Math. Program., Ser. A (2011) 126:179–190
DOI 10.1007/s10107-009-0270-0

FULL LENGTH PAPER

A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse

Ralf Gollmer · Uwe Gotzes · Rüdiger Schultz

Received: 5 February 2007 / Accepted: 8 January 2009 / Published online: 13 February 2009

Abstract We introduce stochastic integer programs with second-order dominance constraints induced by mixed-integer linear recourse. Closedness of the constraint set mapping with respect to perturbations of the underlying probability measure is derived. For discrete probability measures, large-scale, block-structured, mixed-integer linear programming equivalents to the dominance constrained stochastic programs are identified. For these models, a decomposition algorithm is proposed and tested with instances from power optimization.

Keywords Stochastic integer programming · Stochastic dominance · Mixed-integer optimization

Mathematics Subject Classification (2000) 90C15 · 90C11 · 60E15

1 Introduction

In [13] we have studied optimization problems with first-order stochastic dominance constraints where the relevant random variables are given by objective values of stochastic programs with mixed-integer linear recourse. In this note we derive counterparts to the results of [13] when the models are based on second-order stochastic dominance.

Let X and Y be real-valued random variables. Following [16], we say that X is stochastically smaller in first order than $\mathsf{Y}(\mathsf{X} \preceq_1 \mathsf{Y})$ iff $\mathbb{E}h(\mathsf{X}) \leq \mathbb{E}h(\mathsf{Y})$ for all

R. Gollmer · U. Gotzes · R. Schultz (✉)
Department of Mathematics, University of Duisburg-Essen,
Campus Duisburg, 47048 Duisburg, Germany
e-mail: schultz@math.uni-duisburg.de

nondecreasing disutility functions h for which both expectations exist. Investigations in [13] refer to this concept of stochastic order.

X is said to be stochastically smaller than Y in increasing convex order ($\mathsf{X} \preceq_{icx} \mathsf{Y}$) iff $\mathbb{E}h(\mathsf{X}) \leq \mathbb{E}h(\mathsf{Y})$ for all nondecreasing convex disutility functions h for which both expectations exist. For a decision maker who is rational in the sense of preferring less to more $\preceq_1$ and $\preceq_{icx}$ correspond to the traditional first- and second-order dominance rules which assume preference of more to less.

Consider the following dominance constrained stochastic program

$$\min\{g(x) : f(x, z(\omega)) \preceq_{icx} a(\omega), \;\; x \in X\}. \tag{1}$$

The random variables $f(x, z(\omega))$ arise from mixed-integer linear recourse, i.e., they are specified as

$$f(x, z(\omega)) := c^\top x + \Phi(z(\omega) - Tx)$$

where

$$\Phi(t) := \min\left\{q^\top y : Wy = t, \;\; y \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}\right\}. \tag{2}$$

This setting results from the random optimization problem

$$\min\left\{c^\top x + q^\top y : Tx + Wy = z(\omega), \;\; x \in X, \;\; y \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}\right\} \tag{3}$$

together with the nonanticipativity requirement that x must be selected prior to observing $z(\omega)$ and deciding on $y = y(x, \omega)$. For further modeling details on stochastic programs with recourse see [5,15,18,20].

Basic assumptions in (3) are that W has rational entries, and that $X \subseteq \mathbb{R}^m$ is a nonempty polyhedron, possibly involving integer requirements to components of x. Further assumptions assuring well-posedness of Φ will be given later in the text.

With a predesigned benchmark random variable $a(\omega)$, the constraints of (1) single out all those feasible nonanticipative x in (3) which lead to total costs $f(x, z(\omega))$ that, in terms of $\preceq_{icx}$, are stochastically smaller than $a(\omega)$. Compared with the counterpart model in [13], this comparison must "withstand" a smaller class of disutility functions (non-decreasing convex vs. non-decreasing), leading to a weaker requirement on x. The objective function $g : \mathbb{R}^m \to \mathbb{R}$ in (1) shall be at least lower semicontinuous, but will be restricted to a linear function $g^\top x$ in our algorithmic considerations.

The study of stochastic dominance constraints in optimization was initialized in [9–11,17]. Although we address specific random variables $f(x, z(\omega))$, the structural results of [9–11,17] are not applicable here. This is due to the lacking continuity of the mixed-integer value function Φ from (2).

Our paper is organized as follows. In Sect. 2, we derive some basic structural properties of (1). In Sect. 3, we outline algorithmic aspects including computations.

2 Structural properties

Let $\mathcal{P}(\mathbb{R}^s)$, $\mathcal{P}(\mathbb{R})$ be the sets of all Borel probability measures on $\mathbb{R}^s$ and $\mathbb{R}$, and let $\mu \in \mathcal{P}(\mathbb{R}^s)$ and $\nu \in \mathcal{P}(\mathbb{R})$ denote the image measures induced by $z(\omega)$, $a(\omega)$ on $\mathbb{R}^s$ and $\mathbb{R}$, respectively. The constraint $f(x, z(\omega)) \preceq_{icx} a(\omega)$ in (1) can be equivalently expressed as

$$\int_{\mathbb{R}^s} [f(x,z) - \eta]_+ \, \mu(dz) \leq \int_{\mathbb{R}} [a - \eta]_+ \, \nu(da) \quad \forall \eta \in \mathbb{R}, \tag{4}$$

provided all objects in (4) are well-defined, see [16]. Recall that Φ is the value function of a mixed-integer linear program, cf. (2). Assume

(A1) (complete recourse) $W(\mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}) = \mathbb{R}^s$,
(A2) (sufficiently expensive recourse) $\{u \in \mathbb{R}^s : W^\top u \leq q\} \neq \emptyset$.

Then (A1) and (A2), together with the rationality of W, imply that Φ is real-valued and lower semicontinuous on $\mathbb{R}^s$, i.e., $\liminf_{t_n \to t} \Phi(t_n) \geq \Phi(t)$ for all $t \in \mathbb{R}^s$, [2,6]. Moreover, there exist $\alpha > 0$, $\beta > 0$ such that for all $t_1, t_2 \in \mathbb{R}^s$

$$|\Phi(t_1) - \Phi(t_2)| \leq \alpha \|t_1 - t_2\| + \beta. \tag{5}$$

This settles well-posedness of the integrands in (4). For finiteness of the integrals we assume

$$\text{(A3)} \quad \text{(finite first moments)} \qquad \int_{\mathbb{R}^s} \|z\| \, \mu(dz) < \infty, \quad \int_{\mathbb{R}} |a| \nu(da) < \infty.$$

Using (5) and the fact that (A2) implies $\Phi(0) = 0$, we obtain that for fixed x there is a constant $\kappa > 0$ such that

$$|[f(x,z) - \eta]_+| \leq \alpha \|z\| + \kappa \quad \forall z \in \mathbb{R}^s.$$

Hence, (A1)–(A3) imply that the integral on the left in (4) is always finite. For the integral on the right, (A3) ensures this property.

In accordance with (A3) we denote by $\mathcal{P}_1(\mathbb{R}^s)$ the subset of $\mathcal{P}(\mathbb{R}^s)$ with measures having finite first moment. We consider the multifunction $C : \mathcal{P}_1(\mathbb{R}^s) \to 2^{\mathbb{R}^m}$ where

$$C(\nu) := \{x \in \mathbb{R}^m : f(x,z) \preceq_{icx} a, \ x \in X\} \quad \nu \in \mathcal{P}_1(\mathbb{R}). \tag{6}$$

The space $\mathcal{P}_1(\mathbb{R}^s)$ is equipped with weak convergence of probability measures [3]. A sequence $\{\mu_n\}$ in $\mathcal{P}_1(\mathbb{R}^s)$ is said to converge weakly to $\mu \in \mathcal{P}_1(\mathbb{R}^s)$, written $\mu_n \xrightarrow{w} \mu$, if for any bounded continuous function $h : \mathbb{R}^s \to \mathbb{R}$ it holds $\int_{\mathbb{R}^s} h(z)\mu_n(dz) \to \int_{\mathbb{R}^s} h(z)\mu(dz)$ as $n \to \infty$.

Our aim is to show that C is a closed multifunction on $\mathcal{P}_1(\mathbb{R}^s)$. This means that for arbitrary $\mu \in \mathcal{P}_1(\mathbb{R}^s)$ and sequences $\mu_n \in \mathcal{P}_1(\mathbb{R}^s)$, $x_n \in C(\mu_n)$ with $\mu_n \xrightarrow{w} \mu$ and $x_n \to x$ it follows that $x \in C(\mu)$.

Lemma 2.1 *Let $\mu_n, \mu \in \mathcal{P}(\mathbb{R}^s)$ with $\mu_n \xrightarrow{w} \mu$ and $h : \mathbb{R}^s \to \mathbb{R}$ be lower semicontinuous with $h(z) \geq 0\ \forall z \in \mathbb{R}^s$. Then*

$$\int_{\mathbb{R}^s} h(z)\,\mu(dz) \;\leq\; \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz).$$

Proof We start with the bounded case and assume there exist $\underline{h}, \bar{h} \in \mathbb{R}$ such that $\underline{h} < h(z) < \bar{h}\ \forall z \in \mathbb{R}^s$. Without loss of generality we assume $0 < h(z) < 1\ \forall z \in \mathbb{R}^s$ which can be achieved by affine scaling. □

Fix $k \in \mathbb{N}$ and consider the sets $H_i := \{z \in \mathbb{R}^s : i/k < h(z)\},\ i = 0, \ldots, k$. Since h is lower semicontinuous, H_i is open for all i. It holds

$$\sum_{i=1}^{k} \frac{i-1}{k}\, \mu\left[\left\{z : \frac{i-1}{k} < h(z) \leq \frac{i}{k}\right\}\right] \leq \int_{\mathbb{R}^s} h(z)\,\mu(dz)$$
$$\leq \sum_{i=1}^{k} \frac{i}{k}\, \mu\left[\left\{z : \frac{i-1}{k} < h(z) \leq \frac{i}{k}\right\}\right].$$

The sum on the right equals

$$\sum_{i=1}^{k} \frac{i}{k}\, (\mu[H_{i-1}] - \mu[H_i]) = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^{k} \mu[H_i],$$

while the sum on the left is identical with

$$\sum_{i=1}^{k} \frac{i-1}{k}\, (\mu[H_{i-1}] - \mu[H_i]) = \frac{1}{k} \sum_{i=1}^{k} \mu[H_i].$$

Putting this together yields

$$\frac{1}{k} \sum_{i=1}^{k} \mu[H_i] \leq \int_{\mathbb{R}^s} h(z)\,\mu(dz) \leq \frac{1}{k} + \frac{1}{k} \sum_{i=1}^{k} \mu[H_i]. \tag{7}$$

By the Portmanteau Theorem (see [3, Theorem 2.1, pp. 11/12]) we have for all i

$$\mu[H_i] \leq \liminf_n\ \mu_n[H_i]. \tag{8}$$

Applying the left inequality in (7) to μ_n and taking the limes inferior provides

$$\frac{1}{k}\ \liminf_n \sum_{i=1}^{k} \mu_n[H_i] \leq \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz),$$

and, together with (8),

$$\frac{1}{k}\sum_{i=1}^{k}\mu[H_i] \leq \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz).$$

Now we apply the right inequality in (7) and obtain

$$-\frac{1}{k} + \int_{\mathbb{R}^s} h(z)\,\mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz).$$

With $k \to \infty$ this yields the assertion for bounded h. For extension to unbounded non-negative h let $r \in \mathbb{R}_+$ and consider the truncated function $h_r : \mathbb{R}^s \to \mathbb{R}$ with

$$h_r(z) := \begin{cases} h(z), & \text{if } h(z) \leq r \\ r\ , & \text{otherwise.} \end{cases}$$

Lower semicontinuity of h implies lower semicontinuity of h_r for all $r \in \mathbb{R}_+$. The assertion then is valid for h_r, since h_r is bounded. Moreover, $h_r(z) \leq h(z)\,\forall z \in \mathbb{R}^s$. This yields

$$\begin{aligned} \int_{\mathbb{R}^s} h_r(z)\,\mu(dz) &\leq \liminf_n \int_{\mathbb{R}^s} h_r(z)\,\mu_n(dz) \\ &\leq \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz) \quad \forall r \in \mathbb{R}_+. \end{aligned} \tag{9}$$

The Monotone Convergence Theorem (see for instance [4, Theorem 16.2, p. 211]) yields

$$\int_{\mathbb{R}^s} h_r(z)\,\mu(dz) \longrightarrow \int_{\mathbb{R}^s} h(z)\,\mu(dz) \quad \text{for } r \to \infty.$$

Together with (9) this implies

$$\int_{\mathbb{R}^s} h(z)\,\mu(dz) \leq \liminf_n \int_{\mathbb{R}^s} h(z)\,\mu_n(dz),$$

and the proof is complete. □

Proposition 2.2 *Assume (A1)–(A3). Then the multifunction C, as defined in* (6)*, is closed on $\mathcal{P}_1(\mathbb{R}^s)$.*

Proof Let $\mu_n, \mu \in \mathcal{P}_1(\mathbb{R}^s)$ and $x_n \in C(\mu_n)$ such that $\mu_n \xrightarrow{w} \mu$ and $x_n \to x$. Closedness of X then immediately yields $x \in X$. According to (4), $x_n \in C(\mu_n)$ implies

$$\int\limits_{\mathbb{R}^s} [f(x_n, z) - \eta]_+ \, \mu_n(dz) \leq \int\limits_{\mathbb{R}} [a - \eta]_+ \, \nu(da) \quad \forall \eta \in \mathbb{R}. \tag{10}$$

Notice that the integrands $[f(.,.) - \eta]_+$ are non-negative and lower semicontinuous for all $\eta \in \mathbb{R}$. Together with Fatou's Lemma (see for instance [4, Theorem 16.3, p. 212]), this implies

$$\begin{aligned}\int\limits_{\mathbb{R}^s} [f(x, z) - \eta]_+ \mu_n(dz) &\leq \int\limits_{\mathbb{R}^s} \liminf_k \, [f(x_k, z) - \eta]_+ \, \mu_n(dz) \\ &\leq \liminf_k \int\limits_{\mathbb{R}^s} [f(x_k, z) - \eta]_+ \, \mu_n(dz)\end{aligned}$$

for all $\eta \in \mathbb{R}$. Taking the limes inferior with respect to n on both sides we obtain

$$\begin{aligned}\liminf_n \int\limits_{\mathbb{R}^s} [f(x, z) - \eta]_+ \, \mu_n(dz) &\leq \liminf_n \, \liminf_k \int\limits_{\mathbb{R}^s} [f(x_k, z) - \eta]_+ \, \mu_n(dz) \\ &\leq \liminf_n \int\limits_{\mathbb{R}^s} [f(x_n, z) - \eta]_+ \, \mu_n(dz) \\ &\leq \int\limits_{\mathbb{R}} [a - \eta]_+ \, \nu(da) \quad \forall \eta \in \mathbb{R}.\end{aligned}$$

Here the second inequality follows from passing to the diagonal sequence where $n = k$, and the third inequality follows from (10). Applying Lemma 2.1 with $h(z) := [f(x, z) - \eta]_+$ implies

$$\begin{aligned}&\int\limits_{\mathbb{R}^s} [f(x, z) - \eta]_+ \, \mu(dz) \\ &\quad \leq \liminf_n \int [f(x_n, z) - \eta]_+ \, \mu_n(dz) \leq \int\limits_{\mathbb{R}} [a - \eta]_+ \, \nu(da) \quad \forall \eta \in \mathbb{R}\end{aligned}$$

and thus $x \in C(\mu)$. □

For the first-order dominance relation $\preceq_1$, the characterization in the spirit of (4) is based on probabilities of lower left orthants. In [13], when establishing the closedness of C(.), therefore the standard Portmanteau Theorem [3, Theorem 2.1] could be used where it took its extension to expectations (Lemma 2.1) in the above analysis. Moreover, in [13] the lower semicontinuity of the probability measure [4, Theorem 4.1] was employed where it took Fatou's Lemma in the above proof of Proposition 2.2.

Of course, by setting μ_n identical to μ for all n, Proposition 2.2 implies that $C(\mu)$ is a closed subset of $\mathbb{R}^m$ for all $\mu \in \mathcal{P}_1(\mathbb{R}^s)$.

Closedness of the multifunction C is the key to proving lower semicontinuity of the optimal value function given by $\varphi(\mu) := \inf\{g(x) : x \in C(\mu)\}$. For instance, if X is nonempty and compact, g lower semicontinuous, and (A1)–(A3) are valid, then φ is lower semicontinuous at all $\bar{\mu} \in \mathcal{P}_1(\mathbb{R}^s)$ for which the optimization problem defining $\varphi(\bar{\mu})$ is solvable. The proof follows the lines of Berge's classical theory, see for instance [1] or [13].

Approximation schemes in stochastic programming often lead to weakly converging sequences of discrete probability measures. It is generally desired that convergence of approximations implies convergence of optimal values and/or optimal solutions. The above analysis provides some first results along this line. In this way it contributes to justifying that algorithm design is done for models living on finite probability spaces.

3 Algorithmic aspects

3.1 MILP equivalent and decomposition algorithm

The following proposition establishes a mixed-integer linear programming (MILP) equivalent to (1) for finite probability spaces. Its proof, for details see [12], is obtained by adapting the steps of the proof of Proposition 3.1 in [13]. In particular, it is beneficial that, with finitely distributed benchmark $a(\omega)$, the relation $f(x, z(\omega)) \preceq_{icx} a(\omega)$ already holds when (4) is valid for all η in the finite set of mass points of the benchmark.

Proposition 3.1 *Let $z(\omega)$ and $a(\omega)$ in (1) follow discrete distributions with realizations $z_l,\ l = 1, \ldots, L$, and $a_k,\ k = 1, \ldots, K$, as well as probabilities $\pi_l,\ l = 1, \ldots, L$, and $p_k,\ k = 1, \ldots, K$, respectively. Let further $g(x) := g^\top x$ be linear. Assume (A1) and (A2). Then (1) is equivalent to the mixed-integer linear program*

$$\min\left\{g^\top x : \begin{array}{l} c^\top x + q^\top y_{lk} - a_k \le \ \ v_{lk} \quad \forall l\ \forall k \\ Tx + W y_{lk} = z_l \quad \forall l\ \forall k \\ \sum_{l=1}^{L} \pi_l v_{lk} \le \bar{a}_k \quad \forall k \\ x \in X,\ \ y_{lk} \in \mathbb{Z}_+^{\bar{m}} \times \mathbb{R}_+^{m'}, v_{lk} \ge 0 \quad \forall l\ \forall k \end{array}\right\} \tag{11}$$

where $\bar{a}_k := \int_{\mathbb{R}} [a - a_k]_+ \, \nu(da),\ k = 1, \ldots, K$.

Inspecting (11) we observe that the constraints

$$\sum_{l=1}^{L} \pi_l v_{lk} \le \bar{a}_k \quad \forall k \tag{12}$$

are the only ones coupling explicitly second-stage variables, namely v_{lk}, across different scenarios l. In other words, without (12), we arrive at the "usual" block structure in two-stage stochastic programming where nonanticipativity of x is the only condition interlinking constraints from different scenarios, [5,7,15,18–20]. Relaxation of nonanticipativity then leads to decomposition into single-scenario subproblems and to a lower bound to (11). An upper bound to (11) can be obtained by a feasibility heuristics, such as the subsequent Algorithm 3.3.

The bounding procedures are iterated by understanding (11) as an "expanded" representation of the nonconvex global minimization problem (1) and employing a branch-and-bound algorithm, such as the subsequent Algorithm 3.2. In [13] this principal algorithmic approach has also been taken for the first-order counterpart model to (1). The difference is in the coupling constraints (12). For the first-order model they involve probability functionals, thus leading to additional Boolean variables in the MILP equivalent.

By $\mathbf{P}$ we denote a list of problems, and $\varphi_{LB}(P)$ is a lower bound for the optimal value of $P \in \mathbf{P}$. Moreover, $\bar{\varphi}$ denotes the currently best upper bound to the optimal value of (11), and $X(P)$ is the element in the partition of X belonging to P.

Algorithm 3.2

STEP 1 (INITIALIZATION):
Let $\mathbf{P} := \{(11)\}$ *and* $\bar{\varphi} := +\infty$.

STEP 2 (TERMINATION):
If $\mathbf{P} = \emptyset$ *then the* $\bar{x}$ *that yielded* $\bar{\varphi} = g^\top \bar{x}$ *is optimal.*

STEP 3 (BOUNDING):
Select and delete a problem P *from* $\mathbf{P}$. *Compute a lower bound* $\varphi_{LB}(P)$ *and apply a feasibility heuristics to find a feasible point* $\bar{x}$ *of* P.

STEP 4 (PRUNING):
If $\varphi_{LB}(P) = +\infty$ *(infeasibility of a subproblem)) or* $\varphi_{LB}(P) > \bar{\varphi}$ *(inferiority of* P*), then go to Step 2.*
If $\varphi_{LB}(P) = g^\top \bar{x}$ *(optimality for* P*), then check whether* $g^\top \bar{x} < \bar{\varphi}$. *If yes, then* $\bar{\varphi} := g^\top \bar{x}$. *Go to Step 2.*
If $g^\top \bar{x} < \bar{\varphi}$, *then* $\bar{\varphi} := g^\top \bar{x}$.

STEP 5 (BRANCHING):
Create two new subproblems by partitioning the set $X(P)$ *by means of linear inequalities. Add these subproblems to* $\mathbf{P}$ *and go to Step 2.*

The following feasibility heuristics uses as input a bunch $\tilde{x}_l$, $l = 1, \ldots, L$, of first-stage vectors corresponding to the individual scenarios. This is a mild assumption since lower bounding procedures leading to scenario decomposition typically provide such vectors as x_l-parts of optimal solutions to single-scenario subproblems.

Algorithm 3.3

STEP 1:
Using $\tilde{x}_l$, $l = 1, \ldots, L$, *pick a "reasonable candidate"* $\bar{x}$, *for instance one arising most frequently, or average the* $\tilde{x}_l$, $l = 1, \ldots, L$, *and round to integers if necessary.*

STEP 2:
Solve for each $l = 1, \ldots, L$:

$$\left.\begin{aligned} \min\left\{\textstyle\sum_{k=1}^{K} v_{lk}\right. &: c^{\top}\bar{x} + q^{\top}y_{lk} - a_k \ \leq \ v_{lk} \\ T\bar{x} + Wy_{lk} &= z_l \\ y_{lk} \in \mathbb{Z}_{+}^{\bar{m}} \times \mathbb{R}_{+}^{m'}, &\ v_{lk} \geq 0, \quad k = 1, \ldots, K \end{aligned}\right\} \tag{13}$$

If one of these problems fails to be feasible, $\bar{x}$ cannot be feasible for (11), and the heuristics stops with assigning the formal upper bound $+\infty$. Otherwise, go to Step 3.

STEP 3:
Check whether the optimal v_{lk} found in (13) fulfil

$$\sum_{l=1}^{L} \pi_l v_{lk} \leq \bar{a}_k \quad k = 1, \ldots, K.$$

If yes, then a feasible solution to (11) is found. The heuristics stops with the upper bound $g^{\top}\bar{x}$. Otherwise, the heuristics stops without a feasible solution to (11) and assigns the formal upper bound $+\infty$.

We have implemented Algorithm 3.2 using a lower bounding procedure which ignores the nonanticipativity of x and performs Lagrangean relaxation of (12). Upper bounding follows Algorithm 3.3, with the input tuple $\tilde{x}_l$, $l = 1, \ldots, L$, taken from optimal solutions to single-scenario subproblems that arise in the course of the Lagrangean relaxation for optimal or nearly optimal solutions to the Lagrangean dual, for details see [12].

3.2 Computations

For an impression on the numerical performance of our decomposition algorithm we report tests with instances from the operation of a dispersed energy system under uncertainty of consumer demand, infeed from renewables, and fuel as well as power prices. See [14] for a detailed model description. Typically, the uncertainty prone data are known with certainty for an initial time period of the planning horizon. This gives rise to a random optimization problem (3), with nonanticipativity concerning the decision variables of the initial period. The objective function $g(x)$ in (1) is the sum over all start-ups of units in the initial period. The dominance constrained model (1) then aims at minimizing abrasion of the generation units over all technically and economically feasible generation policies leading to overall costs that are stochastically smaller in increasing convex order than a given cost benchmark.

The MILP corresponding to a single-scenario instance of the random optimization problem (3) has about 17,500 variables (9,000 boolean, 8,500 continuous) and

Table 1 Results for instances with 30 data scenarios and 4 benchmark scenarios

Instance	Benchmarks		Time (s)	Cplex		ddsip.vSD	
	Probability	Benchmark value		Upper bound	Lower bound	Upper bound	Lower bound
1	0.085	2895000	697.21	–	29	29	9
	0.14	4851000	2471.63	–	29	29	29
	0.635	7789000	7520.07	29	29	29	29
	0.14	10728000					
2	0.085	2900000	702.31	–	27	27	9
	0.14	4860000	3635.25	–	27	27	27
	0.635	7800000	14905.68	–	27 out of mem.	27	27
	0.14	10740000					
3	0.085	3000000	666.31	–	18	18	9
	0.14	5000000	3907.92	–	18	18	18
	0.635	8000000	7181.68	18	18	18	18
	0.14	11000000					
4	0.085	3500000	500.05	–	11	11	9
	0.14	5500000	1404.96	–	11	11	11
	0.635	8500000	6559.52	11	11	11	11
	0.14	11500000					
5	0.085	4000000	474.68	–	8	8	8
	0.14	6000000	6076.40	8	8	8	8
	0.635	9000000					
	0.14	12000000					

22,000 constraints. With $K = 4$, $L = 30$, the MILP equivalent (11) to the dominance constrained model (1) has about 1,740,000 variables (894,000 Boolean, 846,000 continuous) and 2,220,000 constraints. For $K = 4$, $L = 50$, we have about 2,898,000 variables (1,489,000 Boolean, 1,409,000 continuous) and 3,698,000 constraints.

In Tables 1 and 2 we compare results for the MILP equivalents obtained with the solver Cplex [8] to results computed with our implementation ddsip.vSD of Algorithm 3.2. Computations were done on a Linux-PC with a 3.2 GHz Pentium processor and 2GB RAM. The time limit was set to eight hours.

For instances with different benchmarks, the tables report the times when a first feasible solution was found ("Upper Bound") and the times when optimality was verified (coincidence of values in the "Upper Bound" and "Lower Bound" columns). Algorithm 3.2 was always fastest, both in finding the first feasible solution and in verifying optimality. With growing problem size Cplex had more and more trouble in finding a feasible solution and started to run out of memory. With $L = 50$, the available memory was insufficient to build up the model (lp-)file to be used by Cplex. Algorithm 3.2 was still able to solve these problems to optimality. Altogether, Tables 1 and 2 clearly indicate the superiority of the decomposition approach behind Algorithm 3.2. Further computational experiments supporting this hypothesis can be found in [12].

Table 2 Results for instances with 50 data scenarios and 4 benchmark scenarios

Instance	Benchmarks		Time (s)	Cplex		ddsip.vSD	
	Probability	Benchmark value		Upper bound	Lower bound	Upper bound	Lower bound
1	0.09	2895000	1084.68	–	–	29	9
	0.135	4851000	3747.69	–	–	29	29
	0.67	7789000					
	0.105	10728000					
2	0.09	2900000	1125.39	–	–	27	9
	0.135	4860000	5857.67	–	–	27	27
	0.67	7800000					
	0.105	10740000					
3	0.09	3000000	1041.15	–	–	18	9
	0.135	5000000	6126.89	–	–	18	18
	0.67	8000000					
	0.105	11000000					
4	0.09	3500000	1026.21	–	–	11	9
	0.135	5500000	2872.83	–	–	11	11
	0.67	8500000					
	0.105	11500000					
5	0.09	4000000	1096.69	–	–	8	8
	0.135	6000000					
	0.67	9000000					
	0.105	12000000					

Acknowledgments We wish to thank Christoph Helmberg (Technical University of Chemnitz) for giving us access to the implementation of his spectral bundle method. Further thanks are due to Frederike Neise (University of Duisburg-Essen) for fruitful discussions and her support in computational testing. Parts of this paper were written while the third author was visiting the Centro de Modelamiento Matematico, Universidad de Chile, Santiago. Partial funding for this research was provided by the German Federal Ministry of Education and Research (BMBF) under grants 03SCNIVG and 03SF0312A. We thank these institutions for their support.

References

1. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Non-linear Parametric Optimization. Akademie, Berlin (1982)
2. Bank, B., Mandel, R.: Parametric Integer Optimization. Akademie, Berlin (1988)
3. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
4. Billingsley, P.: Probability and Measure. Wiley, New York (1986)
5. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer, New York (1997)
6. Blair, C.E., Jeroslow, R.G.: The value function of a mixed integer program: I. Discrete Math. **19**, 121–138 (1977)
7. Carøe, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. Oper. Res. Lett. **24**, 37–45 (1999)
8. CPLEX Callable Library 9.1.3, ILOG (2005)

9. Dentcheva, D., Henrion, R., Ruszczyński, A.: Stability and sensitivity of optimization problems with first order stochastic dominance constraints. SIAM J. Optim. **18**, 322–337 (2007)
10. Dentcheva, D., Ruszczyński, A.: Optimization with stochastic dominance constraints. SIAM J. Optim. **14**, 548–566 (2003)
11. Dentcheva, D., Ruszczyński, A.: Optimality and duality theory for stochastic optimization with non-linear dominance constraints. Math. Program. **99**, 329–350 (2004)
12. Gollmer, R., Gotzes, U., Schultz, R.: Second-order stochastic dominance constraints induced by mixed-integer linear recourse, Preprint 644/2007, Department of Mathematics, University of Duisburg-Essen (2007); also available online from http://www.uni-duisburg.de/FB11/disma/preprints.shtml
13. Gollmer, R., Neise, F., Schultz, R.: Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. SIAM J. Optim. **19**, 552–571 (2008)
14. Handschin, E., Neise, F., Neumann, H., Schultz, R.: Optimal operation of dispersed generation under uncertainty using mathematical programming. Int. J. Electr. Power Energy Syst. **28**, 618–626 (2006)
15. Kall, P., Wallace, S.W.: Stochastic Programming. Wiley, Chichester (1994)
16. Müller, A., Stoyan, D.: Comparison Methods for Stochastic Models and Risks. Wiley, Chichester (2002)
17. Noyan, N., Rudolf, G., Ruszczyński, A.: Relaxations of linear programming problems with first order stochastic dominance constraints. Oper. Res. Lett. **34**, 653–659 (2006)
18. Prékopa, A.: Stochastic Programming. Kluwer, Dordrecht (1995)
19. Ruszczyński, A.: Decomposition methods in stochastic programming. Math. Program. **79**, 333–353 (1997)
20. Ruszczyński, A., Shapiro, A. (eds.): Handbooks in Operations Research and Management Science, 10: Stochastic Programming. Elsevier, Amsterdam (2003)