Valid Inequalities and Restrictions for Stochastic Programming Problems with First Order Stochastic Dominance Constraints

Nilay Noyan* Andrzej Ruszczyński†

March 21, 2006

Abstract

Stochastic dominance relations are well-studied in statistics, decision theory and economics. Recently, there has been significant interest in introducing dominance relations into stochastic optimization problems as constraints. In the discrete case, stochastic optimization models involving second order stochastic dominance (SSD) constraints can be solved by linear programming (LP). However, problems involving first order stochastic dominance (FSD) constraints are potentially hard due to the non-convexity of the associated feasible regions. In this paper we consider a mixed 0–1 linear programming formulation of a discrete FSD-constrained optimization model and present an LP relaxation based on SSD constraints. We derive some valid inequalities and restrictions by employing the probabilistic structure of the problem. We also generate cuts that are valid inequalities for the disjunctive relaxations arising from the underlying combinatorial structure of the problem by applying the lift-and-project procedure. We describe three heuristic algorithms to construct feasible solutions, based on conditional SSD-constraints, variable fixing, and conditional value at risk. Finally, we present numerical results for several instances of a portfolio optimization problem.

Keywords: *Stochastic programming, Stochastic dominance, Valid inequalities, Disjunctive cuts, Conditional value at risk.*

Introduction

The notion of first order stochastic dominance (FSD), also called stochastic ordering, has been introduced in statistics (see [22, 24]) and further applied and developed in economics and decision theory (see [33, 17, 19, 35, 15, 38, 23]). A random variable X *dominates* another random variable Y in the first order, which we denote by $X \succeq_{(1)} Y$, if

$$F_X(\eta) \le F_Y(\eta) \quad \text{for all} \quad \eta \in \mathbb{R}, \tag{1}$$

*RUTCOR, Rutgers University, 640 Bartholomew Rd., Piscataway, New Jersey 08854, U.S.A., E-mail: noyan@rutcor.rutgers.edu

†Department of Management Science and Information Systems, Rutgers University, 94 Rockefeller Rd., Piscataway, New Jersey 08854, U.S.A., E-mail: rusz@business.rutgers.edu

where $F_X(\eta) = P[X \leq \eta]$ denotes the distribution function of a random variable X. It is well-known (see, e.g., [27]) that the stochastic ordering relation $X \succeq_{(1)} Y$ can be equivalently expressed as follows:

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$$

holds for all non-decreasing functions $u : \mathbb{R} \to \mathbb{R}$ for which the above expectations are finite.

For two integrable random variables X and Y, $X \succeq_{(2)} Y$ denotes the second order stochastic dominance (SSD) relation:

$$\int_{-\infty}^{\eta} F_X(\alpha)\, d\alpha \leq \int_{-\infty}^{\eta} F_Y(\alpha)\, d\alpha \quad \text{for all} \quad \eta \in \mathbb{R}. \tag{2}$$

It is easy to see that the FSD relation (1) implies SSD via integration over the interval $(-\infty, \eta)$. We are also going to use the following equivalent formulation of the SSD condition, featuring the expected shortfalls of the random variables (see [10]):

$$\mathbb{E}([\eta - X]_+) \leq \mathbb{E}([\eta - Y]_+) \quad \text{for all} \quad \eta \in \mathbb{R}. \tag{3}$$

Dominance relations can be involved in stochastic optimization problems as constraints, allowing us to obtain solutions dominating some random benchmark outcomes. Such models have been introduced and analyzed in [10, 11, 13]. In [12, 13] it is also proved that the convexification of the set defined by the FSD relation is equal to the set defined by the SSD relation, if the probability space is non-atomic. In [10, 11] the properties of stochastic optimization models involving SSD constraints are studied in detail, along with numerical illustrations for the case where the reference random variable Y has a discrete distribution with finitely many realizations.

In this paper we focus on optimization models involving FSD constraints, which are equivalent to a continuum of probabilistic constraints. They pose two main challenges. The first challenge is the potential non-convexity of the associated feasible region. Secondly, the model involving FSD constraints is a semi-infinite optimization problem. The probabilistic nature of the problem allows specific analysis which goes beyond the existing theory of semi-infinite optimization (see, e.g., [20]), both in optimality theory and in the analysis of convex relaxations. We focus on problems with discrete distributions and we develop new valid inequalities by employing the probabilistic and combinatorial structures of our problem. We also describe heuristic algorithms to construct feasible solutions.

In Section 1 we present a mixed 0–1 linear programming formulation with multiple knapsack constraints of a discrete FSD-constrained optimization model. In Section 2 an LP relaxation, based on SSD constraints, is presented, which is different from the usual LP relaxation obtained by dropping the integrality restriction. In Section 2 we also present some valid inequalities and restrictions, which are derived by employing the probabilistic structure of the problem. In Section 3 we derive valid inequalities by considering the combinatorial structure of our problem and by applying the lift-and-project procedure developed in [5, 6], which has been shown to be an effective way of strengthening mixed 0-1 programming formulations (see [7]). Section 5 describes heuristic algorithms for constructing feasible solutions. We conclude the paper by presenting in Section 6 numerical results illustrating the substantial computational efficiency of our methods.

The expected value operator is denoted by $\mathbb{E}$. An abstract probability space is denoted by $(\Omega, \mathscr{F}, P)$, where Ω is the sample space, $\mathscr{F}$ is a σ-algebra on Ω and P is a probability measure on Ω. We denote the cardinality of a set A by $|A|$. For a real number $\eta \in \mathbb{R}$ let $[\eta]_+ = \max(0, \eta)$.

1 Linear optimization problem with first order constraints

Let $(\Omega, 2^{\Omega}, P)$ be a finite probability space, where $\Omega = \{\omega_1, \ldots, \omega_N\}$ with corresponding probabilities $p_1, \ldots, p_N$. Let $\psi : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be an outcome mapping satisfying the condition that $\psi(\cdot, \omega)$ is a linear function for all $\omega \in \Omega$. For a given vector $\mathbf{z} \in \mathbb{R}^n$ let us define the mapping $\psi_{\mathbf{z}} : \Omega \to \mathbb{R}$ by $\psi_{\mathbf{z}}(\omega) = \psi(\mathbf{z}, \omega)$. Let Y be some random variable on Ω (*benchmark outcome*). We consider the following linear stochastic optimization model involving an FSD constraint:

$$\begin{aligned} \max \quad & \mathbf{c}^T \mathbf{z} \\ \text{subject to} \quad & \psi_{\mathbf{z}} \succeq_{(1)} Y, \\ & \mathbf{z} \in Z, \end{aligned} \tag{4}$$

where Z is a compact polyhedron. Without the dominance constraint, problem (4) can be easily formulated as a linear programming problem. The stochastic dominance constraint, however, can render the feasible set nonconvex in the general case.

Let $\hat{y}_i = Y(\omega_i),\ i = 1, \ldots, N$, denote the realizations of the benchmark outcome Y. Without loss of generality we may assume $\hat{y}_1 \le \hat{y}_2 \le \ldots \le \hat{y}_N$. Let us denote the different realizations of the reference outcome Y by $y_1 < y_2 < \ldots < y_D$, and let y_0 be an arbitrary number such that $y_0 < y_1$, implying $P[Y \le y_0] = 0$.

Since the distribution function of the benchmark outcome Y is a right-continuous step function, it is easy to verify that the FSD constraint in (4) is equivalent to

$$P[\psi_{\mathbf{z}} < y_k] \le P[Y \le y_{k-1}], \quad k = 1, \ldots, D. \tag{5}$$

We now reformulate problem (4) by introducing the binary variables

$$\beta_{i,k} = \begin{cases} 1 & \text{if } \psi(\mathbf{z}, \omega_i) < y_k, \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, \ldots, N, \quad k = 1, \ldots, D. \tag{6}$$

Let $M_{i,k} \in \mathbb{R}$ satisfy the inequalities

$$M_{i,k} \ge y_k - \min_{\mathbf{z} \in Z} \psi(\mathbf{z}, \omega_i), \quad i = 1, \ldots, N, \quad k = 1, \ldots, D. \tag{7}$$

Since $\psi(\cdot, \omega_i)$ is linear for all $i = 1, \ldots, N$, and Z is compact, such $M_{i,k}$ values exist.

Now using (5), (6) and (7) we can rewrite (4) as a mixed 0–1 linear programming problem, which we refer to as problem (4):

$$\begin{aligned} \max \quad & \mathbf{c}^T \mathbf{z} \\ \text{subject to} \quad & \sum_{i=1}^{N} p_i \beta_{i,k} \le P[Y \le y_{k-1}], \quad k = 1, \ldots, D, && (8) \\ & \psi(\mathbf{z}, \omega_i) + M_{i,k}\, \beta_{i,k} \ge y_k, \quad i = 1, \ldots, N, \quad k = 1, \ldots, D, && (9) \\ & \beta_{i,k} \in \{0, 1\}, \quad i = 1, \ldots, N, \quad k = 1, \ldots, D, && (10) \\ & \mathbf{z} \in Z. \end{aligned}$$

Note that we can select a sufficiently large common value $M_{i,k} = M,\ i = 1,\dots,N,\ k = 1,\dots,D$.

Inequalities (8) together with the integrality restriction (10) define a set of knapsack constraints. With every item $i \in \{1,\dots,N\}$ we associate a weight $p_i > 0$, and every knapsack $k \in \{1,\dots,D\}$ has a capacity $u_k = P[Y \le y_{k-1}],\ k = 1,\dots,D$.

In the next sections we develop convex approximations by valid inequalities of problem (4). The mixed integer formulation MBLP will be used only for comparison. Our valid inequalities will be formulated exclusively in terms of problem (4), without any use of the binary variables $\beta_{i,k}$.

2 Valid Inequalities based on Second Order Stochastic Dominance

2.1 Second order constraints as valid inequalities

Since FSD implies SSD and, as we are going to see below, SSD constraints can be represented by a system of linear inequalities, the SSD constraints can be viewed as valid inequalities for the feasible set of MBLP. This fact leads to an LP relaxation which is different from the usual one obtained by relaxing the integrality restriction. In [28] we provide examples which show that neither of these relaxations is stronger than the other.

Since Y has a discrete distribution, inequalities (3) are equivalent to (see [10]):

$$\mathbb{E}([y_k - X]_+) \le \mathbb{E}([y_k - Y]_+), \quad k = 1,\dots,D. \tag{11}$$

Let us introduce the following functions, representing the shortfall values of the outcome $\psi_{\mathbf{z}}$:

$$s_{\mathbf{z}}(i,k) = [y_k - \psi(\mathbf{z},\omega_i)]_+, \quad i = 1,\dots,N,\ k = 1,\dots,D. \tag{12}$$

When the policy $\mathbf{z}$ is clear from the context, we also use the simplified notation $s_{i,k} = s_{\mathbf{z}}(i,k)$. Then the next system of linear inequalities is equivalent to the SSD relation (11):

$$\sum_{i=1}^{N} p_i s_{i,k} \le \mathbb{E}([y_k - Y]_+), \quad k = 1,\dots,D. \tag{13}$$

Let $v_k = \mathbb{E}([y_k - Y]_+)$. By introducing the decision vector $\bar{\mathbf{s}} \in \mathbb{R}^{ND}$ to represent the shortfall values we obtain the corresponding LP relaxation of problem (4):

$$\max \quad \mathbf{c}^T\mathbf{z} \tag{14}$$

$$\text{subject to} \quad \psi(\mathbf{z},\omega_i) + \bar{s}_{i,k} \ge y_k, \quad i = 1,\dots,N, \quad k = 1,\dots,D, \tag{15}$$

$$\sum_{i=1}^{N} p_i \bar{s}_{i,k} \le v_k, \quad k = 1,\dots,D, \tag{16}$$

$$\bar{s}_{i,k} \ge 0, \quad i = 1,\dots,N, \quad k = 1,\dots,D, \tag{17}$$

$$\mathbf{z} \in Z. \tag{18}$$

We refer to the above problem as *the SSD-based LP relaxation* of problem (4).

Observation 1. *For every feasible solution* $(\mathbf{z},\bar{\mathbf{s}})$ *of* (14)-(18) *we have* $\bar{\mathbf{s}} \ge \mathbf{s}_{\mathbf{z}}$, *and the pair* $(\mathbf{z},\mathbf{s}_{\mathbf{z}})$ *is also feasible for* (14)-(18).

2.2 Conditional second order stochastic dominance constraints

The second order stochastic dominance relation is a convex constraint. Our objective is to construct valid inequalities, by using this relation for conditional distributions.

Proposition 1. *If $X \succeq_{(1)} Y$ then for every $t \in \mathbb{R}$ and for all $\eta \ge t$*

$$(\eta-t)P[X < t]+P[X \ge t]\,\mathbb{E}\big\{[\eta-X]_+ \,|\, X \ge t\big\} \le (\eta-t)P[Y < t]+P[Y \ge t]\,\mathbb{E}\big\{[\eta-Y]_+ \,|\, Y \ge t\big\}. \tag{19}$$

Proof. We start from the following simple observation: if $X \succeq_{(1)} Y$ then for every $t \in \mathbb{R}$

$$\int_t^{\eta} F_X(\alpha)\, d\alpha \le \int_t^{\eta} F_Y(\alpha)\, d\alpha \quad \text{for all} \quad \eta \ge t. \tag{20}$$

It is immediately implied by (1) via integration on the interval $[t, \eta]$.

By converting the left-hand side of (20) to an iterated double integral and changing the order of integration by Fubini's theorem, we obtain

$$\begin{aligned}\int_t^{\eta} F_X(\alpha)\, d\alpha &= \int_t^{\eta}\int_{-\infty}^{\alpha} dF_X(z)\, d\alpha = \int_{(-\infty,\, t)}\int_t^{\eta} d\alpha\, dF_X(z) + \int_{[t,\, \eta]}\int_z^{\eta} d\alpha\, dF_X(z)\\ &= \int_{(-\infty,\, t)} (\eta-t)\, dF_X(z) + \int_{[t,\, \eta]} (\eta-z)\, dF_X(z)\\ &= (\eta-t)P[X < t] + P[X \ge t]\,\mathbb{E}\big\{[\eta-X]_+ \,|\, X \ge t\big\}.\end{aligned}$$

Similarly, the right-hand side of (20) can be transformed as follows:

$$\int_t^{\eta} F_Y(\alpha)\, d\alpha = (\eta-t)P[Y < t]) + P[Y \ge t]\,\mathbb{E}\big\{[\eta-Y]_+ \,|\, Y \ge t\big\}.$$

Substituting to (20) we obtain the assertion. □

Inequality (19) is particularly strong, if $P[X \ge t] = P[Y \ge t]$. In this case, we have the following corollary which means that the conditional distribution of X, given $X \ge t$ dominates in the second order the conditional distribution of Y, given $Y \ge t$.

Corollary 1. *If $X \succeq_{(1)} Y$ and $P[X \ge t] = P[Y \ge t]$, then*

$$\mathbb{E}\big\{[\eta-X]_+ \,|\, X \ge t\big\} \le \mathbb{E}\big\{[\eta-Y]_+ \,|\, Y \ge t\big\} \quad \textit{for all} \quad \eta \ge t. \tag{21}$$

Observe also that when $t \to -\infty$ inequality (21) implies the SSD condition formulated in terms of the expected shortfalls of the random variables:

$$\mathbb{E}([\eta-X]_+) \le \mathbb{E}([\eta-Y]_+) \quad \text{for all} \quad \eta \in \mathbb{R}.$$

Indeed, for an integrable random variable X we have $\lim_{t\to-\infty} t\,P[X < t] = 0$, and the SSD relation follows.

Let us now focus on problem (4). Suppose $\psi_{\mathbf{z}} \succeq_{(1)} Y$. Then it follows from Corollary 1 that for every t such that $P[\psi_{\mathbf{z}} \geq t] = P[Y \geq t]$ and with $k^* \in \{1, \ldots, D\}$ such that $y_{k^*-1} < t \leq y_{k^*}$, we have

$$\mathbb{E}\big\{[y_k - \psi_{\mathbf{z}}]_+ \mid \psi_{\mathbf{z}} \geq t\big\} \leq \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\} \quad k = k^*, \ldots, D. \tag{22}$$

We can extend this observation as follows. Suppose that we fix some t and a point $\bar{\mathbf{z}}$ such that $P[\psi_{\bar{\mathbf{z}}} \geq t] = P[Y \geq t]$. We can define the set

$$\Omega^* = \big\{1 \leq i \leq N : \psi(\bar{\mathbf{z}}, \omega_i) \geq t\big\}.$$

Finally, let the index k^* be defined as above, and let

$$v_k^* = \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\}, \quad k = k^*, \ldots, D.$$

By Proposition 1, for every feasible solution $\mathbf{z}$ of problem (4) such that

$$\big\{\omega : \psi(\mathbf{z}, \omega) \geq t\big\} = \Omega^* \tag{23}$$

the following system of linear inequalities holds true:

$$\begin{aligned} &\psi(\mathbf{z}, \omega_i) + \bar{s}_{i,k} \geq y_k, \quad i \in \Omega^*, \quad k = k^*, \ldots, D, \\ &\sum_{i \in \Omega^*} \frac{p_i}{\sum_{l \in \Omega^*} p_l} \bar{s}_{i,k} \leq v_k^*, \quad k = k^*, \ldots, D, \\ &\bar{s}_{i,k} \geq 0, \quad i \in \Omega^*, \quad k = k^*, \ldots, D. \end{aligned} \tag{24}$$

It should be stressed that system (24) has been derived under the assumption that we consider only a restricted set of solutions which includes the solutions satisfying (23). We shall use it to derive lower bounds for the optimal value of problem (4).

2.2.1 Valid inequalities based on conditional second order stochastic dominance

The inequalities derived in the previous section can be used only under additional conditions, like (23). In this subsection we derive valid convex inequalities for the entire feasible set, by employing convex approximations of the function $1/P[X \geq t]$.

Proposition 2. *If $X \succeq_{(1)} Y$ then for every $t \in \mathbb{R}$, for all $\eta \geq t$ and all $\alpha \in (0, 1)$*

$$(\eta - t)\left(\frac{-P[X \geq t] + 2\alpha - \alpha^2}{\alpha^2}\right) + \mathbb{E}\big\{[\eta - X]_+ \mid X \geq t\big\} \leq (\eta - t)\frac{P[Y < t]}{P[Y \geq t]} + \mathbb{E}\big\{[\eta - Y]_+ \mid Y \geq t\big\}. \tag{25}$$

Proof. By (19), if $X \succeq_{(1)} Y$ then for every $t \in \mathbb{R}$ and for all $\eta \geq t$

$$(\eta - t)\frac{P[X < t]}{P[X \geq t]} + \mathbb{E}\big\{[\eta - X]_+ \mid X \geq t\big\} \leq (\eta - t)\frac{P[Y < t]}{P[X \geq t]} + \frac{P[Y \geq t]}{P[X \geq t]}\,\mathbb{E}\big\{[\eta - Y]_+ \mid Y \geq t\big\}. \tag{26}$$

As the relation $X \succeq_{(1)} Y$ implies that $P[X \geq t] \geq P[Y \geq t]$, the coefficient in front of the last term in (26) can be replaced by one. Simple manipulation then yields

$$(\eta - t)\left(\frac{1}{P[X \geq t]} - 1\right) + \mathbb{E}\big\{[\eta - X]_+ \mid X \geq t\big\} \leq (\eta - t)\frac{P[Y < t]}{P[Y \geq t]} + \mathbb{E}\big\{[\eta - Y]_+ \mid Y \geq t\big\}. \quad (27)$$

Since $\left(\frac{1}{u} - 1\right)$ is a convex function of u and $\frac{-u+2\alpha-\alpha^2}{\alpha^2}$ is its tangent at α, we have

$$\left(\frac{-P[X \geq t] + 2\alpha - \alpha^2}{\alpha^2}\right) \leq \left(\frac{1}{P[X \geq t]} - 1\right),$$

for every $\alpha \in (0, 1)$. Substituting this inequality into (27) we obtain the assertion. □

Let us now consider our FSD-constrained problem with discrete distributions.

Corollary 2. *If $\psi_{\mathbf{z}} \succeq_{(1)} Y$ then for every $t \in \mathbb{R}$ and $k^* \in \{1, \dots, D\}$ such that $y_{k^*-1} < t \leq y_{k^*}$, and for all $\alpha \in (0, 1)$ we have*

$$\begin{aligned}(y_k - t)\left(\frac{-P[\psi_{\mathbf{z}} \geq t] + 2\alpha - \alpha^2}{\alpha^2}\right) &+ \mathbb{E}\big\{[y_k - \psi_{\mathbf{z}}]_+ \mid \psi_{\mathbf{z}} \geq t\big\} \\ &\leq (y_k - t)\frac{P[Y < t]}{P[Y \geq t]} + \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\}, \quad k = k^*, \dots, D. \quad (28)\end{aligned}$$

Let us introduce the following binary variables:

$$\delta_{i,t} = \begin{cases} 1 & \text{if } \psi(\mathbf{z}, \omega_i) < t, \\ 0 & \text{otherwise,} \end{cases} \qquad i = 1, \dots, N. \quad (29)$$

Let $M_{i,t} \in \mathbb{R}$ satisfy the inequalities

$$M_{i,t} \geq t - \min_{\mathbf{z} \in Z} \psi(\mathbf{z}, \omega_i), \quad i = 1, \dots, N. \quad (30)$$

Since $\psi(\cdot, \omega_i)$ is linear for all $i = 1, \dots, N$, and Z is compact, such $M_{i,t}$ values exist.

Since $\sum_{i=1}^{N} p_i \delta_{i,t} = P[\psi_{\mathbf{z}} < t]$, using (29) and (30), we can represent (28) by the following system of inequalities:

$$\begin{aligned}&(y_k - t)\left(\frac{\sum_{i=1}^{N} p_i \delta_{i,t} - 1}{\alpha^2} + \frac{2}{\alpha} - 1\right) + \mathbb{E}\big\{[y_k - \psi_{\mathbf{z}}]_+ \mid \psi_{\mathbf{z}} \geq t\big\} \\ &\quad \leq (y_k - t)\frac{P[Y < t]}{P[Y \geq t]} + \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\}, \quad k = k^*, \dots, D, \qquad (31) \\ &\psi(\mathbf{z}, \omega_i) + M_{i,t}\,\delta_{i,t} \geq t, \quad i = 1, \dots, N, \\ &\delta_{i,t} \in \{0, 1\}, \quad i = 1, \dots, N.\end{aligned}$$

Remark 1. For every decision vector $\mathbf{z}$ satisfying (28), there exist a vector $\delta \in \mathbb{R}^N$ such that the pair $(\delta, \mathbf{z})$ satisfy (29) and (31). And also if there exists a pair $(\delta, \mathbf{z})$ satisfying (29) and (31), then the decision vector $\mathbf{z}$ satisfies (28).

Remark 2. The system of inequalities (31) can be equivalently expressed by linear inequalities by introducing variables representing shortfall values, as in (24).

Suppose that $\bar{\mathbf{z}}$ is the solution of the current LP relaxation of problem (4). Let $t = \psi(\bar{\mathbf{z}}, \omega_{i^{**}})$ for some $i^{**} \in \{1, \dots, N\}$ such that $P[\psi_{\bar{\mathbf{z}}} \geq \psi(\bar{\mathbf{z}}, \omega_{i^{**}})] = P[Y \geq y_{k^*}]$ for some k^*, for which the FSD constraint (8) is violated, and let $\alpha = P[\psi_{\bar{\mathbf{z}}} \geq t]$.

In this case, by Remark 2 the system of inequalities (31) provides us with a set of valid inequalities which cut the current optimal solution. No restriction on the set of feasible solutions is involved here, contrary to the previous subsection. To describe the way such inequalities are obtained, observe that inequality (31) for the index k^* is equivalent to:

$$(y_{k^*} - t)\left(\frac{-P[\psi_{\mathbf{z}} \geq t]}{P[\psi_{\bar{\mathbf{z}}} \geq t]^2} + \frac{2}{P[\psi_{\bar{\mathbf{z}}} \geq t]} - 1\right) + \mathbb{E}\big\{[y_{k^*} - \psi_{\mathbf{z}}]_+ \mid \psi_{\mathbf{z}} \geq t\big\}$$
$$\leq (y_{k^*} - t)\frac{P[Y < t]}{P[Y \geq t]} + \mathbb{E}\big\{[y_{k^*} - Y]_+ \mid Y \geq t\big\}.$$

At the current solution $\bar{\mathbf{z}}$ we have

$$(y_{k^*}-t)\left(\frac{1}{P[\psi_{\bar{\mathbf{z}}} \geq t]} - 1\right) + \mathbb{E}\big\{[y_{k^*}-\psi_{\bar{\mathbf{z}}}]_+ \mid \psi_{\bar{\mathbf{z}}} \geq t\big\} \leq (y_{k^*}-t)\left(\frac{1}{P[Y \geq t]} - 1\right) + \mathbb{E}\big\{[y_{k^*}-Y]_+ \mid Y \geq t\big\}.$$

Since $P[\psi_{\bar{\mathbf{z}}} \geq t] = P[Y \geq t]$ and $\mathbb{E}\big\{[y_{k^*} - Y]_+ \mid Y \geq t\big\} = 0$ by the definition of t, we conclude that

$$\mathbb{E}\big\{[y_{k^*} - \psi_{\bar{\mathbf{z}}}]_+ \mid \psi_{\bar{\mathbf{z}}} \geq t\big\} = 0.$$

However, by the definition of t and k^*, respectively, $P[\psi_{\bar{\mathbf{z}}} \geq t] = P[Y \geq t]$ and the FSD constraint (8) is violated for the index k^*. Thus, there exists at least one index $i \in \{1, \dots, N\}$ such that $t \leq \psi(\bar{\mathbf{z}}, \omega_i) < y_{k^*}$, i.e., $s_{\bar{\mathbf{z}}}(i, k^*) = [y_{k^*} - \psi(\bar{\mathbf{z}}, \omega_i)]_+ > 0$. Therefore for the current solution

$$\mathbb{E}\big\{[y_{k^*} - \psi_{\bar{\mathbf{z}}}]_+ \mid \psi_{\bar{\mathbf{z}}} \geq t\big\} > 0.$$

It should be stressed that the property of separating the current point $\bar{z}$ from the feasible set is guaranteed only when the integrality restriction on the vector δ is present. If this restriction is relaxed, separation may not occur.

3 Valid inequalities based on disjunctions

3.1 Cover Inequalities

Consider the mixed-integer formulation MBLP with knapsack constraints (8). We start from some basic definitions to fix notation.

Definition 1. *A set $C \subseteq \{1, \dots, N\}$ is called a* cover *with respect to knapsack $k \in \{1, \dots, D\}$, if $\sum_{j \in C} p_j > u_k$. The cover C is* minimal *with respect to knapsack k if $\sum_{j \in C \setminus \{s\}} p_j \leq u_k$ for all $s \in C$.*

The set of all covers for knapsack k will be denoted $\mathscr{C}_k$. For any cover $C \in \mathscr{C}_k$ we have the following well-known valid inequality [3, 18, 31, 39], called the cover inequality corresponding to C:

$$\sum_{j \in C} \beta_{j,k} \leq |C| - 1. \tag{32}$$

We are going to derive from it valid inequalities for the original formulation (4).

In this section we first obtain valid simple disjunctions, namely, elementary disjunctions in nonnegative variables, involving linear inequalities by restating the inequalities arising from the underlying combinatorial structure of the problem as logical conditions. Then we derive a useful class of valid inequalities from the corresponding disjunction relaxations by applying the lift-and-project procedure.

3.2 New valid disjunctions using cover inequalities

Obviously, $\bigcup_{k=1}^{D} \mathscr{C}_k$ is the set of all minimal covers corresponding to the inequalities (8) of MBLP.

Proposition 3. *The random outcome $\psi_{\mathbf{z}}$ dominates the benchmark discrete random variable Y in the first order if and only if*

$$y_k \leq \max_{j \in C} \psi(\mathbf{z}, \omega_j) \tag{33}$$

holds for all $C \in \mathscr{C}_k$ and for all $k = 1, \dots, D$.

Proof. Suppose that there exists a cover $C \in \mathscr{C}_k$ such that $y_k > \max_{i \in C} \psi(\mathbf{z}, \omega_i)$ implying $y_k > \psi(\mathbf{z}, \omega_i)$ for all $i \in C$. Then by definition (6), we have $\beta_{i,k} = 1$ for all $i \in C$ and by the definition of a cover $\sum_{i \in C} p_i \beta_{i,k} = \sum_{i \in C} p_i > u_k$. Therefore, inequality (8) for the index k is violated and so $\psi_{\mathbf{z}}$ does not dominate Y in the first order.

In order to prove the converse, suppose that $\psi_{\mathbf{z}}$ does not dominate Y in the first order, i.e., there exist an index $k \in \{1, \dots, N\}$ such that $\sum_{i=1}^{N} p_i \beta_{i,k} > u_k$. This implies that $C = \{i \in \{1, \dots, N\} : \beta_{i,k} = 1\} \in \mathscr{C}_k$. By the construction of C, we have $y_k > \psi(\mathbf{z}, \omega_i)$ for all $i \in C$ implying $y_k > \max_{i \in C} \psi(\mathbf{z}, \omega_i)$. □

Enumeration of all covers is practically impossible. We therefore focus on covers for which the condition given in Proposition 3 is violated. They will be used to generate cutting planes.

Proposition 4. *For any cover $C \in \mathscr{C}_k$ the cover inequality*

$$\sum_{j \in C} \beta_{j,k} \leq |C| - 1, \tag{34}$$

together with (9) *and* (10)*, is equivalent to*

$$y_k \leq \max_{j \in C} \psi(\mathbf{z}, \omega_j), \tag{35}$$

together with (9) *and* (10).

Proof. Suppose that $\sum_{j \in C} \beta_{j,k} > |C| - 1$. Then by definition (6), $\beta_{j,k} = 1$ and $y_k > \psi(\mathbf{z}, \omega_j)$ for all $j \in C$. Therefore, $y_k > \max_{j \in C} \psi(\mathbf{z}, \omega_j)$.

In order to prove the converse, suppose that $y_k > \max_{j \in C} \psi(\mathbf{z}, \omega_j)$ implying $y_k > \psi(\mathbf{z}, \omega_j)$ for all $j \in C$. Then by definition (6) we have $\beta_{j,k} = 1$ for all $j \in C$. Therefore, $\sum_{j \in C} \beta_{j,k} > |C| - 1$. □

It is easy to see that when the integrality restriction (10) is relaxed, inequality (35) together with (9) is stronger than (34) together with (9). However, adding the inequalities of form (35) to the set of constraints defining the feasible region of the LP relaxation of MBLP will result in defining a nonconvex feasible region. Our objective is to find a way to use inequalities (35) while generating promising cutting planes without destroying the convexity of the feasible region. Valid inequalities can be derived by restating (35) as a logical condition:

$$y_k \leq \psi(\mathbf{z}, \omega_j) \text{ for at least one index } j \in C. \tag{36}$$

This can also be written as a disjunction:

$$\bigvee_{j \in C} (y_k \leq \psi(\mathbf{z}, \omega_j)). \tag{37}$$

3.3 Generating cutting planes from valid disjunctions

We will generate cuts that are valid inequalities for some disjunctive relaxations of the FSD-constrained problem (4) by applying the lift-and-project procedure and we obtain disjunctive relaxations by using the valid disjunctions in the form of (37). The lift-and-project procedure is a systematic way to generate an optimal (in a specific sense) disjunctive cut for a given disjunctive relaxation and it involves solving a higher dimensional cut generating linear program (CGLP). At the heart of the CGLP formulation lies the fundamental Disjunctive Cut Principle as formulated in [2, 8]. For an introduction to disjunctive programming (optimization under logical constraints involving linear inequalities) see [1, 4].

According to Proposition 3, if for the current policy $\mathbf{z}$ the random outcome $\psi_{\mathbf{z}}$ does not dominate Y in the first order, then there exists a cover $C \in \mathscr{C}_k$ such that $y_k > \max_{j \in C} \psi(\mathbf{z}, \omega_j)$. In particular, to generate a cutting plane we consider the following disjunctive relaxation of problem (4):

$$\begin{aligned} \max\ & \mathbf{c}^T \mathbf{z} \\ \text{s.t. } & \mathbf{z} \in \bar{Z}, \\ & \bigvee_{j \in C} (\psi(\mathbf{z}, \omega_j) \geq y_k), \end{aligned} \tag{38}$$

where $\bar{Z}$ is the compact polyhedron defined by the constraints of the current LP relaxation of problem (4), which is based on SSD constraints.

For convenience of notation, we assume that $\bar{Z} = \{\mathbf{z} : A\mathbf{z} \leq b,\ \mathbf{z} \geq 0\}$. It will be clear how to restate the resulting valid inequalities for more complicated definitions of the polyhedron $\bar{Z}$. We also write $\psi(\mathbf{z}, \omega_j) = d_j^T \mathbf{z}$. We can rewrite problem (38) in a more explicit form:

$$\begin{aligned} &\max \mathbf{c}^T \mathbf{z} \\ &\text{s.t } A\mathbf{z} \le b, \\ &\quad \mathbf{z} \ge 0, \\ &\quad \bigvee_{j \in C} (d_j^T \mathbf{z} \ge y_k). \end{aligned} \tag{39}$$

Let F^{FP}, F^{DP} and F^{LP} denote the feasible regions of problem (4), the disjunctive relaxation and the LP relaxation of problem (4), respectively. Then we have $F^{FP} \subset F^{DP} \subset F^{LP}$. Furthermore, an optimal solution to the LP relaxation of problem (4) will be denoted by $\mathbf{x}^{LP}$.

We are interested in the family of inequalities $\alpha \mathbf{x} \le \gamma$ implied by the constraint set of the disjunctive relaxation (39). The family of all such inequalities includes all valid cutting planes for (39). On the other hand, the set of points satisfying all members of the family is precisely the closed convex hull of the set of feasible solutions to disjunctive relaxation (39). A characterization of this family is given in the next theorem, which is a special case of the general disjunctive cut theorem, due to Balas [2] and Blair and Jeroslow [8].

Theorem 1. *The inequality $\alpha^T \mathbf{z} \le \gamma$ is a consequence of the constraints*

$$\begin{aligned} A\mathbf{z} &\le b, \\ \mathbf{z} &\ge 0, \\ \bigvee_{j \in C} &(d_j^T \mathbf{z} \ge y_k), \end{aligned}$$

if and only if there exist vectors $\theta_j \in \mathbb{R}_+^m$ and scalars $\sigma_j \ge 0$, $j \in C$, such that

$$\alpha \le A^T \theta_j - \sigma_j d_j, \ \forall j \in C, \tag{40}$$

$$\gamma \ge b^T \theta_j - \sigma_j y_k, \ \forall j \in C. \tag{41}$$

Our objective is to determine a cutting plane of the form $\alpha^T \mathbf{z} \le \gamma$ which is valid for F^{DP} but cuts off parts of F^{LP}, and in particular point $\mathbf{z}^{LP} \in F^{LP}$.

The cut, which maximizes the amount $\alpha^T \mathbf{z}^{LP} - \gamma$ by which $\mathbf{z}^{LP}$ is cut off, is derived by solving the lift-and-project CGLP:

$$\begin{aligned} &\max \alpha^T \mathbf{z}^{LP} - \gamma \\ &\text{s.t. } \alpha \le A^T \theta_j - \sigma_j d_j, \ \forall j \in C, \\ &\quad \gamma \ge b^T \theta_j - \sigma_j y_k, \ \forall j \in C, \\ &\quad \theta_j \ge 0, \ \sigma_j \ge 0, \ \forall j \in C, \\ &\quad (\alpha, \gamma) \in \mathbb{W}, \end{aligned}$$

where $\mathbb{W}$ is defined by a certain normalization constraint to truncate the feasible cone (see [7, 9, 32]). From many possible normalizations we choose the constraint:

$$\sum_{j \in C} \theta_j^T \mathbf{e} + \sigma_j = 1,$$

where $\mathbf{e}$ is the unit vector.

4 Inequalities based on Conditional Value-at-Risk

Let us define the first quantile function $F_X^{(-1)} : (0, 1) \to \mathbb{R}$ corresponding to a random variable X as the left-continuous inverse of the cumulative distribution function F_X:

$$F_X^{(-1)}(\alpha) = \inf\{\eta : F_X(\eta) \geq \alpha\}.$$

In the financial literature, the quantile $F_X^{(-1)}(\alpha)$ is called the Value at Risk (VaR) and denoted by $\mathrm{VaR}_\alpha(X)$.

By the definition of the FSD relation we have

$$X \succeq_{(1)} Y \Leftrightarrow F_X^{(-1)}(\alpha) \geq F_Y^{(-1)}(\alpha) \quad \text{for all } 0 < \alpha < 1. \tag{42}$$

Consider the second quantile function:

$$F_X^{(-2)}(\alpha) = \int_0^\alpha F_X^{(-1)}(u)\, du \quad \text{for } 0 < \alpha \leq 1,$$

For completeness, we set $F_X^{(-2)}(0) = 0$ and $F_X^{(-2)}(\alpha) = +\infty$ for $\alpha \notin [0, 1]$. The function $F_X^{(-2)}$ is convex. The second quantile function is closely related to the Conditional Value at Risk (CVaR) defined as (see [34]):

$$\mathrm{CVaR}_\alpha(X) = \sup_\eta \left\{ \eta - \frac{1}{\alpha} \mathbb{E}([\eta - X]_+) \right\}. \tag{43}$$

We always have $\mathrm{CVaR}_\alpha(X) = F_X^{(-2)}(\alpha)/\alpha$ (see [30]). It is well known (see [30, 14]) that the SSD relation is equivalent to a continuum of CVaR constraints:

$$\left[X \succeq_{(2)} Y \right] \Leftrightarrow \left[\mathrm{CVaR}_\alpha(X) \geq \mathrm{CVaR}_\alpha(Y) \quad \text{for all } 0 < \alpha < 1 \right]. \tag{44}$$

Thus the constraints at the right hand side of (44) are convex relaxations of the FSD constraint. They just represent in a different way the SSD constraint discussed earlier.

But we can also obtain a convex approximation of the set defined by the FSD constraint from within. By construction, $\mathrm{CVaR}_\alpha(X) \leq \mathrm{VaR}_\alpha(X) = F_X^{(-1)}(\alpha)$. Therefore, by (42), we can construct a convex restriction of the FSD constraint:

$$\left[\mathrm{CVaR}_\alpha(X) \geq F_Y^{(-1)}(\alpha) \text{ for all } 0 < \alpha \leq 1 \right] \Rightarrow \left[X \succeq_{(1)} Y \right]. \tag{45}$$

Consider a fixed decision vector $\mathbf{z}$ and the resulting outcome $X = \psi_{\mathbf{z}}$. By (43), in the discrete case $\text{CVaR}_\alpha(\psi_{\mathbf{z}})$ can be calculated by solving the following linear programming problem (see [34]):

$$\max \quad \eta\alpha - \sum_{i=1}^{N} p_i \nu_i \tag{46}$$

$$\text{subject to} \quad \nu_i \geq \eta - \psi(\mathbf{z}, \omega_i), \quad i = 1, \ldots, N, \tag{47}$$

$$\nu_i \geq 0, \quad i = 1, \ldots, N, \tag{48}$$

$$\eta \in \mathbb{R}. \tag{49}$$

Thus the set of vectors $\mathbf{z} \in Z$ such that

$$\text{CVaR}_\alpha(\psi(\mathbf{z})) \geq F_Y^{(-1)}(\alpha) \tag{50}$$

can be represented by the following inequalities:

$$\begin{aligned} &\eta\alpha - \sum_{i=1}^{N} p_i \nu_i \geq \alpha F_Y^{(-1)}(\alpha) \\ &\nu_i \geq \eta - \psi(\mathbf{z}, \omega_i), \quad i = 1, \ldots, N, \\ &\nu_i \geq 0, \quad i = 1, \ldots, N, \\ &\mathbf{z} \in Z, \\ &\eta \in \mathbb{R}. \end{aligned} \tag{51}$$

Such a system has to be constructed for every $\alpha \in (0, 1)$. But when the benchmark outcome Y is a discrete random variable with realizations $y_1 < y_2 < \ldots < y_D$, relation (45) is equivalent to

$$\left[\text{CVaR}_{(F_Y(y_k))}(X) \geq y_k \text{ for all } k = 1, \ldots, D\right] \Rightarrow \left[X \succeq_{(1)} Y\right]. \tag{52}$$

Then the system (51) needs to be constructed for $\alpha_k = F_Y(y_k)$, $k = 1, \ldots, D$. Of course, each of these systems involves different variables η_k and $\nu_{i,k}$.

Condition (45) (or (52)) is very conservative, especially for larger values of α. If $\alpha \to 1$, then $\text{CVaR}_\alpha(X) \to \mathbb{E}(X)$ and (45) would require that

$$\mathbb{E}(X) \geq F_Y^{(-1)}(\alpha) \quad \text{for all} \quad \alpha \in (0, 1),$$

which implies that the entire support of Y has to be below $\mathbb{E}(X)$. This is unrealistically conservative.

The only practical application of (45) is for small values of α in a heuristic procedure for constructing feasible solutions.

5 Heuristic algorithms for constructing feasible solutions

Optimization problems with FSD constraints are too difficult for the standard MIP solvers such as CPLEX. In practice, it is important to construct feasible solutions which provide good lower bounds on the optimal objective value.

Let us represent the SSD-based LP relaxation solved at the ℓth stage of the solution process as follows:

$$
\begin{aligned}
\max \quad & \mathbf{c}^T\mathbf{z} && (53)\\
\text{subject to} \quad & \psi(\mathbf{z},\omega_i)+\bar{s}_{i,k}\geq y_k, \quad i=1,\dots,N, \quad k=1,\dots,D, && (54)\\
& \sum_{i=1}^{N} p_i\bar{s}_{i,k}\leq v_k, \quad k=1,\dots,D, && (55)\\
& \bar{s}_{i,k}\geq 0, \quad i=1,\dots,N, \quad k=1,\dots,D, && (56)\\
& (\mathbf{z},\bar{\mathbf{s}})\in S^\ell. && (57)
\end{aligned}
$$

Here (54)–(56) are the SSD constraints, and (57) represents all deterministic constraints, additional cuts and restrictions added during the solution process. In the simplest case of the pure SSD-based LP relaxation, at the first stage (for $\ell = 1$) we have $S^1 = Z \times \mathbb{R}^{ND}$. If additional disjunctive cuts are employed, the definition of S^1 involves these cuts as well.

Denote by $(\mathbf{z}^\ell, \bar{\mathbf{s}}^\ell)$ the solution of problem (53)–(57) at the ℓth stage of the solution process. Let $\mathbf{s}^\ell = \mathbf{s}_{\mathbf{z}^\ell}$ be defined as in (12). If $\mathbf{z}^\ell$ satisfies the FSD constraint,

$$
P[\psi_{\mathbf{z}^\ell} < y_k] \leq P[Y \leq y_{k-1}], \quad k=1,\dots,D, \tag{58}
$$

a feasible solution has been found. Otherwise, we restrict the set S^ℓ. We consider three techniques for restricting this set.

Heuristic 1. ***Fixing***

The fixing method is very similar to diving techniques employed in general mixed-integer programming.

Step 1. Solve (53)–(57). If the FSD constraint (58) is satisfied or if the problem is infeasible, then stop. Otherwise continue.

Step 2. Find the smallest index k^* for which the constraint (58) is violated.

Step 3. Define the set: $I^* = \{1 \leq i \leq N : s^\ell_{i,k^*} > 0\}$.

Step 4. Select the index $i^* \in I^*$ such that $s^\ell_{i^*,k^*} = \min_{i\in I^*} s^\ell_{i,k^*}$.

Step 5. Define the new set $S^{\ell+1} = S^\ell \cap \big\{(\mathbf{z}, \bar{\mathbf{s}}) \in Z \times \mathbb{R}^{ND} : \bar{s}_{i^*,k^*} = 0\big\}$.

Step 6. Increase ℓ by one and go to Step 1.

As the number of the shortfall variables $\bar{s}_{i,k}$ is finite, this procedure is finite. It either ends at some feasible point $\mathbf{z}^\ell$ or discovers an empty feasible set in (54)–(57).

Heuristic 2. ***Conditional second order stochastic dominance constraints***

This method is based on Corollary 1.

The method attempts to guess the value of $t = \psi(\mathbf{z}^\ell, \omega_{i^{**}})$ for some $i^{**} \in \{1,\dots,N\}$ such that $P[\psi_{\mathbf{z}^\ell} \geq \psi(\mathbf{z}^\ell, \omega_{i^{**}})] = P[Y \geq y_{k^*}]$ for some k^*, and to guess the set of events $\Omega^* = \big\{1 \leq i \leq N :$

$\psi(\mathbf{z}^\ell, \omega_i) \geq \psi(\mathbf{z}^\ell, \omega_{i^{**}})\big\}$. It is most convenient to describe it for the case of equally likely realizations, that is for $p_i = 1/N$, $i = 1, \ldots, N$.

Step 1. Solve (53)–(57). If the constraint (58) is satisfied or if the problem is infeasible, then stop. Otherwise continue.

Step 2. Find the smallest index k^* for which the constraint (58) is violated.

Step 3. Define the set: $I^* = \big\{1 \leq i \leq N : s^\ell_{i,k^*} > 0\big\}$.

Step 4. Select the index $i^{**} \in I^*$ such that $P\big[\psi^\ell_{\mathbf{z}} \geq \psi(\mathbf{z}^\ell, \omega_{i^{**}})\big] = P\big[Y \geq y_{k^*}\big]$. The point $\psi(\mathbf{z}^\ell, \omega_{i^{**}})$ will play the role of t in (22).

Step 5. Define the set $\Omega^* = \big\{1 \leq i \leq N : \psi(\mathbf{z}^\ell, \omega_i) \geq \psi(\mathbf{z}^\ell, \omega_{i^{**}})\big\}$, and let $N^* = |\Omega^*|$. Calculate the re-scaled probabilities corresponding to the realizations of the benchmark outcome Y:

$$p^\ell_j = \frac{p_j}{\sum_{s=k^*}^{D} p_s}, \quad j = k^*, \ldots, D,$$

and calculate the conditional shortfall values:

$$v^*_k = \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\} = \sum_{j=k^*}^{k} p^\ell_j (y_k - y_j), \quad k = k^*, \ldots, D.$$

Step 6. Define the new set $S^{\ell+1}$ by augmenting the definition of the set S^ℓ with the second order constraints for the conditional distributions:

$$\begin{aligned} &\psi(\mathbf{z}, \omega_i) + \bar{s}_{i,k} \geq y_k, \quad i \in \Omega^*, \quad k = k^*, \ldots, D, \\ &\frac{1}{N^*} \sum_{i \in \Omega^*} \bar{s}_{i,k} \leq v^*_k, \quad k = k^*, \ldots, D, \\ &\bar{s}_{i,k} \geq 0, \quad i \in \Omega^*, \quad k = k^*, \ldots, D. \end{aligned} \tag{59}$$

Step 7. Increase ℓ by one and go to Step 1.

Proposition 5. *After finitely many iterations the algorithm either finds a feasible solution, or encounters infeasibility.*

Proof. If the algorithm does not stop at iteration ℓ, it is easy to see that Steps 2-5 can be executed. It remains to analyze Step 6.

We claim that

$$y_{k^*-1} < \psi(\mathbf{z}^\ell, \omega_{i^{**}}) < y_{k^*}. \tag{60}$$

By the definition of i^{**} we have $s^\ell_{i^{**},k^*} > 0$ implying $\psi(\mathbf{z}^\ell, \omega_{i^{**}}) < y_{k^*}$. Suppose that $\psi(\mathbf{z}^\ell, \omega_{i^{**}}) \leq y_{k^*-1}$, then by the definition of i^{**} and the fact that $y_{k^*-2} < y_{k^*-1}$ we get $P\big[\psi^\ell_{\mathbf{z}} < \psi(\mathbf{z}^\ell, \omega_{i^{**}})\big] > P\big[Y \leq$

$y_{k^*-2}]$. This implies that the constraint (58) is violated at least for index $k^* - 1$, and therefore we have a contradiction with the definition of k^*.

By the definition of i^{**} and (60) , we have

$$P\big[\psi_{\mathbf{z}}^{\ell} \geq \psi(\mathbf{z}^{\ell}, \omega_{i^{**}})\big] = P\big[Y \geq \psi(\mathbf{z}^{\ell}, \omega_{i^{**}})\big]. \tag{61}$$

Since the conditions (60) and (61) hold true, $\psi(\mathbf{z}^{\ell}, \omega_{i^{**}})$ plays the role of t in (22) and we obtain

$$\mathbb{E}\big\{[y_k - \psi_{\mathbf{z}}]_+ \mid \Omega^*\big\} \leq \mathbb{E}\big\{[y_k - Y]_+ \mid Y \geq y_{k^*}\big\} \quad k = k^*, \dots, D. \tag{62}$$

In the discrete case with equally likely realizations, condition (62) takes the form of the following system of linear inequalities:

$$\begin{aligned}
&\psi(\mathbf{z}, \omega_i) + \bar{s}_{i,k} \geq y_k, \quad i \in \Omega^*, \quad k = k^*, \dots, D,\\
&\frac{1}{N^*} \sum_{i \in \Omega^*} \bar{s}_{i,k} \leq v_k^*, \quad k = k^*, \dots, D,\\
&\bar{s}_{i,k} \geq 0, \quad i \in \Omega^*, \quad k = k^*, \dots, D.
\end{aligned}$$

By the argument above the algorithm is well defined. As the imposition of the conditional second order dominance constraints (59) involves the requirement that

$$\bar{s}_{i,k^*} = 0, \quad i \in \Omega^*, \tag{63}$$

the system of inequalities (59) cuts the current optimal solution and the procedure is finite. Therefore, the algorithm stops at Step 1 at some iteration ℓ^* with the solution $\mathbf{z}^{\ell^*}$ for which random outcome $\psi_{\mathbf{z}^{\ell^*}}$ dominates benchmark outcome Y in the first order. □

The method appears to be even more greedy than the fixing method, because it fixes a group of shortfall variables, not just one of them. However, the conditional second order constraint (59) is also a set of additional valid inequalities involving the outcomes for $i \in \Omega^*$, given that the fixing occurred. This significantly tightens the feasible set of the next problem (53)–(57), without cutting any feasible solutions, in addition to the solutions possibly cut off by the constraints (63). Thus, the number of iterations required for the Heuristic 2 algorithm to find a feasible solution is less or equal to the number of iterations required for the Heuristic 1 algorithm. According to the computational results, the difference in the number of iterations appears significant.

Heuristic 3. ***Progressive Conditional Value-at-Risk***

This method is based on the implication (45).

The method attempts to find the smallest $\alpha \in (0, 1)$ for which (50) is violated and to augment the definition of the set S^{ℓ} with the set of constraints given in (51).

Step 1. Solve (53)–(57). If the constraint (58) is satisfied or if the problem is infeasible, then stop. Otherwise continue.

Step 2. Find the smallest index k^* for which the constraint (58) is violated.

Step 3. Define $\alpha_\ell = F_Y(y_{k^*})$.

Define the new set S^ℓ by augmenting the definition of the set S^ℓ with the following constraints for the Conditional Value-at-Risk at level α_ℓ (50):

$$
\begin{aligned}
&\eta_\ell \alpha_\ell - \sum_{i=1}^{N} p_i \nu_{i,\ell} \geq \alpha_\ell F_Y^{(-1)}(\alpha_\ell) \\
&\nu_{i,\ell} \geq \eta_\ell - \psi(\mathbf{z}, \omega_i), \quad i = 1, \dots, N, \\
&\nu_{i,\ell} \geq 0, \quad i = 1, \dots, N, \\
&\eta \in \mathbb{R}.
\end{aligned}
\tag{64}
$$

Step 4. Increase ℓ by one and go to Step 1.

We verify that CVaR constraint (50) always cuts the current solution for $\alpha_\ell = F_Y(y_{k^*})$. Since the constraint (58) is violated for the index k^*, by the use of (42)

$$
\mathrm{VaR}_{\alpha_\ell}(\psi_{\mathbf{z}^\ell}) = F_{\psi_{\mathbf{z}^\ell}}^{(-1)}(\alpha_\ell) < F_Y^{(-1)}(\alpha_\ell).
$$

We also know that $\mathrm{CVaR}_{\alpha_\ell}(\psi_{\mathbf{z}^\ell}) \leq \mathrm{VaR}_{\alpha_\ell}(\psi_{\mathbf{z}^\ell})$. Therefore $\mathrm{CVaR}_{\alpha_\ell}(\psi_{\mathbf{z}^\ell}) < F_Y^{(-1)}(\alpha_\ell)$ for $\alpha_\ell = F_Y(y_{k^*})$, which implies that adding the constraints (64) will cut off the current optimal solution.

The CVaR constraints are very conservative. The number of iterations required for Heuristic 3 algorithm to find a feasible solution or to obtain an empty feasible set is usually very small.

6 Computational Results

To illustrate the computational efficiency of the SSD-based LP relaxation, the disjunctive cuts and the heuristic algorithms for generating feasible solutions, we consider a portfolio optimization problem.

6.1 Portfolio optimization problem

The problem of optimizing a portfolio of finitely many assets is a classical problem in the theoretical and computational finance. The practice of portfolio optimization under risk uses mean–risk models [25, 26, 29, 30].

Another risk-averse approach to the portfolio selection problem is that of stochastic dominance. In [10, 11, 14] the dominance-constrained portfolio optimization problem involves the SSD relation.

Here we consider the portfolio optimization model under the condition that the portfolio return stochastically dominates a reference return, for example, the return of an index, in the first order. As the first order relation carries over to expectations of monotone utility functions, no rational decision maker will prefer a portfolio with return Y over a portfolio with return X which dominates Y in the first order.

Let $R_1, R_2, \ldots, R_n$ be random returns of assets $1, 2, \ldots, n$. We assume that $\mathbb{E}[|R_j|] < \infty$ for all $j = 1, \ldots, n$. We denote the fractions of the initial capital invested in assets $j = 1, \ldots, n$ by $z_1, \ldots, z_n$ and then the total return is $R(\mathbf{z}) = R_1 z_1 + \ldots + R_n z_n$. Clearly, the set of possible asset allocations is:

$$Z = \{\mathbf{z} \in \mathbb{R}^n : z_1 + \cdots + z_n = 1,\ z_j \geq 0,\ j = 1, \ldots, n\}.$$

We consider several problem instances of different sizes, obtained from historical data on realizations of joint daily returns of 719 assets in four years (1999, 2000, 2001 and 2002) ([37]). We use the returns in each day as equally probable realizations.

The benchmark outcome was constructed as the return of a certain index of the selected assets. For the purpose of comparison we have selected equally weighted indexes composed of the q assets having the highest average return in this period. Let us denote the set of those q assets by Q. Thus we have chosen the benchmark random return Y as the return of the equally weighted portfolio $\bar{\mathbf{z}}$, where $\bar{z}_i = 1/q$ if $i \in Q$, 0 otherwise, and the return realizations are

$$\hat{y}_i = \frac{1}{q} \sum_{j \in Q} r_{j,i}, \quad i = 1, \ldots, N,$$

where $r_{j,i}$ denotes the return of asset j in day i.

We solved all the problem instances on a 2.00 GHz Pentium 4 PC with 1.00 GB of RAM, by using the AMPL modeling language [16] and the version 9.1 of the CPLEX solver [21].

6.2 The SSD-based LP relaxation versus the usual LP relaxation

Let $\mathbf{z}^*$ be the best available feasible solution. To compare the performance of the usual and the SSD-based LP relaxations, we calculate the relative optimality gap as follows:

$$\text{Relative optimality gap (Gap)} = \frac{\mathbf{c}^T \mathbf{z}^{LP} - \mathbf{c}^T \mathbf{z}^*}{\mathbf{c}^T \mathbf{z}^*}.$$

The best feasible solutions $\mathbf{z}^*$ are obtained by the heuristic algorithms described in Section 5. The SSD-based LP relaxation was solved by the specialized cutting plane approach of [36].

The results for the problem from the second row of Table 1 are illustrated graphically in Figures 1 through 4.

Figure 1 shows the distribution functions of the benchmark random outcome and the random outcomes obtained by using the usual LP and the SSD-based LP relaxations. Figure 3 shows the distribution functions of the random outcomes obtained by using the best available feasible solution, the usual LP and the SSD-based LP relaxations. Figures 2 and 4 show the magnified view of a portion of the left tail of Figures 1 and 3, respectively.

6.3 Computational testing of the disjunctive cuts

We compare the performance of CPLEX with and without adding the disjunctive cuts described in Section 3.3 to the MBLP formulation of the instances of the portfolio optimization problem. We have implemented

Number of Scenarios	Number of Continuous Variables	Number of 0-1 Variables	Gap Usual LP Relaxation	Gap SSD-based LP Relaxation	Reduction in Relative Optimality Gap
52	719	2704	5.95%	1.61%	72.93%
82	719	6724	13.67%	0.18%	98.69%
104	719	10816	12.93%	0.54%	95.79%
248	719	61504	?	0.65%	
82	719	6724	35.41%	0.50%	98.59%
104	719	10816	31.11%	0.40%	98.73%
252	719	63504	?	0.47%	
82	719	6724	9.28%	0.44%	95.28%
104	719	10816	13.67%	0.07%	99.47%
252	719	63504	?	0.32%	
82	300	6724	16.21%	0.62%	96.15%
104	300	10816	14.90%	1.00%	93.30%
248	300	61504	26.10%	0.32%	98.78%
82	300	6724	38.53%	1.81%	95.29%
104	300	10816	30.87%	1.53%	95.04%
252	300	63504	16.82%	0.78%	95.36%
82	300	6724	10.68%	0.44%	95.88%
104	300	10816	14.25%	0.20%	98.60%
252	300	63504	13.65%	0.37%	97.25%
52	719	2704	5.33%	0.49%	90.79%
82	300	6724	7.46%	0.42%	94.43%
104	300	10816	11.86%	5.81%	50.98%
248	300	61504	?	0.17%	

Table 1: Reduction in Optimality Gap by the SSD-based LP Relaxation. (? : “unrecoverable failure: insufficient memory”.)

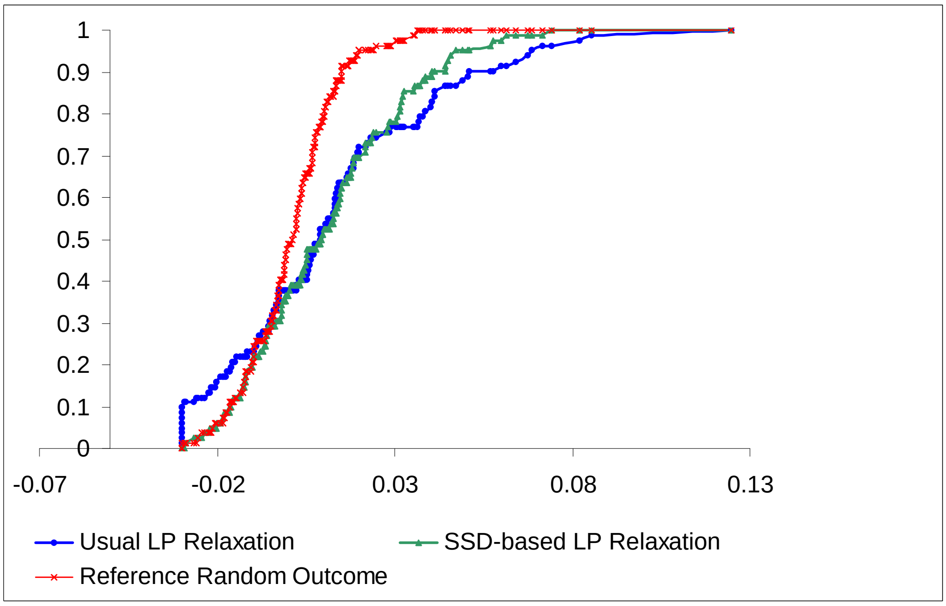

Figure 1: Distribution Functions: The benchmark random outcome and the random outcomes obtained by the usual LP and the SSD-based LP relaxations.

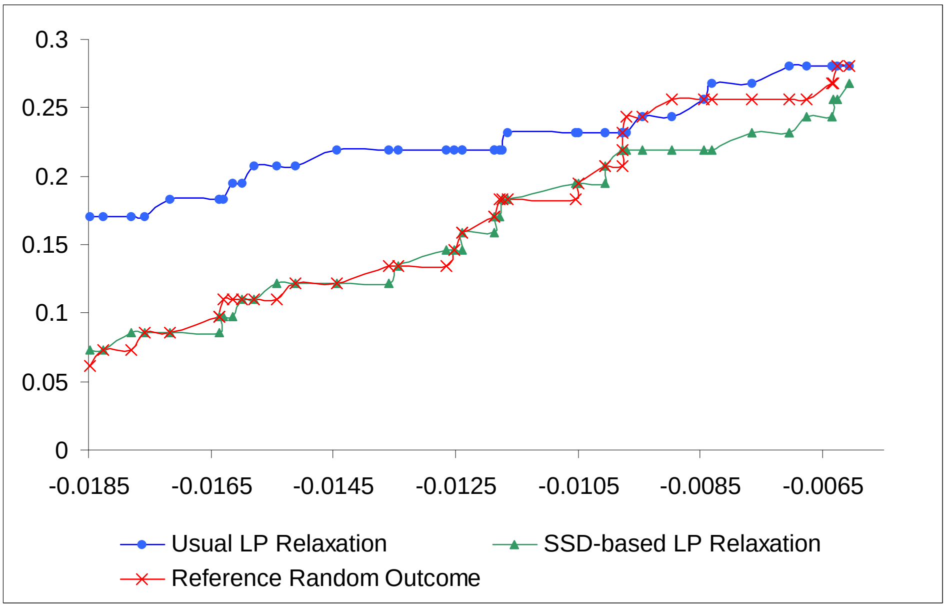

Figure 2: Magnified view of a portion of the left tail of Figure 1.

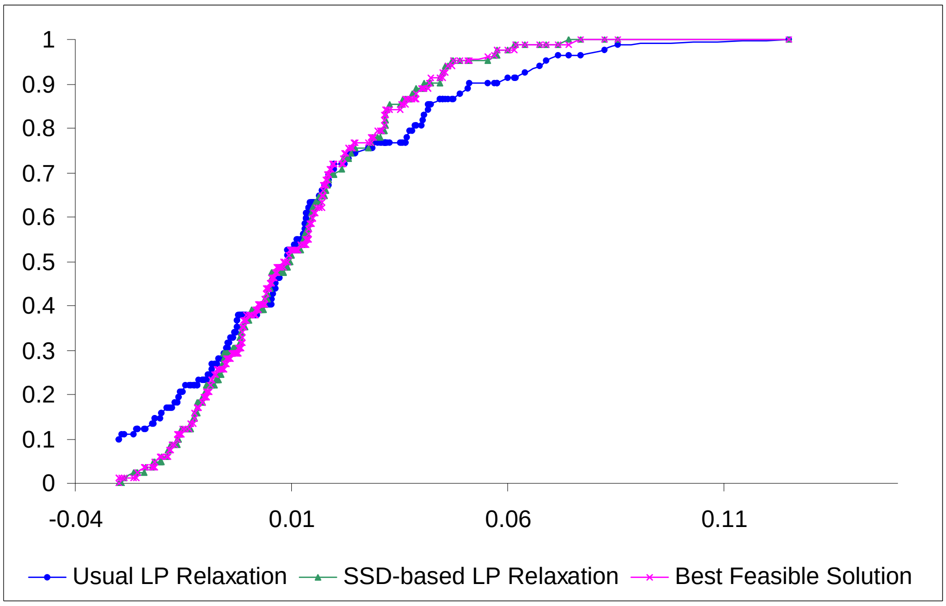

Figure 3: Distribution Functions: Best feasible solution, the usual LP and the SSD-based LP relaxations.

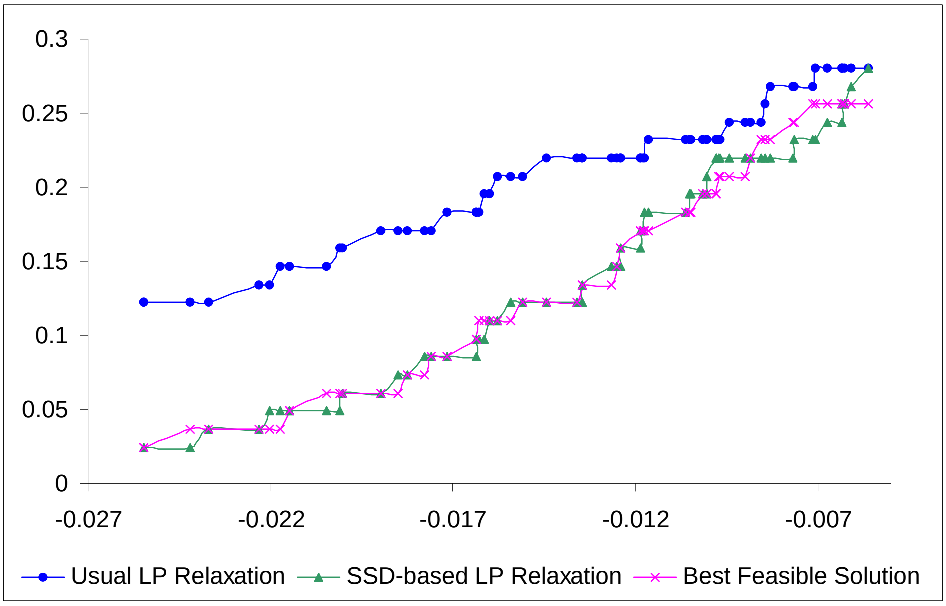

Figure 4: Magnified view of a portion of the left tail of Figure 3.

Number of Scenarios	Number of Continuous Variables	Number of 0-1 Variables	Relative Optimality Gap CPLEX (Nodes=5000)	Relative Optimality Gap CPLEX + the Disjunctive Cuts (Nodes=5000)	Reduction in Relative Optimality Gap
52	719	2704	1.177%	0.702%	40.357%
82	719	6724	5.717%	0.049%	99.137%
82*	719	6724	19.249%	7.514%	60.965%
82	719	6724	4.379%	0.908%	79.268%
82	300	6724	7.356%	0.739%	89.952%
82	300	6724	100.000%	13.019%	86.981%
82	300	6724	8.704%	2.265%	73.979%
104	300	10816	11.480%	2.914%	74.619%
104**	719	10816	10.006%	2.985%	70.171%
52	300	2704	2.184%	0.241%	88.952%
82	300	6724	3.619%	2.488%	31.266%

Table 2: Reduction in Optimality Gap by the Disjunctive Cuts for CPLEX. (* , ** : Only 30 and 50 disjunctive cuts are added, respectively.)

the lift-and-project procedure in the modeling language AMPL and CPLEX 9.1. was used as the LP solver. CPLEX was terminated after 5000 nodes were processed by the branch-and-bound algorithm.

Let $\mathbf{z}_t^*$ be the best feasible solution found and $\mathbf{z}_t^{LP}$ be the solution of the usual LP relaxation with the lowest optimal objective function value after t nodes are processed.

We calculate the relative optimality gap as follows:

$$\text{Relative optimality gap} = \frac{\mathbf{c}^T\mathbf{z}_t^{LP} - \mathbf{c}^T\mathbf{z}_t^*}{\mathbf{c}^T\mathbf{z}_t^{LP}}.$$

When no integer solution is found, the optimal objective function value of the best solution is taken as 0, and so the relative optimality gap is equal to 100%.

First, we solved the SSD-based LP relaxation. For the violated FSD constraints with two smallest indices, we generated the corresponding disjunctive cuts. We added these cuts to the current LP relaxation and continued to solve the LP relaxation and to generate disjunctive cuts. In this way, we generated 100 disjunctive cuts and added them to the formulation of MBLP. Table 2 shows the relative optimality gap values obtained with and without the disjunctive cuts and also the reduction in relative optimality gap by the use of such cuts.

Adding the disjunctive cuts to the formulation of MBLP tightens the feasible region, which results in reducing the optimality gap, as seen in Table 2. The reduction in optimality gap is due to generally improving both the upper bound and lower bound on the optimal objective function value. (For the last problem instance only the upper bound was improved.)

To illustrate the progress of different approaches, we consider an instance of the portfolio optimization problem, from the 11th row of Table 1, with the dimensions given in Table 4. For this specific instance of our problem we solved the SSD-based and the usual LP relaxations with and without different number of disjunctive cuts.

Number of Cuts	Gap Usual LP Relaxation Without the Disjunctive Cuts	Gap Usual LP Relaxation With the Disjunctive Cuts	Usual LP Relaxation Reduction in Gap	Gap SSD-based LP Relaxation Without the Disjunctive Cuts	Gap SSD-based LP Relaxation With the Disjunctive Cuts	SSD-based LP Relaxation Reduction in Gap
20	16.207%	1.156%	92.865%	0.624%	0.427%	31.554%
30	16.207%	1.012%	93.757%	0.624%	0.411%	34.142%
50	16.207%	0.469%	97.109%	0.624%	0.390%	37.476%
100	16.207%	0.385%	97.623%	0.624%	0.367%	41.171%
150	16.207%	0.369%	97.724%	0.624%	0.360%	42.334%
200	16.207%	0.368%	97.726%	0.624%	0.359%	42.518%

Table 3: Reduction in Optimality Gap by the Disjunctive Cuts for the LP Relaxations.

Here we calculate the relative optimality gap as follows:

$$\text{Relative optimality gap (Gap)} = \frac{\mathbf{c}^T\mathbf{z}^{LP} - \mathbf{c}^T\mathbf{z}_t^*}{\mathbf{c}^T\mathbf{z}_t^*}.$$

Number of Scenarios (N)	82
Number of Stocks (n)	300
Number of binary variables	6724
Number of continuous variables	300
Number of linear constraints	6807

Table 4: Dimensions of the Problem.

Table 3 demonstrates the effectiveness of the disjunctive cuts in strengthening the LP relaxations of problem (4) and reducing the optimality gap.

Figure 6 illustrates that the disjunctive cuts accelerate the convergence to a better relative optimality gap, in addition to improving the lower and upper bounds on optimal objective function value.

Also adding the disjunctive cuts to the SSD-based LP relaxation may help the heuristic algorithms to provide us with better feasible solutions.

The computational results presented in this section show the effectiveness of the disjunctive cutting planes in improving the upper and lower bounds on the optimal objective function value, i.e., in strengthening the LP relaxations of problem (4) and constructing better feasible solutions.

6.4 Computational testing of the heuristic algorithms

The problems listed in Table 1 are too difficult for the standard MIP solver of CPLEX.

Computationally, the importance of constructing feasible solutions is to obtain a good lower bound on the optimal objective function value. If a standard MIP solver is provided with a lower bound, the

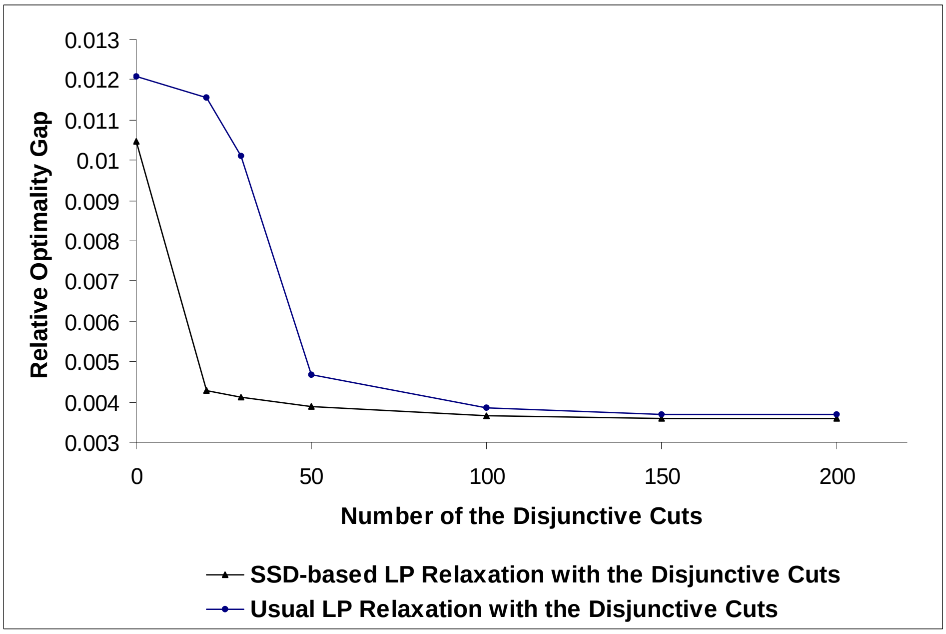

Figure 5: Reduction in Optimality Gap by Different Number of Disjunctive Cuts.

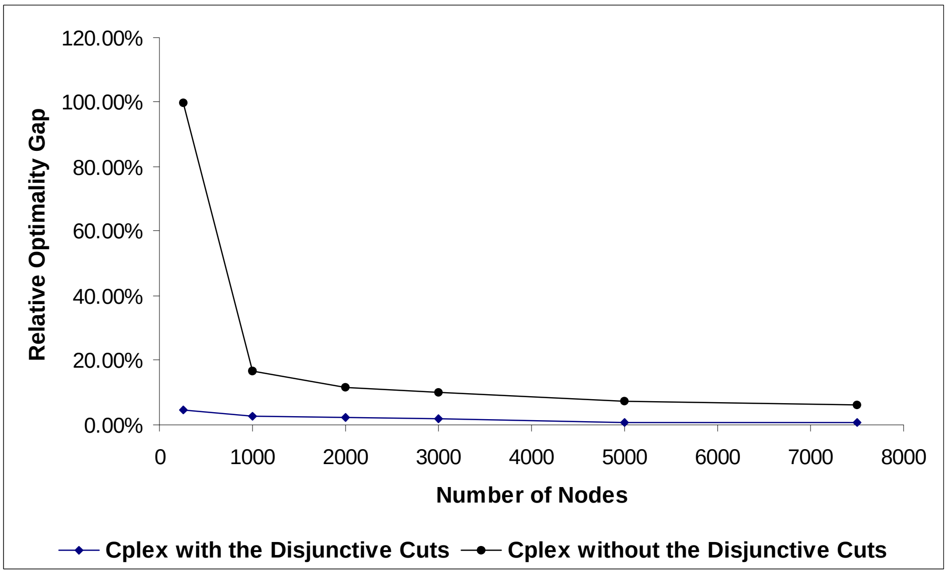

Figure 6: Reduction in Optimality Gap by the Disjunctive Cuts for Different Number of Nodes.

Number of Scenarios	Number of Continuous Variables	Number of 0-1 Variables Mixed 0-1 LP Problem (MBLP)	Relative Optimality Gap Heuristic 1. Fixing	Relative Optimality Gap Heuristic 2. Conditional SSD	Relative Optimality Gap Heuristic 3. Progressive CVaR
52	719	2704	1.585%	1.641%	12.877%
82	719	6724	0.432%	0.179%	10.418%
104	719	10816	1.158%	0.541%	Infeasible
248	719	61504	1.047%	0.644%	13.987%
82	719	6724	0.548%	0.498%	Infeasible
104	719	10816	0.606%	0.393%	Infeasible
252	719	63504	0.490%	0.464%	Infeasible
82	719	6724	0.741%	0.436%	11.294%
104	719	10816	0.104%	0.072%	12.220%
252	719	63504	0.398%	0.320%	23.071%
82	300	6724	0.633%	0.620%	Infeasible
104	300	10816	1.618%	0.988%	Infeasible
248	300	61504	0.462%	0.317%	15.395%
82	300	6724	2.926%	1.781%	Infeasible
104	300	10816	2.093%	1.508%	Infeasible
252	300	63504	0.774%	0.797%	Infeasible
82	300	6724	1.730%	0.438%	Infeasible
104	300	10816	0.244%	0.199%	Infeasible
252	300	63504	0.578%	0.373%	24.769%
52	719	2704	0.680%	0.488%	Infeasible
82	300	6724	0.414%	1.086%	Infeasible
104	300	10816	5.620%	5.495%	Infeasible
248	300	61504	0.171%	0.180%	Infeasible

Table 5: Relative Optimality Gap for the Heuristics.

noncritical nodes can be cut off without being processed during the branch-and-bound method, and therefore finding a lower bound can be very effective in improving the efficiency of the MIP solvers.

Let $\mathbf{z}^*$ be the best feasible solution obtained by the described heuristic algorithms. For the computational performance of the heuristic algorithms we define the relative optimality gap as follows:

$$\text{Relative optimality gap} = \frac{\mathbf{c}^T\mathbf{z}^{\text{SSD}} - \mathbf{c}^T\mathbf{z}^*}{\mathbf{c}^T\mathbf{z}^{\text{SSD}}}.$$

Table 5 provides information about the performance of different heuristic algorithms for generating feasible solutions. Both diving and conditional SSD constraints are quire efficient, with slightly higher accuracy of the second one. The CVaR constraints, as expected, are too conservative.

The example, whose dimensions given in Table 4, was also solved (in the mixed 0–1 formulation (MBLP)) by CPLEX. The reduction of the optimality gap for different number of nodes in the branch and bound tree is illustrated in Table 6. This is compared in quality of the bounds and in time to Heuristic 2 (conditional SSD constraints). A dramatic difference in favor of SSD-based cuts can be observed. The progress of the branch and bound method is graphically illustrated in Figure 7.

Number of NODES	Relative Optimality Gap CPLEX	Relative Optimality Gap Heuristic 2.	Reduction in Relative Optimality Gap	Heuristic 2. CPU (sec)	CPLEX CPU (sec)
250	100.000%	0.620%	99.380%	334	723
500	23.012%	0.620%	97.306%	334	981
750	17.422%	0.620%	96.442%	334	1250
1000	16.868%	0.620%	96.325%	334	1553
2000	11.709%	0.620%	94.706%	334	2210
3000	10.278%	0.620%	93.969%	334	2715
5000	7.356%	0.620%	91.572%	334	3930
7500	6.227%	0.620%	90.045%	334	5105
10000	5.645%	0.620%	89.019%	334	6289
50000	4.417%	0.620%	85.964%	334	28491

Table 6: Relative Optimality Gap for Different Number of Nodes.

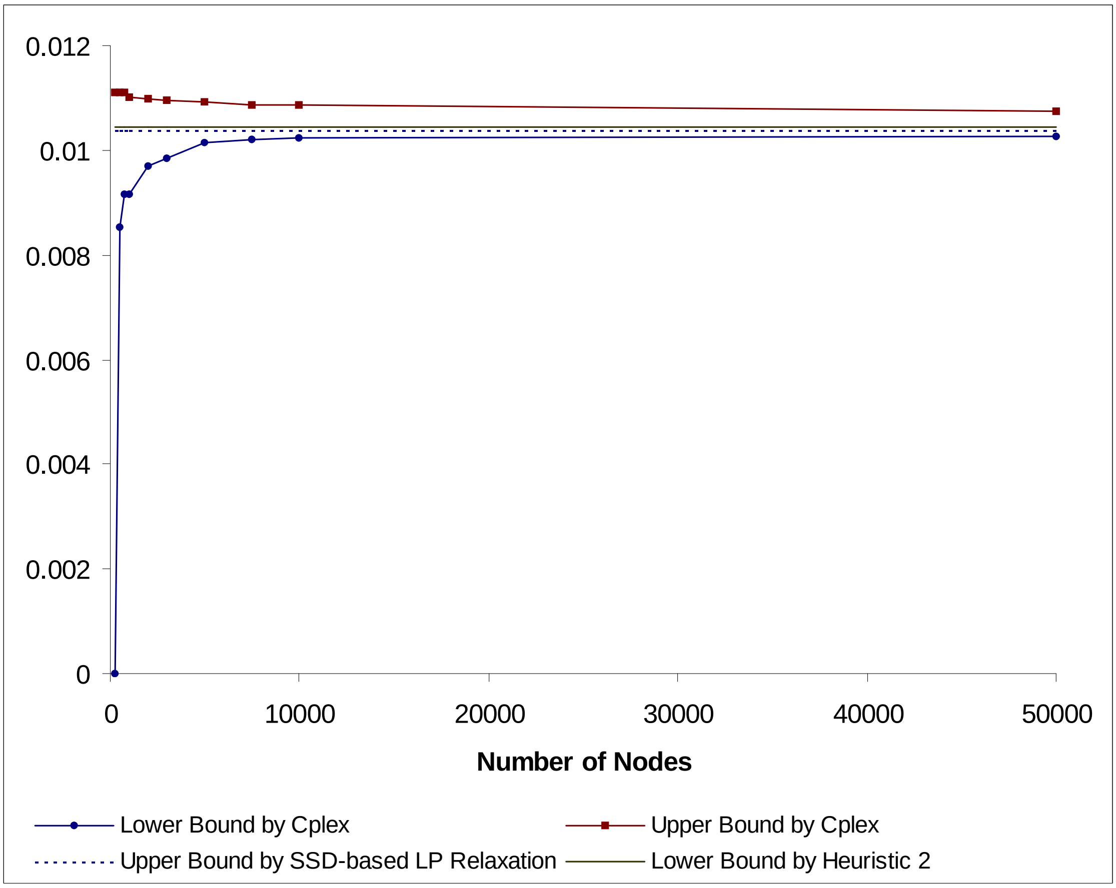

Figure 7: Upper and Lower Bounds on the Optimal Objective Function Value.

Number of Continuous Variables	Number of 0-1 Variables	Number of Variables	Heuristic 1. CPU (sec)	Heuristic 2. CPU (sec)	CPLEX (Nodes=5000) CPU (sec)	Relative Opt. Gap CPLEX	Relative Opt. Gap Heuristic 2.
719	2704	3423	154	448	4983	1.177%	1.641%
300	6724	7024	234	333	3930	7.356%	0.620%
719	6724	7443	392	448	8690	5.717%	0.179%
300	10816	11116	529	480	6554	11.480%	0.988%
719	10816	11535	1053	1194	16282	10.006%	0.541%
300	61504	61804	2636	3137	?	?	0.317%
719	61504	62223	7430	7289	?	?	0.644%
300	63504	63804	8117	7810	?	?	0.373%
719	63504	64223	8217	9442	?	?	0.464%

Table 7: CPU and the Relative Optimality Gap Results for the Heuristics and CPLEX after 5000 Nodes. (? : "unrecoverable failure with no integer solution".)

The relative optimality gap for CPLEX was calculated as follows:

$$\text{Relative optimality gap} = \frac{\mathbf{c}^T\mathbf{z}_t^{LP} - \mathbf{c}^T\mathbf{z}_t^*}{\mathbf{c}^T\mathbf{z}_t^{LP}}.$$

For Heuristic 2, the gap was calculated as:

$$\text{Relative optimality gap} = \frac{\mathbf{c}^T\mathbf{z}^{\text{SSD}} - \mathbf{c}^T\mathbf{z}^*}{\mathbf{c}^T\mathbf{z}^{\text{SSD}}}.$$

As can be seen from Table 6, even after 50000 nodes were processed by CPLEX solver (in 28491 sec.) the relative optimality gap is still 4.4%. On the other hand, Heuristic 2 provides us with the relative optimality gap of 0.62% and 85.96% reduction in relative optimality gap in just 334 sec.

Table 7 shows that for CPLEX MIP solver it takes substantially more computational time to find "good enough" feasible solutions. Whereas, the heuristic algorithms find "better" feasible solutions in less computational time than CPLEX.

As seen from the computational results presented in this section, Heuristic 1 and Heuristic 2 turn out to be computationally practical and significantly effective in finding feasible solutions of hard problem (4). Heuristic 3 was disregarded because of its poor performance and frequent infeasibility.

7 Conclusion

A linear stochastic optimization problems with first order stochastic dominance constraints for discrete random variables can be formulated as a mixed 0–1 programming problems. Our contribution to the solution of such problems is twofold. First, we decrease the upper bound on the optimal objective value by considering the SSD-based LP relaxation instead of the usual LP relaxation, and by adding new cutting planes arising from combinatorial and probabilistic structures of the problem. Secondly, we increase the lower bound by constructing better feasible solutions using special heuristic procedures, which are also derived from the probabilistic structure of the problem. In this way we substantially reduce the gap between the upper and lower bounds, as compared to standard mixed integer programming techniques.

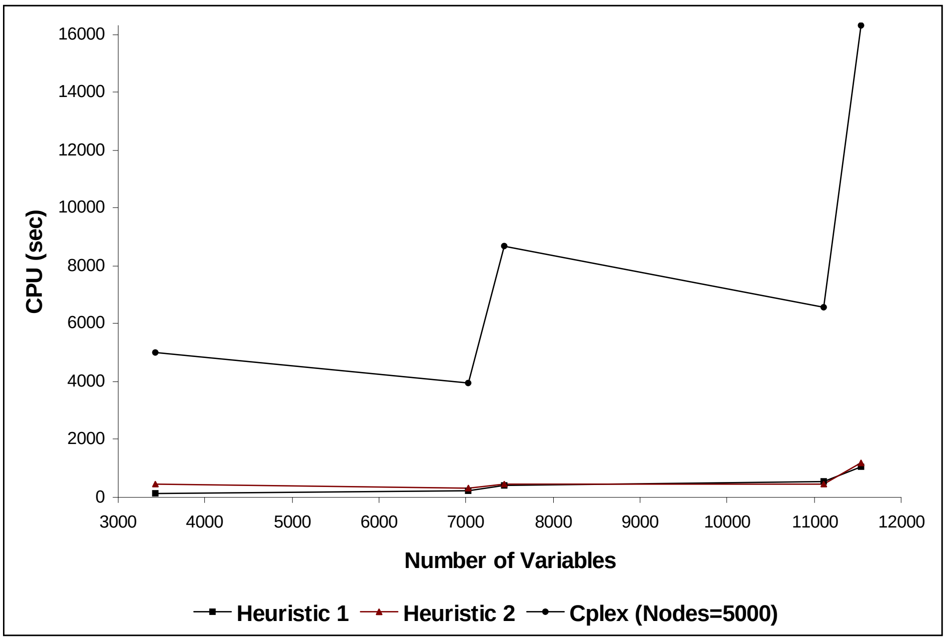

Figure 8: CPU Results for Problem Instances of Different Sizes.

References

[1] Balas, E., 1974, Disjunctive programming: Properties of the convex hull of feasible points, technical report MSRR 348, Carnegie Mellon University.

[2] Balas, E., 1975, Disjunctive programming: cutting planes from logical conditions, in: O. L. Mangasarian, R. R. Meyer and S. M. Robinson (eds.), Nonlinear Programming 2, Academics Press, New York, 279–312.

[3] Balas, E., 1975, Facets of the knapsack polytope, Mathematical Programming 8, 146–164.

[4] Balas, E., 1979, Disjunctive programming, Ann. Discr. Math. 5, 3–51.

[5] Balas, E., S. Ceria, and G. Cornuèjols, 1993, A lift-and-project cutting plane algorithm for mixed 0-1 programs, Mathematical Programming 58, 295–324.

[6] Balas, E., S. Ceria, and G. Cornuèjols, 1993, Solving Mixed 0-1 Programs by a Lift-and-Project Method, SODA 1993, 232-242.

[7] Balas, E., S. Ceria, and G. Cornuèjols, 1996, Mixed 0-1 programming by lift-and-project in a branch-and-cut framework, Management Science 42, 1229–1246.

[8] Blair, C. E., and R. G. Jeroslow, 1978, A converse for disjunctive constraints, Journal of Optimization Theory and Applications 25, 195–206.

[9] Ceria, S., and J. Soares, 1997, Disjunctive cut generation for mixed 0-1 programs: duality and lifting, Working paper, Graduate School of Business, Columbia University.

[10] Dentcheva, D., and A. Ruszczyński, 2003, Optimization with stochastic dominance constraints, SIAM Journal on Optimization 14, 548–566.

[11] Dentcheva, D., and A. Ruszczyński, 2004, Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints, Mathematical Programming 99, 329–350.

[12] Dentcheva, D., and A. Ruszczyński, 2004, Convexification of stochastic ordering, Comptes Rendus de l'Academie Bulgare des Sciences 57, No. 4, 7–14.

[13] Dentcheva, D., and A. Ruszczyński, 2004, Semi-Infinite Probabilistic Optimization: First Order Stochastic Dominance Constraints, Optimization 53, 583–601.

[14] Dentcheva, D., and A. Ruszczyński, 2006, Portfolio Optimization with First Order Stochastic Dominance Constraints, Journal of Banking and Finance 30/2, 433–451.

[15] Fishburn, P. C., 1970, Utility Theory for Decision Making, Wiley, New York.

[16] Fourer, R., D. M. Gay and B. W. Kernighan, 1993, AMPL: A Modelling Language for Mathematical Programming, The Scientific Press.

[17] Hadar, J., and W. Russell, 1969, Rules for ordering uncertain prospects, Amer. Econom. Rev. 59, 25–34.

[18] Hammer, P. L, E. L. Johnson and U. N. Peled, 1975, Facets of regular 0-1 polytopes, Mathematical Programming 8, 179–206.

[19] Hanoch, G., and H. Levy, 1969, The efficiency analysis of choices involving risk, Rev. Econom. Stud. 36, 335–346.

[20] Klatte, D. and R. Henrion, 1998, Regularity and stability in nonlinear semi-infinite optimization, Nonconvex optimization and Applications 25, 69–102.

[21] ILOG CPLEX, 2000, CPLEX 7.0 Users Manual and Reference Manual, ILOG CPLEX Division, Incline Village, NV.

[22] Lehmann, E., 1955, Ordered families of distributions, Annals of Mathematical Statistics 26, 399–419.

[23] Levy, H., 1992, Stochastic dominance and expected utility: survey and analysis, Management Science, 38, 555–593.

[24] Mann, H. B., and D. R. Whitney, 1947, On a test of whether one of two random variables is stochastically larger than the other, Annals of Mathematical Statistics 18, 50–60.

[25] Markowitz, H. M., 1952, Portfolio Selection, Journal of Finance 7, 77-91.

[26] Markowitz, H. M., 1959, Portfolio Selection, John Wiley & Sons, New York.

[27] Müller A., and D. Stoyan, 2002, Comparison Methods for Stochastic Models and Risks, John Wiley & Sons, Chichester.

[28] Noyan, N., G. Rudolf, and A. Ruszczyński, 2005, Relaxations of Linear Programming Problems with First Order Stochastic Dominance Constraints, accepted for publication in Operations Research Letters, available online at www.sciencedirect.com.

[29] Ogryczak, W. and A. Ruszczyński, 1999, From stochastic dominance to mean-risk models: semideviations as risk measures, European Journal of Operations Research 116, 33–50.

[30] Ogryczak, W. and A. Ruszczyński, 2001, Dual stochastic dominance and related mean-risk models, SIAM Journal of Optimization 13, 60–78.

[31] Padberg, M. W, 1975, A note on zero-one programming, Operations Research 23, 833–837.

[32] Perregaard, M., 2003, Generating Disjunctive Cuts for Mixed Integer Programs, Doctoral Dissertation, Graduate School of Industrial Administration, Schenley Park, Pittsburgh, Carnegie Mellon University.

[33] Quirk, J. P., and R. Saposnik, 1962, Admissibility and measurable utility functions, Review of Economic Studies 29, 140–146.

[34] Rockafellar, R. T. and S. Uryasev, 2002, Optimization of Conditional Value-at-Risk, Journal of Risk, 21–41.

[35] Rothschild, M., and J. E. Stiglitz, 1969, Increasing Risk: I. A definition, J. Econom. Theory 2, 225–243.

[36] Rudolf, G. and A. Ruszczyński, 2006, A Dual Approach to Linear Stochastic Optimization Problems with Second Order Dominance Constraints, submitted for publication.

[37] Ruszczynski, A. and R. J. Vanderbei, 2003, Frontiers of Stochastically Nondominated Portfolios, Econometrica 71, 1287–1297.

[38] Whitmore, G. A. and M. C. Findlay, 1978, Stochastic Dominance: An Approach to Decision-Making Under Risk, D.C.Heath, Lexington, MA.

[39] Wolsey, L. A, 1975, Faces for a linear inequality in 0-1 variables, Mathematical Programming 8, 165–178.