leq 500$. However, MIP-MD returns the best average optimality gaps on the remaining instances generated with method lit and is the only configuration that solves all the instances generated with new and with $|\mathcal{S}| = 1000$ to optimality. Again, when the instances are generated with new and the number of scenarios is 3000, MIP-MD provides four optimal solutions as it is the case for MIP-Q as well. However, the latter returns better optimality gaps for the instance not solved to optimality. The observation made for the instances with n-m being 20-1 regarding the number of violated inequalities detected by the configurations applies here as well: Both MIP-MD and MIP-Q detect few violated inequalities while MIP-MS detects several of them.

Table 4 reports the results of the four configurations for the CCMKP instances, where the x-variables are binary. For configuration MIP-MD, we evaluate both separation heuristics for the MD-VIs, namely the ADM and the C-wise approach. The instances are grouped as in Table 2 and the columns' headings are the same as in Table 3.

The structure of these instances induce many more violated valid inequalities compared to the CKVLP instances. Indeed, in general, MIP-MD (regardless from the separation heuristic) and MIP-Q detect tenths of violated inequalities and MIP-MS detects several hundredths and up to 1700 of them, on average.

The instances with n-m being 20-10 are easy to be solved if the number of scenarios is less than 1000. Indeed, all configurations can solve all those instances to optimality in comparable computational time. MIP-MD with separation heuristic C-wise and MIP-Q are the only configurations that solve to optimality all the instances generated with lit and with $|\mathcal{S}| = 3000$. However, the former provides these optimal solutions in half of the computational time required by MIP-Q. MIP-MD with

TABLE 4. Results of the MIP-tests with MD-VIs, the mixing-set inequalities, and the quantile cuts for the CCMKP instances, where the x-variables are binary.

Instances			MIP	MIP-MD				MIP-MS		MIP-Q	
				ADM		C-wise					
n-m	Gen.	$\|\mathcal{S}\|$	gap/time	cuts	gap/time	cuts	gap/time	cuts	gap/time	cuts	gap/time
20-10	lit	100	9.28s	17.80	7.47s	18.60	**4.31s**	202.80	6.34s	15.60	5.49s
		500	110.42s	12.80	46.33s	20.80	**21.53s**	380.80	60.40s	14.60	39.13s
		1000	642.60s	14.00	**111.55s**	17.20	120.40s	674.00	260.68s	18.20	164.71s
		3000	12.03%(1)	10.80	2.19%(4)	17.80	**696.69s**	1538.00	4.60%(4)	17.60	1500.19s
	new	100	14.43s	20.60	13.01s	10.20	8.40s	166.20	**8.01s**	3.60	13.11s
		500	91.75s	29.80	**62.04s**	12.40	70.99s	503.60	72.04s	6.60	91.51s
		1000	388.66s	14.80	**254.66s**	7.20	296.24s	593.40	356.02s	6.80	442.39s
		3000	41.46%(0)	21.20	**32.92%(0)**	9.40	33.83%(1)	1552.40	37.82%(0)	7.40	52.73%(0)
40-30	lit	100	10.17%(0)	38.80	9.16%(1)	46.60	**8.32%(1)**	214.80	11.73%(0)	44.00	12.05%(0)
		500	29.22%(0)	20.00	5.73%(2)	41.20	**1.81%(4)**	430.00	14.44%(0)	35.60	15.76%(0)
		1000	52.17%(0)	46.40	**20.85%(0)**	46.80	24.03%(0)	801.00	29.35%(0)	52.60	34.95%(0)
		3000	41.63%(0)	19.80	**15.07%(0)**	38.80	18.14%(0)	1412.20	24.66%(0)	34.80	29.13%(0)
	new	100	19.70%(0)	62.80	22.23%(0)	13.60	**18.79%(0)**	185.00	21.98%(0)	6.60	19.05%(0)
		500	39.24%(0)	30.80	**20.45%(0)**	11.00	32.42%(0)	387.00	27.21%(0)	10.00	34.72%(0)
		1000	48.89%(0)	38.80	**29.81%(0)**	16.20	55.00%(0)	700.40	39.56%(0)	16.80	52.94%(0)
		3000	49.05%(0)	31.20	**27.68%(0)**	17.20	48.18%(0)	1694.80	42.26%(0)	16.00	50.00%(0)

separation heuristic ADM shows the best average optimality gaps for the instances generated with new method and with $|\mathcal{S}| = 3000$.

The instances with n-m being 40-30 are harder to solve. Indeed, there is no group of instances for which all instances are solved to optimality by any configuration. However, separating the MD-VIs yields the best results. Indeed, MIP-MS and MIP-Q do not provide any optimal solutions, whereas MIP-MD is able to provide at least four out of five. In addition, the winner configuration w.r.t. optimality gaps is MIP-MD with either the ADM or the C-wise heuristic. The ADM heuristic performs better on the instances generated with method new. However, it is not clear which of the two separation heuristics performs better in general.

Table 5 reports the results of the four configurations on CCMKP instances, for which the x-variables are continuous. The rows and the columns' headings are the same as in Table 4. The instances of the CCMKP, which are characterized by multiple constraints per scenario and continuous variables are the hardest to be solved. All configurations are able to solve to optimality all the instance of two groups, only, namely the ones with n-m being 20-10 and $|\mathcal{S}| = 100$. The computational times to obtain the optimal solutions are comparable. For the remaining 14 groups of instances, MIP-MD provides 4 optimal solutions regardless of the separation heuristic, whereas MIP-MS and MIP-Q provide 5 optimal solutions each. However, separating the MD-VIs leads to the best optimality gaps on 10 out of the 14 groups of instances. It seems that the more consistent separation approach for the MD-VIs is ADM—at least on the largest instances, i.e., those with n-m being 40-30. The number of violated inequalities detected by the configurations follow the trend observed in Table 4.

7. CONCLUSION

In this paper we studied mixed-integer linear chance-constrained problems (CCPs) in which the feasible region is parameterized by a random vector, which is supposed

TABLE 5. Results of the MIP with MD-VIs, the mixing-set inequalities, and the quantile cuts for the CCMKP instances, where the x-variables are continuous.

Instances			MIP	MIP-MD				MIP-MS		MIP-Q	
				ADM		C-wise					
n-m	Gen.	$\|\mathcal{S}\|$	gap/time	cuts	gap/time	cuts	gap/time	cuts	gap/time	cuts	gap/time
20-10	lit	100	204.87s	17.80	172.64s	18.60	**114.96s**	202.80	202.04s	15.60	117.55s
		500	3.35%(0)	12.80	2.76%(0)	20.80	**2.20%(0)**	380.80	2.60%(0)	14.60	2.31%(0)
		1000	11.92%(0)	14.00	8.61%(0)	17.20	**6.99%(0)**	674.00	7.12%(0)	18.20	8.85%(0)
		3000	32.96%(0)	10.80	**9.40%(0)**	17.80	10.69%(0)	1538.00	11.27%(0)	17.60	14.21%(0)
	new	100	313.56s	20.60	**227.12s**	10.20	305.02s	166.20	368.67s	3.60	292.92s
		500	10.10%(0)	29.80	10.83%(0)	12.40	**9.47%(0)**	503.60	9.73%(0)	6.60	9.87%(0)
		1000	16.77%(0)	14.80	14.60%(0)	7.20	13.25%(0)	593.40	**13.17%(0)**	6.80	13.21%(0)
		3000	44.58%(0)	21.20	**19.84%(0)**	9.40	26.31%(0)	1552.40	29.96%(0)	7.40	24.86%(0)
40-30	lit	100	0.15%(2)	38.80	0.21%(3)	46.60	0.12%(3)	214.80	0.14%(4)	44.00	**0.08%(4)**
		500	10.35%(0)	20.00	8.49%(0)	41.20	8.67%(0)	430.00	6.95%(0)	35.60	**6.56%(0)**
		1000	24.77%(0)	46.40	**13.21%(0)**	46.80	14.44%(0)	801.00	15.64%(0)	52.60	14.38%(0)
		3000	35.04%(0)	19.80	**11.91%(0)**	38.80	15.44%(0)	1412.20	18.89%(0)	34.80	20.17%(0)
	new	100	1.47%(1)	62.80	**0.80%(1)**	13.60	1.26%(1)	185.00	1.31%(1)	6.60	0.94%(1)
		500	16.61%(0)	30.80	15.14%(0)	11.00	14.36%(0)	387.00	**13.36%(0)**	10.00	13.60%(0)
		1000	30.26%(0)	38.80	**19.85%(0)**	16.20	21.40%(0)	700.40	24.04%(0)	16.80	21.83%(0)
		3000	41.97%(0)	31.20	**18.78%(0)**	17.20	33.37%(0)	1694.80	36.25%(0)	16.00	30.81%(0)

to have finite support. We introduced two families of valid inequalities. We coined the inequalities of the first family as *primal-dual* valid inequalities, while we called the second family of inequalities *covering* valid inequalities. If re-scaling the latter inequalities is possible, we obtained what we called the *multi-disjunctive* valid inequalities. These generalize the simple-disjunctive valid inequalities previously proposed in the literature [34].

From a theoretical point of view, we conducted a dominance analysis among the families of valid inequalities and we proved that covering valid inequalities dominate primal-dual valid inequalities. Moreover, we presented theoretical results on the closure of the multi-disjunctive valid inequalities in comparison to the closure of the so-called quantile cuts known from the literature [28, 34, 42]. Finally, we showed that the problem of separating covering valid inequalities is NP-hard.

From a computational point of view, driven by the theoretical complexity of the separation problem of the covering valid inequalities, we proposed heuristic procedures to find violated cuts. Lastly, with the aim to compare the impact of the valid inequalities with those from the literature (e.g., the mixing-set valid inequalities and the quantile cuts), we conducted extensive numerical experiments on instances for the covering-type k-violation linear program (CKVLP) and the chance-constrained multidimensional knapsack problem. To this end, we generated two sets of instances: one with a method taken from the literature and a second, newly generated, set. We then ran two computational tests. One to evaluate the impact of the multi-disjunctive valid inequalities on improving the dual bound at the root node and one to evaluate the performance of a MIP solver to solve formulation (3) when multi-disjunctive valid inequalities are added to the root node of the branch-and-bound tree. Both computational tests highlighted the effectiveness of the proposed valid inequalities.

Last but not least, our heuristic separation procedures are generic and may be applied to other linear CCPs with finite support. This improves over the previous literature. Indeed, even if heuristic procedures to separate the quantile cuts are available, experiments revealed that they are rather ineffective for the instances that we considered. On the contrary, separating mixing-set inequalities poses two major challenges; selecting an appropriate direction η and computing the β_η^s coefficients. Even if η is selected, computing the β coefficients generally requires solving $|\mathcal{S}|$ linear problems, which can be computationally prohibitive as $|\mathcal{S}|$ grows.

References

[1] S. Ahmed, J. Luedtke, Y. Song, and W. Xie. “Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs.” In: *Mathematical Programming* 162.1 (2017), pp. 51–81. DOI: 10.1007/s10107-016-1029-z.

[2] S. Ahmed and A. Shapiro. “Solving Chance-Constrained Stochastic Programs via Sampling and Integer Programming.” In: *INFORMS Tutorials in Operations Research* (2008), pp. 261–269. DOI: 10.1287/educ.1080.0048.

[3] G. J. Alexander and A. M. Baptista. “A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model.” In: *Management Science* 50.9 (2004), pp. 1261–1273.

[4] G. J. Alexander and A. M. Baptista. “Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis.” In: *Journal of Economic Dynamics and Control* 26.7-8 (2002), pp. 1159–1193.

[5] S. Benati and R. Rizzi. “A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem.” In: *European Journal of Operational Research* 176 (2007), pp. 423–434. DOI: 10.1016/j.ejor.2005.07.020.

[6] P. Beraldi and A. Ruszczynski. “A Branch and Bound Method for Stochastic Integer Problems under Probabilistic Constraints.” In: *Optimization Methods and Software* 17 (2002), pp. 359–382. DOI: 10.1080/1055678021000033937.

[7] P. Beraldi and A. Ruszczynski. “The Probabilistic Set-Covering Problem.” In: *Operations Research* 50.6 (2002), pp. 956–967.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. 2011. DOI: 10.1561/2200000016.

[9] D. Cattaruzza, M. Labbé, M. Petris, M. Roland, and M. Schmidt. “Exact and Heuristic Solution Techniques for Mixed- Integer Quantile Minimization Problems.” In: *INFORMS Journal on Computing* 36.4 (2024), pp. 1084–1107. DOI: 10.1287/ijoc.2022.0105.

[10] A. Charnes and W. Cooper. “Chance-Constrained Programming.” In: *Management Science* 6.1 (1959), pp. 73–79. DOI: 10.1287/mnsc.6.1.73.

[11] A. Charnes, W. Cooper, and G. Symonds. “Cost Horizon and Certainty Equivalents: An Approach to Stochastic Programming Heating Oil.” In: *Management Science* 4.3 (1958), pp. 253–263. DOI: 10.1287/mnsc.4.3.235.

[12] J. Cho and A. Papavasiliou. “Exact Mixed-Integer Programming Approach for Chance-Constrained Multi-Area Reserve Sizing.” In: *IEEE Transactions on Power Systems* 39.3 (2024), pp. 3310–3323. DOI: 10.1109/TPWRS.2023.3279692.

[13] D. Dentcheva, A. Prèkopa, and A. Ruszczynski. “Concavity and efficient points of discrete distributions in probabilistic programming.” In: *Mathematical Programming* 89 (2000), pp. 55–77. DOI: 10.1007/PL00011393.

[14] A. A. Gaivoronski and G. Pflug. “Finding optimal portfolios with constraints on value at risk.” In: *Proceedings III Stockholm seminar on risk behavior and risk management*. 1999.

[15] A. A. Gaivoronski and G. Pflug. "Value-at-Risk in Portfolio Optimization: Properties and Computational Approach." In: *Journal of Risk* 7.2 (2005), pp. 1–31.

[16] M. Garey and D. Johnson. *Computers and Intractability*. Vol. 179. San Francisco: Freeman, 1979. DOI: 10.5555/574848.

[17] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving Highly Detailed Gas Transport MINLPs: Block Separability and Penalty Alternating Direction Methods." In: *INFORMS Journal on Computing* 30.2 (2018), pp. 309–323. DOI: 10.1287/ijoc.2017.0780.

[18] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving power-constrained gas transportation problems using an MIP-based alternating direction method." In: *Computers & Chemical Engineering* 82 (2015), pp. 303–317. DOI: 10.1016/j.compchemeng.2015.07.005.

[19] R. Glowinski and A. Marroco. "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires." In: *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* 9.R2 (1975), pp. 41–76.

[20] O. Günlük and Y. Pochet. "Mixing mixed-integer inequalities." In: *Mathematical Programming* 90 (2001), pp. 429–457. DOI: 10.1007/s101070100224.

[21] N. Jiang and W. Xie. "ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs." In: *INFORMS Journal on Computing* 70.6 (2022), pp. 3581–3600. DOI: 10.1287/opre.2021.2225.

[22] T. Kleinert and M. Schmidt. "Computing Feasible Points of Bilevel Problems with a Penalty Alternating Direction Method." In: *INFORMS Journal on Computing* 33.1 (2019), pp. 198–215. DOI: 10.1287/ijoc.2019.0945.

[23] S. Kücükyavuz. "On mixing set arising in chance-constrained programming." In: *Mathematical Programming* 132 (2012), pp. 31–56. DOI: 10.1007/s10107-010-0385-3. Springer.

[24] S. Kücükyavuz and R. Jiang. "Chance-constrained optimization under limited distributional information: A review of reformulations based on sampling and distributional robustness." In: *EURO Journal on Computational Optimization* 10.100030 (2022), pp. 1–45. DOI: 10.1016/j.ejco.2022.100030. Elsevier.

[25] A. Lejeune. "Pattern-Based Modeling and Solution of Probabilistically Constrained Optimization Problems." In: *Operations Research* 60.6 (2012), pp. 1356–1372. DOI: 10.1287/opre.1120.1120.

[26] C.-C. Lin. "Comments on "A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem"." In: *European Journal of Operational Research* 194 (2009), pp. 339–341.

[27] R. Liu, Z. Lin, and Z. Su. "Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning." In: *Asian Conference on Machine Learning*. PMLR. 2013, pp. 116–132. DOI: 10.1007/s10994-014-5469-5.

[28] J. Luedtke. "A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support." In: *Mathematical Programming* 146.1–2 (2014), pp. 219–244. DOI: 10.1007/s10107-013-0684-6.

[29] J. Luedtke, S. Ahmed, and G. L. Nemhauser. "An integer programming approach for linear programs with probabilistic constraints." In: *Mathematical Programming* 122 (2010), pp. 247–272. DOI: 10.1007/s10107-008-0247-4.

[30] R. Mansini, W. Ogryczak, and M. G. Speranza. "LP solvable models for portfolio optimization: A classification and computational comparison." In:

IMA Journal of Management Mathematics 14.3 (2003), pp. 187–220. DOI: 10.1093/imaman/14.3.187.
[31] B. L. Miller and H. M. Wagner. "Chance Constrained Programming with Joint Constraints." In: *Operations Research* 13.6 (1965), pp. 930–945. DOI: 10.1287/opre.13.6.930.
[32] A. Nemirovski and A. Shapiro. "Convex Approximations of Chance Constrained Programs." In: *SIAM Journal on Optimization* 17.4 (2007), pp. 969–996. DOI: 10.1137/050622328.
[33] A. Porras, C. Domìnguez, J. M. Morales, and S. Pineda. "Tight and Compact Sample Average Approximation for Joint Chance-Constrained Problems with Applications to Optimal Power Flow." In: *INFORMS Journal on Computing* 35.6 (2023), pp. 1454–1469. DOI: 10.1287/ijoc.2022.0302.
[34] F. Qiu, S. Ahmed, S. S. Dey, and L. A. Wolsey. "Covering linear programming with violations." In: *INFORMS Journal on Computing* 26.3 (2014), pp. 531–546. DOI: 10.1287/ijoc.2013.0582.
[35] R. T. Rockafellar and S. Uryasev. "Optimization of conditional value-at-risk." In: *Journal of Risk* 2.3 (2000), pp. 21–41. DOI: 10.21314/JOR.2000.038.
[36] M. Roland, A. Forel, and T. Vidal. "Adaptive partitioning for chance-constrained problems with finite support." In: *SIAM Journal on Optimization* 33.1 (2025), pp. 531–546. DOI: 10.1137/24M1632772.
[37] L. Schewe, M. Schmidt, and D. Weninger. "A decomposition heuristic for mixed-integer supply chain problem." In: *Operations Research Letters* 48.3 (2020), pp. 225–232. DOI: 10.1016/j.orl.2020.02.006.
[38] A. Sexena, V. Goyal, and M. A. Lejeune. "MIP reformulations of the probabilistic set covering problem." In: *Mathematical Programming* 121 (2010), pp. 1–31. DOI: 10.1007/s10107-008-0224-y.
[39] Y. Song and J. R. Luedtke. "Branch-and-cut approaches for chance-constrained formulations of reliable network design problems." In: *Mathematical Programming Computation* 5 (2013), pp. 397–432. DOI: 10.1007/s12532-013-0058-3.
[40] Y. Song, J. R. Luedtke, and S. Kücükyavuz. "Chance-Constrained Binary Packing Problems." In: *INFORMS Journal on Computing* 26.4 (2014), pp. 735–747. DOI: 10.1287/ijoc.2014.0595.
[41] M. W. Tanner and L. Ntaimo. "IIS branch-and-cut for joint chance-constrained stochastic programs and application to optimal vaccine allocation." In: *European Journal of Operational Researcgh* 207.1 (2010), pp. 290–296. DOI: 10.1016/j.ejor.2010.04.019.
[42] W. Xie and S. Ahmed. "On quantile cuts and their closure for chance constrained optimization problems." In: *Mathematical Programming* 172 (2018), pp. 621–646. DOI: 10.1007/s10107-017-1190-z.
[43] Z. Zhang, C. Gao, and J. Luedtke. "New valid inequalities and formulations for the static joint Chance-constrained Lot-sizing problem." In: *Mathematical Programming* 199 (2023), pp. 639–669. DOI: 10.1007/s10107-022-01847-y.

Appendix A. No Dominance Relation Between Inequalities (7) and Inequalities (8)

The basic setup is given by $\mathcal{S} = \{1, 2\}$, $p_1 = p_2 = \frac{1}{2}$, and $\tau = \frac{1}{2}$.

A.1. Case 1.

We consider

$$A^1 = \begin{bmatrix} 1 & 3 \\ 2.5 & 2 \end{bmatrix}, \quad A^1_{\text{scaled}} = \begin{bmatrix} 0.8 & 2.4 \\ 5 & 4 \end{bmatrix}, \quad b^1 = \begin{pmatrix} 1.25 \\ 0.5 \end{pmatrix}$$

and

$$A^2 = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}, \quad A^2_{\text{scaled}} = \begin{bmatrix} 1.333 & 2 \\ 1.333 & 0.667 \end{bmatrix}, \quad b^2 = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix}.$$

Furthermore, we assume $\bar{\mathcal{S}} = \mathcal{S}$, $\lambda^1 = \lambda^2 = (0.5, 0.5)$ and obtain $U_1(\lambda, \bar{\mathcal{S}}) = 2$, $U_2(\lambda, \bar{\mathcal{S}}) = 2.5$, as well as $L(\lambda, \bar{\mathcal{S}}) = 0.875$, while $V_1(\lambda, \bar{\mathcal{S}}) = 2.9$, $V_2(\lambda, \bar{\mathcal{S}}) = 3.2$. Thus,

$$\begin{aligned} U_1(\lambda, \bar{\mathcal{S}})x_1 + U_2(\lambda, \bar{\mathcal{S}})x_2 &\geq L(\lambda, \bar{\mathcal{S}}), \\ 2x_1 + 2.5x_2 &\geq 0.875, \\ 2.285714286x_1 + 2.857142857x_2 &\geq 1 \end{aligned}$$

dominates

$$\begin{aligned} V_1(\lambda, \bar{\mathcal{S}})x_1 + V_2(\lambda, \bar{\mathcal{S}})x_2 &\geq 1, \\ 2.9x_1 + 3.2x_2 &\geq 1. \end{aligned}$$

A.2. **Case 2.** We now consider

$$A^1 = \begin{bmatrix} 1 & 3 \\ 2.5 & 2 \end{bmatrix}, \quad A^1_{\text{scaled}} = \begin{bmatrix} 0.8 & 2.4 \\ 2.0833 & 1.6667 \end{bmatrix}, \quad b^1 = \begin{pmatrix} 1.25 \\ 1.2 \end{pmatrix}$$

and

$$A^2 = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}, \quad A^2_{\text{scaled}} = \begin{bmatrix} 1.333 & 2 \\ 1.667 & 0.8333 \end{bmatrix}, \quad b^2 = \begin{pmatrix} 1.5 \\ 1.2 \end{pmatrix}.$$

We again assume $\bar{\mathcal{S}} = \mathcal{S}$ and $\lambda^1 = \lambda^2 = (0.5, 0.5)$ and obtain $U_1(\lambda, \bar{\mathcal{S}}) = 2$, $U_2(\lambda, \bar{\mathcal{S}}) = 2.5$, as well as $L(\lambda, \bar{\mathcal{S}}) = 1.225$, while $V_1(\lambda, \bar{\mathcal{S}}) = 1.5$, $V_2(\lambda, \bar{\mathcal{S}}) = 2.033$. Thus,

$$\begin{aligned} U_1(\lambda, \bar{\mathcal{S}})x_1 + U_2(\lambda, \bar{\mathcal{S}})x_2 &\geq L(\lambda, \bar{\mathcal{S}}), \\ 2x_1 + 2.5x_2 &\geq 1.225, \\ 1.632653061x_1 + 2.040816327x_2 &\geq 1 \end{aligned}$$

is dominated by

$$\begin{aligned} V_1(\lambda, \bar{\mathcal{S}})x_1 + V_2(\lambda, \bar{\mathcal{S}})x_2 &\geq 1, \\ 1.5x_1 + 2.0333x_2 &\geq 1. \end{aligned}$$

A.3. **Case 3.** Finally, we consider

$$A^1 = \begin{bmatrix} 1 & 3 \\ 2.5 & 2 \end{bmatrix}, \quad A^1_{\text{scaled}} = \begin{bmatrix} 0.8333 & 2.5 \\ 1.7857149 & 1.42857143 \end{bmatrix}, \quad b^1 = \begin{pmatrix} 1.2 \\ 1.4 \end{pmatrix}$$

and

$$A^2 = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}, \quad A^2_{\text{scaled}} = \begin{bmatrix} 1.333 & 2 \\ 1.667 & 0.8333 \end{bmatrix}, \quad b^2 = \begin{pmatrix} 1.5 \\ 1.2 \end{pmatrix}.$$

We again have $\bar{\mathcal{S}} = \mathcal{S}$ and $\lambda^1 = \lambda^2 = (0.5, 0.5)$ and obtain $U_1(\lambda, \bar{\mathcal{S}}) = 2$, $U_2(\lambda, \bar{\mathcal{S}}) = 2.5$, as well as $L(\lambda, \bar{\mathcal{S}}) = 1.3$, while $V_1(\lambda, \bar{\mathcal{S}}) = 1.5$, $V_2(\lambda, \bar{\mathcal{S}}) = 1.964285714$. Thus,

$$\begin{aligned} U_1(\lambda, \bar{\mathcal{S}})x_1 + U_2(\lambda, \bar{\mathcal{S}})x_2 &\geq L(\lambda, \bar{\mathcal{S}}), \\ 2x_1 + 2.5x_2 &\geq 1.3, \\ 1.538461538x_1 + 1.923076923x_2 &\geq 1 \end{aligned}$$

and

$$\begin{aligned} V_1(\lambda, \bar{\mathcal{S}})x_1 + V_2(\lambda, \bar{\mathcal{S}})x_2 &\geq 1, \\ 1.5x_1 + 1.964285714x_2 &\geq 1, \end{aligned}$$

which shows that there is no dominance relation between the two constraints.

Appendix B. On the Benefit of $\sum_{i=1}^{n} \lambda_i^s = 1$

Proposition 8. *Model (13) without the normalization (13c) is unbounded.*

Proof. We show that there exists a solution to Model (13) for which the objective function can be driven to $-\infty$. Let $\bar{x}$ be a point that is infeasible for Model (13). Define $S_I(\bar{x}) \subseteq S$ as the set of scenarios that $\bar{x}$ does not satisfy, i.e.,

$$S_I(\bar{x}) = \left\{ s \in S \colon \exists j \in [m] : A_j^s \bar{x} \leq b_j^s \right\}.$$

Let $\bar{S} = S_I(\bar{x})$. Select an arbitrary scenario $s' \in \bar{S}$. For this scenario s', there exists an index $k \in [m]$ such that $A_{k\cdot}^{s'} \bar{x} \leq b_k^{s'}$. Let $\mu \geq 0$ be an arbitrary constant. We now define the multipliers λ as follows:

$$\lambda_k^{s'} = \mu \quad \text{and} \quad \lambda_j^s = 0 \quad \text{for all other } s \in S,\, j \in [m].$$

By the definitions of $U_i(\lambda, \bar{S})$ and $L(\lambda, \bar{S})$ we have

$$\begin{aligned} &\sum_{i \in [n]} U_i(\lambda, \bar{S}) \bar{x}_i - L(\lambda, \bar{S}) \\ &= \sum_{i \in [n]} \lambda_k^{s'} (p(\bar{S}) - \tau)(A_{ki}^{s'}) \bar{x}_i - \lambda_k^{s'} (p(\bar{S}) - \tau)(b_k^{s'}) \\ &= \mu (p(\bar{S}) - \tau) \left(\sum_{i \in [n]} A_{ki}^{s'} \bar{x}_i - b_k^{s'} \right). \end{aligned}$$

Since $\sum_{i \in [n]} A_{ki}^{s'} \bar{x}_i - b_k^{s'} \leq 0$, the right-hand side of the above equation satisfies

$$\mu (p(\bar{S}) - \tau) \left(\sum_{i \in [n]} A_{ki}^{s'} \bar{x}_i - b_k^{s'} \right) \leq 0.$$

As $\mu \to \infty$, the right-hand side tends to $-\infty$ and we obtain the desired result. □

(D. Cattaruzza) (A) Department of Mathematics, Computer Science and Physics, University of Udine, Via delle Scienze 206, 33100 Udine, Italy (B) Univ. Lille, CNRS, Centrale Lille, Inria, UMR 9189 - CRIStAL Lille, France

Email address: diego.cattaruzza@uniud.it

(M. Labbé) (A) Université Libre de Bruxelles, Department of Computer Science, Boulevard du Triomphe, CP212, 1050 Brussels, Belgium; (B) Inria Lille - Nord Europe, Parc scientifique de la Haute Borne, 40, av. Halley - Bât A - Park Plaza, 59650 Villeneuve d'Ascq, France

Email address: martine.labbe@ulb.be

(M. Petris) CERMICS, ENPC, Institut Polytechnique de Paris, Marne-la-Vallée, France

Email address: matteo.petris@enpc.fr

(M. Roland) (A) Polytechnique Montréal, André-Aisenstadt Pavillon, 2920 Tour Road, H3T 1N8 Montréal, Canada; (B) Inria Lille - Nord Europe, Parc scientifique de la Haute Borne, 40, av. Halley - Bât A - Park Plaza, 59650 Villeneuve d'Ascq, France

Email address: mmmroland@gmail.com

(M. Schmidt) Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany

Email address: martin.schmidt@uni-trier.de