

ALGORITHMS FOR MEAN-RISK STOCHASTIC INTEGER PROGRAMS IN ENERGY

Rüdiger Schultz, Frederike Neise

Department of Mathematics

University of Duisburg-Essen, Campus Duisburg

Lotharstr. 65, D-47048 Duisburg, Germany

schultz@math.uni-duisburg.de, neise@math.uni-duisburg.de

ABSTRACT –

We introduce models and algorithms suitable for including risk aversion into stochastic programming problems in energy. For a system with dispersed generation of power and heat we present computational results showing the superiority of our decomposition algorithm over a standard mixed-integer linear programming solver.

RESUMEN:

Introducimos modelos y algoritmos adecuados para incluir la aversión al riesgo en problemas de programación estocástica en la energía. Para un sistema con una dispersión de la generación de potencia y calor presentamos resultados computacionales que muestran la superioridad de cuatro algoritmos de descomposición sobre los típicos resolutores de la programación lineal entera mixta.

KEY WORDS : Dispersed storage and generation, Cogeneration, Renewable Resources, Mathematical Programming, Uncertainty, Decomposition methods, Risk aversion

MSC: 90C15

1. INTRODUCTION

Incomplete information is becoming a more and more prevailing issue in energy optimization problems. Power market liberalization and the impending decentralization of electricity supply, the latter with substantial infeed from renewable resources, are just two phenomena that have fostered this development. In addition, the appropriate handling of risk is of growing importance in the power industry.

Stochastic programming is a proven tool for handling data uncertainty in optimization problems, [1]. The mentioned facts have motivated an extension of traditional stochastic programming methodology in various directions, of which, in the present paper, we address the inclusion of integer decision variables and the transition from risk neutral models to those incorporating risk aversion.

Our paper is organized as follows: In Section 2 we introduce mean-risk extensions of risk neutral stochastic integer programs. In Section 3 we study the block structures that arise in the equivalent mixed-integer linear programs if the underlying probability distributions are discrete. Section 4 is devoted to decomposition and approximation algorithms exploiting these block structures. Computational tests for illustrative examples based on an energy system with dispersed generation of power and heat are reported in Section 5. Finally, we have a conclusions section.

2. STOCHASTIC INTEGER PROGRAMS WITH RISK AVERSION

To explain basic features of extending risk neutral into mean-risk stochastic integer programs let us start out from the following random optimization problem

$$\min \{c^T x + q(\omega)^T y : T(\omega)x + W(\omega)y = z(\omega), x \in X, y \in Y\}, \quad (1)$$

together with the information constraint that x must be selected without anticipation of the realization of the random data $(q, T, W, z)(\omega)$. This leads to a two-stage scheme of alternating decision and observation: The decision on x is followed by observing $(q, T, W, z)(\omega)$ and then y is taken, thus

depending on x and ω . Accordingly, x and y are called first- and second-stage decisions, respectively.

In energy optimization such a setting is plausible in many contexts. For example when planning a dispersed generation system the decisions on what types of and how many generators to install correspond to x , and operation decisions (under stochastic demand, power prices, or infeed from renewable resources) correspond to y , [2]. Or, as another example, take day-ahead trading (with one round of bidding) for a utility. Here the first stage corresponds to the bids, and the second, e.g., to production on the next day. The random input stems from the bids of the foreign utilities, [3].

Coming back to (1) we assume that X and Y are polyhedra, possibly involving integer requirements to vector components, such that (1) is a mixed-integer linear program under uncertainty. The mentioned two-stage dynamics of (1) becomes more explicit by the following reformulation

$$\begin{aligned} & \min_x \left\{ c^T x + \min_y \{ q(\omega)^T y : W(\omega)y = z(\omega) - T(\omega)x, y \in Y \} : x \in X \right\} \\ & = \min_x \{ c^T x + \Phi(x, \omega) : x \in X \}. \end{aligned}$$

The problem of finding a best nonanticipative decision x in (1) thus turns into finding a best member in the indexed family of random variables

$$\left(c^T x + \Phi(x, \omega) \right)_{x \in X}.$$

This leads to the topic of comparing or ranking random variables which is studied in stochastics, both conceptually and computationally, [4]. Ranking random variables according to statistical parameters reflecting mean values and/or dispersions is particularly attractive from the computational perspective. In the risk neutral case this parameter is the expectation IE , whereas in the risk averse case mean-risk models come to the fore that are based on a weighted sum of IE and a, later to be specified, parameter \mathfrak{R} reflecting risk. Denoting $f(x, \omega) := c^T x + \Phi(x, \omega)$ this paradigm leads to the following mean-risk stochastic integer program derived from (1)

$$\min \{ Q_{MR}(x) : x \in X \} \tag{2}$$

where

$$\begin{aligned} Q_{MR}(x) & := (\text{IE} + \rho \cdot \mathfrak{R})[f(x, \omega)] \\ & = \text{IE}[f(x, \omega)] + \rho \cdot \mathfrak{R}[f(x, \omega)] \\ & = Q_{\text{IE}}(x) + \rho \cdot Q_{\mathfrak{R}}(x) \end{aligned}$$

with some fixed weight $\rho > 0$. In two respects, this model goes beyond traditional linear stochastic programs: It allows for integer decision variables, and, by the term $\rho \cdot Q_{\mathfrak{R}}(x)$ it includes risk aversion into the objective. The fact that (2) is a nonlinear optimization problem carries only little weight algorithmically. Its objective is a multivariate integral whose integrand is the optimal value of another (namely the second-stage) optimization problem. This leads to insurmountable numerical difficulties when computing (with complicated probability distributions) function values or gradients of Q_{MR} . The latter, by the way, not even need to exist. Far worse, Q_{MR} is discontinuous in general, and (2) may have a finite infimum that is not attained. This may happen when \mathfrak{R} is specified as variance, see [5] for an (academic) example.

These observations indicate that the risk measure \mathfrak{R} , beside being meaningful for the practitioner, should also fulfil requirements regarding the mathematical structure it induces in Q_{MR} and the algorithmic possibilities for solving (2). Since the initial random optimization problem (1) is mixed-integer and linear, our aim is to find risk measures that enable (for discrete probability distributions) the solution of (2) by mixed-integer linear programming methodology. We will see that, for dimensionality reasons, it will not be sufficient to formulate just some mixed-integer linear program (MILP). Rather it will be important that the MILP obeys block structure amenable to decomposition or approximation. Specifications of \mathfrak{R} fulfilling these requirements are:

- *Excess Probability*,
which reflects the probability of exceeding a prescribed target level $\eta \in \mathbb{R}^s$:

$$Q_{IP^\eta}(x) := IP[\{\omega : f(x, \omega) > \eta\}].$$

- *Conditional Value-at-Risk*,
which, for a given probability level $\alpha \in]0, 1[$, reflects the expectation of the $(1 - \alpha) \cdot 100\%$ worst outcomes. There are different ways to formalize this quantity, see for instance [6]. One possibility is

$$Q_{\alpha CVaR}(x) := \min_{\eta \in \mathbb{R}} g(\eta, x), \quad (3)$$

where

$$g(\eta, x) := \eta + \frac{1}{1 - \alpha} IE[\max\{f(x, \omega) - \eta, 0\}].$$

- *Expected Excess*,
which reflects the expected value of the excess over a given target $\eta \in \mathbb{R}^s$:

$$Q_{D^\eta}(x) := IE[\max\{f(x, \omega) - \eta, 0\}].$$

- *Semideviation*,
which is similar in spirit to the Expected Excess, but with the pre-fixed target replaced by the mean:

$$Q_{D^+}(x) := IE[\max\{f(x, \omega) - Q_{IE}(x), 0\}].$$

Detailed expositions of mathematical structures and algorithms associated with specifications of (2) using these risk measures can be found in [5,7,8].

3. BLOCK STRUCTURES OF MEAN-RISK STOCHASTIC INTEGER PROGRAMS

Assume that $(q, T, W, z)(\omega)$ is discretely distributed with finitely many scenarios (q_j, T_j, W_j, z_j) and probabilities $\pi_j, j = 1, \dots, J$. It is well-known that the risk neutral model $\min\{Q_{IE}(x) : x \in X\}$ then admits the following equivalent representation as a mixed-integer linear program

$$\min \left\{ c^T x + \sum_{j=1}^J \pi_j q_j^T y_j : T_j x + W_j y_j = z_j, x \in X, y_j \in Y, j = 1, \dots, J \right\}. \quad (4)$$

The constraint matrix of (4) has the block structure depicted in Figure 1. The essential feature of this structure is that there is no constraint with two second stage variables y_{j_1} and y_{j_2} , for different scenarios $j_1, j_2 \in \{1, \dots, J\}$, in common.

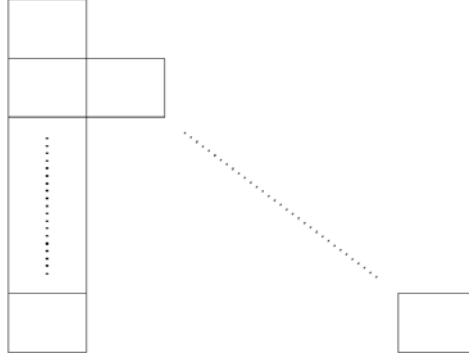


Figure 1: Block structure of constraint matrix

Second-stage variables are coupled only implicitly via x . This fact is crucial in decomposition schemes for the solution of (4). Considering the mean-risk model (2) the question arises on how MILP equivalents for different specifications of \mathfrak{R} look like and whether the principal matrix structure of Figure 1 is preserved.

Let $\mathfrak{R} := Q_{IP^\eta}$ (Excess Probability) and $\rho \geq 0$. If X is bounded it can be shown, cf. [5], that there exists a constant $M > 0$ such that (2) is equivalent to

$$\min \left\{ c^T x + \sum_{j=1}^J \pi_j q_j^T y_j + \rho \cdot \sum_{j=1}^J \pi_j \theta_j : T_j x + W_j y_j = z_j, \right. \\ \left. c^T x + q_j^T y_j - \eta \leq M \cdot \theta_j, \right. \\ \left. x \in X, y_j \in Y, \theta_j \in \{0,1\}, j = 1, \dots, J \right\}. \quad (5)$$

So the structure of Figure 1 is maintained. The specific nature of the risk measure, however, has led to additional Boolean variables θ_j that can be seen as second-stage variables.

Let $\mathfrak{R} := Q_{\alpha CVaR}$ (Conditional Value-at-Risk) and $\rho \geq 0$. Then, cf. [8], (2) is equivalent to

$$\min \left\{ c^T x + \sum_{j=1}^J \pi_j q_j^T y_j + \rho \cdot \left(\eta + \frac{1}{1-\alpha} \sum_{j=1}^J \pi_j v_j \right) : T_j x + W_j y_j = z_j, \right. \\ \left. c^T x + q_j^T y_j - \eta \leq v_j, \right. \\ \left. x \in X, \eta \in \mathbb{R}, y_j \in Y, v_j \in \mathbb{R}_+, \right. \\ \left. j = 1, \dots, J \right\}. \quad (6)$$

Here the matrix structure of Figure 1 is also maintained, but a new continuous variable η appears that can be understood as a first-stage variable. Furthermore, there are additional continuous second-stage variables v_j stemming from a resolution of a max-expression.

Let $\mathfrak{R} := Q_{D^\eta}$ (Expected Excess) and $\rho \geq 0$. Then, cf. [8], (2) is equivalent to

$$\min \left\{ c^T x + \sum_{j=1}^J \pi_j q_j^T y_j + \rho \cdot \sum_{j=1}^J \pi_j v_j : T_j x + W_j y_j = z_j, \right. \\ \left. c^T x + q_j^T y_j - \eta \leq v_j, \right. \\ \left. x \in X, y_j \in Y, v_j \in \mathbb{R}_+, j = 1, \dots, J \right\}. \quad (7)$$

This model is very similar to (4). The principal matrix structure is maintained, and only the resolution of a max-expression led to new constraints and new second-stage variables v_j .

Let $\mathfrak{R} := Q_{D^+}$ (Semideviation) and $\rho \in [0,1]$. Then, cf. [7], (2) is equivalent to

$$\begin{aligned} \min \{ & (1-\rho)c^T x + (1-\rho) \sum_{j=1}^J \pi_j q_j^T y_j + \rho \sum_{j=1}^J \pi_j v_j : T_j x + W_j y_j = z_j, \\ & c^T x + q_j^T y_j \leq v_j, \\ & c^T x + \sum_{j=1}^J \pi_j q_j^T y_j \leq v_j, \\ & x \in X, y_j \in Y, v_j \in \mathbb{R}, \\ & j = 1, \dots, J \}. \end{aligned} \quad (8)$$

Here the principal structure of Figure 1 is not maintained. The constraints

$$c^T x + \sum_{j=1}^J \pi_j q_j^T y_j \leq v_j, j = 1, \dots, J \quad (9)$$

explicitly couple second-stage variables from different scenarios. Summing up, we see that the choice of the risk measure in (2) is a sensitive issue regarding the block structures in the arising MILP equivalents. Consequences may be mild as in (7), the risk measure may be the source of additional integrality requirements as in (5), or it even may generate additional coupling in the model as in (8).

4. DECOMPOSITION AND APPROXIMATION ALGORITHMS

The mean-risk models (2) for the different specifications of \mathfrak{R} all are non-convex non-linear optimization problems in general. Even when alleviating numerical integration problems by using discrete probability distributions the problem remains to find a global minimizer in a non-convex optimization problem. Analytical properties of Q_{MR} are particularly poor when imposing a discrete μ , see [5,7,8]. For lack of smoothness (even lack of continuity), hence, local (sub-) gradient based descent approaches to minimizing Q_{MR} are not promising.

The equivalent problem formulations of the previous section provide an alternative. Although problem dimension was increased considerably, now there is the possibility of finally resorting to the well-developed algorithmic methodology of mixed-integer linear programming. Direct solution of (5) - (8) with general-purpose mixed-integer linear programming software however is prohibitive due to dimensionality reasons. Rather, the mixed-integer linear programming techniques are imbedded into lower bounding procedures of a branch-and-bound scheme.

4.1. Branch-and-Bound Framework

To solve (2) by branch-and-bound the set X is partitioned with increasing granularity. Linear inequalities are used for this partitioning to maintain the (mixed-integer) *linear* description. On the current elements of the partition upper and lower bounds for the optimal objective function value are sought. This is embedded into a coordination procedure to guide the partitioning and to prune elements due to infeasibility, optimality, or inferiority.

For the following generic description of the branch-and-bound method let \mathbb{P} denote a list of problems and $\varphi_{LB}(P)$ be a lower bound for the optimal value of $P \in \mathbb{P}$. By $\bar{\varphi}$ we denote the currently best upper bound to the optimal value of (2).

- *Step 1 (Initialization):*
Let $\mathbb{P} := (2)$ and $\bar{\varphi} := +\infty$.
- *Step 2 (Termination):*

If $P := \emptyset$ then the \bar{x} that yielded $\bar{\varphi} = Q_{MR}(\bar{x})$ is optimal.

- **Step 3 (Bounding):**
Select and delete a problem P from P . Compute a lower bound $\varphi_{LB}(P)$ and find a feasible point \bar{x} of P .
- **Step 4 (Pruning):**
If $\varphi_{LB}(P) = +\infty$ (infeasibility of a subproblem)
or
 $\varphi_{LB}(P) = Q_{MR}(\bar{x})$ (optimality of a subproblem)
or
 $\varphi_{LB}(P) > \bar{\varphi}$ (inferiority), then go to Step 2.
If $Q_{MR}(\bar{x}) < \bar{\varphi}$ then $\bar{\varphi} := Q_{MR}(\bar{x})$.
- **Step 5 (Branching):**
Create two new subproblems by partitioning the feasible set of P . Add these subproblems to P and go to Step 2.

Step 4 is to be understood accordingly if no feasible point \bar{x} was found in Step 3: Then none of the criteria applies, and the algorithm goes to Step 5.

The partitioning in Step 5, in principle, can be carried out by adding (in a proper way) arbitrary linear constraints. The most popular way though is to branch along coordinates. This means to pick a component $x_{(k)}$ of x and add the inequalities $x_{(k)} \leq a$ and $x_{(k)} \geq a + 1$, with some integer a , if $x_{(k)}$ is an integer component, or, otherwise, add $x_{(k)} \leq a - \varepsilon$ and $x_{(k)} \geq a + \varepsilon$, with some real number a , where $\varepsilon > 0$ is some tolerance parameter to avoid endless branching.

When passing the lower bound $\varphi_{LB}(P)$ of P to its two “children” created in Step 5, the difference, or gap, $\bar{\varphi} - \min_{P \in \mathcal{P}} \varphi_{LB}(P)$ provides information of the quality of the currently best solution. Step 2 often is modified by terminating in case the relative size of this gap drops below some threshold expressed in per cent.

4.2. Decomposition by Lagrangean Relaxation of Nonanticipativity

Lower bounding is a critical part in the above branch-and-bound framework. This is the place where decomposition of the relevant stochastic programs into scenario-specific subproblems will become crucial. Assume the stochastic program has the block structure of Figure 2, introduce copies $x_j, j = 1, \dots, J$ of x , and reformulate the problem equivalently by adding the constraints $x_1 = \dots = x_J$. This is nothing but formulating explicitly the nonanticipativity (NA) of x , i.e., the requirement that x has to be independent of the scenarios. The constraint matrix of the reformulated problem has the principal structure shown in Figure 2.

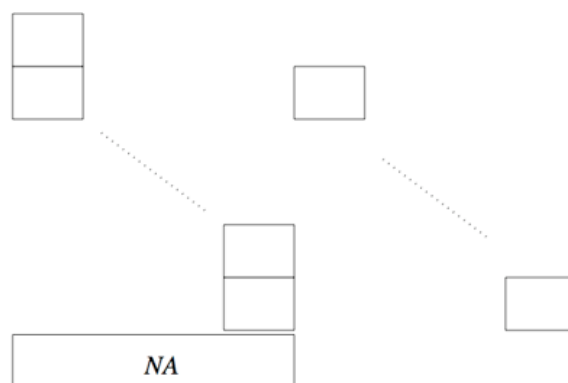


Figure 2: Block structure of reformulation

Relaxation of the NA-block, or in other words, of the nonanticipativity of the first-stage variables, hence, decomposes the constraints according to scenarios. In what follows we spell this out in detail in

terms of Lagrangean relaxation of NA for the Conditional-Value-at-Risk model (6). Recall that in (6) additionally to the decision variable x the auxiliary variable η is independent on j and, hence, can be understood as a first-stage variable that has to meet nonanticipativity. We express the nonanticipativity of x and η by the equations

$$\sum_{j=1}^J H'_j x_j = 0, \quad \sum_{j=1}^J H''_j \eta_j = 0$$

with suitable $l' \times m$ matrices H'_j and $l'' \times 1$ vectors $H''_j, j = 1, \dots, J$. The following Lagrangean function results

$$L(x, y, v, \eta, \lambda) := \sum_{j=1}^J L_j(x_j, y_j, v_j, \eta_j, \lambda)$$

where $\lambda = (\lambda', \lambda'')$ and

$$L_j(x_j, y_j, v_j, \eta_j, \lambda', \lambda'') := \pi_j (c^T x_j + q_j^T y_j + \rho \eta_j + \rho \frac{1}{1-\alpha} v_j) + \lambda'^T H'_j x_j + \lambda''^T H''_j \eta_j.$$

The Lagrangean dual reads

$$\max \{ D(\lambda) : \lambda \in \mathbb{R}^{l'+l''} \} \quad (10)$$

where

$$D(\lambda) = \min \left\{ \sum_{j=1}^J L_j(x_j, y_j, v_j, \eta_j, \lambda) : \begin{aligned} &T_j x_j + W_j y_j = z_j, \\ &c^T x_j + q_j^T y_j - \eta_j \leq v_j, \\ &x_j \in X, \eta_j \in \mathbb{R}, y_j \in Y, v_j \in \mathbb{R}_+, j = 1, \dots, J \end{aligned} \right\}.$$

This optimization problem is separable with respect to the individual scenarios, i.e.,

$$D(\lambda) = \sum_{j=1}^J D_j(\lambda)$$

where, for $j = 1, \dots, J$,

$$D_j(\lambda) = \min \left\{ L_j(x_j, y_j, v_j, \eta_j, \lambda) : \begin{aligned} &T_j x_j + W_j y_j = z_j, \\ &c^T x_j + q_j^T y_j - \eta_j \leq v_j, \\ &x_j \in X, \eta_j \in \mathbb{R}, y_j \in Y, v_j \in \mathbb{R}_+ \end{aligned} \right\}. \quad (11)$$

The Lagrangean dual (10) is a non-smooth concave maximization (or convex minimization) problem with piecewise linear objective for whose solution advanced bundle methods, see for instance [9,10], can be applied. In this way, solving the dual, or in other words, finding a desired lower bound, reduces to function value and subgradient computations for $-D(\lambda)$ (when adopting a convex minimization setting in (10)). A subgradient of $-D$ at λ is given by

$$-\left(\sum_{j=1}^J H'_j x_j^\lambda, \sum_{j=1}^J H''_j \eta_j^\lambda \right)$$

where x_j^λ and η_j^λ are the corresponding components in an optimal solution vector to the optimization problem defining $D_j(\lambda)$. As a consequence, the desired lower bound in Step 3 of the generic framework can be computed by solving the single-scenario mixed-integer linear programs in (11) instead of working with the full-size model (6). This decomposition is instrumental since (11) may be tractable for MILP solvers while (6) may not.

Results of the dual optimization also provide the basis for finding “promising” feasible points in Step 3 of the generic framework. Indeed, starting from an optimal or nearly optimal λ the components $x_j^\lambda, \eta_j^\lambda, j = 1, \dots, J$, of optimal solutions to (11) are seen as J proposals for a nonanticipative first-stage solution (x, η) . A “promising” point then is selected, for instance, by deciding for the most frequent one or by averaging and rounding to integers if necessary.

The specific nature of η as an auxiliary variable and argument in an optimization problem, see (3), allows to improve this heuristic. Instead of selecting right away a candidate for (x, η) , a candidate \bar{x} for x is fixed first, and then η is computed as the best possible value, namely as optimal solution to

$$\min\{g(\eta, \bar{x}) : \eta \in \mathbb{R}\}$$

which is equivalent to

$$\min\left\{\eta + \frac{1}{1-\alpha} \sum_{j=1}^J \pi_j v_j : c^T \bar{x} + \Phi(\bar{x}, \omega_j) - \eta \leq v_j, \right. \\ \left. \eta \in \mathbb{R}, v_j \in \mathbb{R}_+, j = 1, \dots, J\right\}.$$

The input quantities $\Phi(\bar{x}, \omega_j), j = 1, \dots, J$ are readily computed as optimal values of

$$\min\{q_j^T y : W_j y = z_j - T_j \bar{x}, y \in Y\}.$$

If $\Phi(\bar{x}, \omega_j) = +\infty$ (infeasibility of a second-stage subproblem), then \bar{x} is discarded.

This concludes the consideration of (6) as an illustration for how to achieve decomposition in the lower bounding of (5) - (7).

4.3. Approximation by Decomposable Lower Bounds

Lagrangean relaxation of nonanticipativity, of course, is possible for (8) too. However, the counterpart to (11) then is no longer separable in the scenarios due to the presence of the constraints (9). An alternative are approximate lower bounds to (8) that decompose into scenario specific subproblems after relaxation of nonanticipativity. The expectation problem $\min\{Q_{IE}(x) : x \in X\}$ is an immediate such bound, although neglecting risk effects completely. An improvement is the following relation, established in [7]. For $\eta \leq Q_{IE}(x)$ and all $\rho \in [0, 1]$ it holds that

$$Q_{IE}(x) \leq (1-\rho)Q_{IE}(x) + \rho Q_{D^n} + \rho \eta \leq Q_{IE}(x) + \rho Q_{D^+(x)}.$$

A possible specification of η fulfilling $\eta \leq Q_{IE}(x)$ is given by the wait-and-see solution of the expectation model. This is the expected value of

$$\Phi_{ws}(\omega) := \min\{c^T x + q(\omega)^T y : T(\omega)x + W(\omega)y = z(\omega), x \in X, y \in Y\}.$$

Clearly, $\Phi_{ws}(\omega) \leq f(x, \omega)$ for all $x \in X$ and all ω , such that $\eta = IE\Phi_{ws}(\omega) \leq Q_{IE}(x)$ for all $x \in X$.

When wishing to solve (8) the bounding steps in the generic branch-and-bound framework from Subsection 4.1 must be specified as follows. Lower bounds are obtained via Lagrangean relaxation of nonanticipativity applied to

$$\min \{(1 - \rho)Q_{IE}(x) + \rho Q_{D^+}(x) + \rho \eta : x \in X\} \quad (12)$$

(with X , of course, replaced by partition elements at later stages of the algorithm). Since Q_{D^+} leads to decoupled single-scenario models after relaxation of nonanticipativity, see (7), the counterpart to (11) enjoys decomposition in the scenarios. The results of the dual optimization again serve as inputs for the generation of upper bounds. Nonanticipative proposals, however, are not inserted into the objective of (12) but into the original objective $Q_{IE}(x) + \rho Q_{D^+}(x)$.

5. DISPERSED GENERATION - NUMERICAL EXPERIMENTS

To illustrate models and algorithms of the previous sections let us consider a system with dispersed generation (DG) of power and heat. The DG system, run by a German utility, consists of 5 engine-based cogeneration stations, each equipped with a thermal storage and a cooling device, and altogether involving 8 gas boilers, 9 gas motors and one gas turbine. The latter are combined heat and power (CHP) units. The DG system is completed by 12 wind turbines and one hydroelectric power plant. Power is fed into a global grid, enabling electricity trade at energy markets. Heat is fed into local networks around each cogeneration station.

Uncertainty is present both at the input and the output sides of the system. At the input side, this concerns the infeed from the wind turbines. At the output side, power and heat demand as well as power prices typically are prone to uncertainty.

Taking into account all relevant operational constraints, such as minimum and maximum production levels for the DG units, bounds for the fill of the storage, demand fulfilment, fuel consumption and minimum up times of the units, the DG system is modeled as a mixed-integer linear program, with on-off decisions of units as the major sources of integrality. The objective is to minimize fuel costs (minus revenues from trading). The planning horizon is 24 hours, with a discretization into quarter-hourly subintervals. As a random optimization problem, or in other words, a specification of (1), the model has roughly 17.500 variables (9.000 Boolean, 8.500 continuous) and 22.000 constraints. For a detailed description of the model see [11].

With standard software for mixed-integer programs such as ILOG-CPLEX [12] instances of the random optimization problem with stochastic entities set to fixed values can be solved on a Linux-PC with a 3,0 GHz Pentium processor and 2,0 GB RAM in less then 20 seconds and with gaps of less than 0,1 %.

As a next step we impose the information constraint that the random data of the model are known with certainty in the first four hours of the time horizon, and are uncertain from then on. This leads to a two-stage planning problem under uncertainty where the first-stage variables are the decisions from the initial 16 quarter-hourly intervals, and the second-stage is given by the remaining variables.

Scenarios	Boolean variables	Continuous variables	Constraints
1	8.959	8.453	22.196
5	38.719	36.613	96.084
10	75.919	71.813	188.444
50	373.519	353.413	927.324

Table 1: Dimension of a risk neutral model

The specification of the risk neutral expectation-based model (4) then no longer is tractable in acceptable time by ILOG-CPLEX. Table 1 shows problem dimensions, and Table 2 computational results.

The stopping criterion has been a gap of less than 0,01 %. While, for the 50-scenario instance, the decomposition method from Subsections 4.1 and 4.2 reaches this gap in roughly 10 minutes, ILOG-CPLEX did not succeed in 24 hours.

Number of scenarios	Decomposition		
	Objective	Gap (%)	Time (sec.)
5	7.150.903	0,0067	8
10	5.987.578	0,0087	52
20	5.837.559	0,0088	172
30	5.928.542	0,0093	364
40	5.833.191	0,0093	487
50	5.772.268	0,0094	611
Number of scenarios	ILOG-CPLEX		
	Objective	Gap (%)	Time (sec.)
5	7.150.945	0,0072	11
10	5.987.642	0,0099	253
20	5.837.584	0,0093	1.769
30	5.928.500	0,0091	9.486
40	5.833.156	0,0107	26.286
50	5.772.335	0,0129	22.979

Table 2: Solution times for the risk neutral model

The difficulty of finding optimal nonanticipative decisions by means of scenario analysis is illustrated in Table 3. For the heat production of the boilers in one cogeneration station, we have listed an optimal nonanticipative solution (of a risk neutral stochastic program) against optimal solutions for different single-scenario models. Clearly, the optimal nonanticipative solution, displayed in the last row of the table, cannot simply be deduced from the collection of the anticipative first-stage decisions in the rows above.

Scenario number	Output of the boilers in station 1 in time interval				
	12	13	14	15	16
1	3000	3270	4768	6500	3000
2	4829	6500	3000	5652	6500
3	4810	4770	4731	4754	4764
4	4810	4770	4731	5333	6500
5	4810	4770	4731	6500	6500
6	4810	4770	6144	6500	6500
7	4810	4770	4731	6009	6500
8	4810	4770	5076	6500	3000
9	3000	3270	4768	6500	3000
10	3000	6500	6500	3000	3089
Optimal	6500	3000	3188	4754	4764

Table 3: Dispersion of first-stage decisions

The following results for mean-risk models incorporating Expected Excess, Excess Probability, or Conditional Value-at-Risk show some effects of including risk aversion and confirm the superiority of decomposition over standard mixed-integer linear programming solvers.

Table 4 shows contributions to the optimal objective values of the individual scenarios, computed with a mean-risk model involving Expected Excess and varying risk weights ρ . The target value η was set to 6.807.892. It becomes apparent, that, with increasing weight, optimal values are getting above or near target for an increasing number of scenarios.

Scenario number	Weight for risk measure			
	0,0001	20	10.000	infinity
1	5.828.840	5.828.850	5.828.800	6.807.550
2	6.808.230	6.808.130	6.807.950	6.807.980
3	7.787.410	7.787.390	7.787.320	7.787.300
4	8.766.730	8.766.730	8.766.710	8.766.690
5	9.746.580	9.746.570	9.746.550	9.746.530
6	10.726.500	10.726.500	10.726.400	10.726.400
7	4.849.210	4.849.630	4.849.290	6.697.350
8	3.870.510	3.870.780	3.870.900	6.447.340
9	2.893.830	2.894.010	2.894.700	6.389.640
10	1.948.510	1.949.060	1.949.900	6.074.710
Q_{IE}	5.987.630	5.987.730	5.987.740	7.185.900
$Q_{\mathcal{R}}$	489.812	489.795	489.763	489.762

Table 4: Optimal solutions for different risk weights

Computations for a model using the Conditional Value-at-Risk as the objective are reported in Table 5. For instances with 20 and more scenarios the decomposition approach reaches the required gap of 0,005 % in less computational time than ILOG-CPLEX does. Appropriate solutions are found up to 8 times faster with the decomposition method.

Number of scenarios	Decomposition		
	$Q_{\mathcal{R}}$	Gap (%)	Time (sec.)
5	9.256.642	0,0010	37
10	9.060.737	0,0013	79
20	8.522.057	0,0023	436
30	8.996.822	0,0033	594
40	8.795.064	0,0039	1.038
50	8.557.755	0,0039	1.286
Number of scenarios	ILOG-CPLEX		
	$Q_{\mathcal{R}}$	Gap (%)	Time (sec.)
5	9.256.610	0,0007	23
10	9.060.910	0,0033	67
20	8.522.190	0,0040	504
30	8.996.950	0,0049	5.395
40	8.795.120	0,0049	7.366
50	8.557.813	0,0050	9.685

Table 5: Results for the Conditional Value-at-Risk

Table 6 and Table 7 show results for a mean-risk model involving Excess Probability. Here, time limits increasing with the number of scenarios were used as stopping criteria. The tables illustrate, that for instances with more than 10 scenarios the decomposition method finds solutions with smaller gaps

than ILOG-CPLEX does. For instances with 20 and more scenarios ILOG-CPLEX fails to reach a feasible solution at all.

Number of scenarios	Time (sec.)	ρ	Q_{IE}	Q_{IR}	Gap (%)
5	300	0	7.150.866	–	0,0062
		0,0001	7.150.913	0,400	0,0066
		10.000	7.150.930	0,400	0,0348
10	450	0	5.987.526	–	0,0078
		0,0001	5.987.554	0,370	0,0084
		10.000	5.987.556	0,370	0,0302
20	600	0	5.837.500	–	0,0078
		0,0001	5.837.560	0,205	0,0089
		10.000	5.837.556	0,205	0,0183
50	1500	0	5.772.226	–	0,0087
		0,0001	5.772.271	0,205	0,0096
		10.000	5.772.287	0,205	0,0151

Table 6: Results for the Excess Probability using the Decomposition

Number of scenarios	Time (sec.)	ρ	Q_{IE}	Q_{IR}	Gap (%)
5	300	0	7.150.850	–	0,0057
		0,0001	7.150.900	0,400	0,0064
		10.000	7.150.880	0,400	0,0342
10	450	0	5.987.670	–	0,0104
		0,0001	5.987.630	0,370	0,0098
		10.000	5.987.630	0,370	0,0315
20	600	0	infeasible		infinity
		0,0001	infeasible		infinity
		10.000	infeasible		infinity
50	1500	0	infeasible		infinity
		0,0001	infeasible		infinity
		10.000	infeasible		infinity

Table 7: Results for the Excess Probability using ILOG-CPLEX

6. CONCLUSION

Traditional risk neutral stochastic integer programs can be extended into mean-risk models by the help of risk measures such as Excess Probability, Conditional Value-at-Risk, Expected Excess, or Semideviation.

These risk measures induce sound mathematical structures and allow for numerical treatment of the resulting mean-risk models in the framework of mixed-integer linear programming.

For the latter, block structures arising within equivalent MILPs are essential. These structures depend on the risk measures selected and, either, enable a direct decomposition via Lagrangean relaxation of nonanticipativity, or, require approximation by decomposable models.

The operation of energy systems with decentralized generation of power and heat and with infeed from renewable resources is an activity whose optimization is inherently infected by uncertainty. Mean-risk models of the mentioned types are proper tools for optimizing nonanticipative decisions in this context. The introduced decomposition algorithms are computationally promising and outperform general purpose mixed-integer linear programming solvers at this class of problems.

7. ACKNOWLEDGEMENT

This research has been supported by the German Federal Ministry for Education and Research under grant 03SCNIVG.

REFERENCES

- RUSZCZNYSKI, A; SHAPIRO, A. (EDS.) (2003): **Handbooks in Operations Research and Management Science, 10: Stochastic Programming**, Elsevier, Amsterdam.
- VERHOEVEN, M. (2005): **Optimierung der Konfiguration von dezentralen Energieumwandlungsanlagen mit Kraft-Wärme-Kopplung**, Master Thesis, Department of Mathematics, University of Duisburg- Essen.
- NOWAK, M.P.; SCHULTZ, R.; WESTPHALEN, M. (2005): A stochastic integer programming model for incorporating day-ahead trading of electricity into hydro-thermal unit commitment, **Optimization and Engineering 6**, 163 - 176.
- MÜLLER, A.; STOYAN, D. (2002): **Comparison Methods for Stochastic Models and Risks**, Wiley, Chichester.
- SCHULTZ, R.; TIEDEMANN, S. (2003): Risk Aversion via Excess Probabilities in Stochastic Programs with Mixed-Integer Recourse, **SIAM Journal on Optimization 14**, 115 -138.
- ROCKAFELLAR, R.T.; URYASEV, S. (2003): Conditional Value-at-Risk for General Loss Distributions, **Journal of Banking & Finance 26**, 1443 -1471.
- MÄRKERT, A.; SCHULTZ, R. (2005): On deviation measures in stochastic integer programming, **Operations Research Letters 33**, 441-449.
- SCHULTZ, R., TIEDEMANN, S. (2006): Conditional value-at-risk in stochastic programs with mixed-integer recourse, **Mathematical Programming 105**, 365-386.
- HELMBERG, C.; KIWIEL, K.C. (2002): A spectral bundle method with bounds, **Mathematical Programming 93**, 173-194.
- KIWIEL, K.C. (1990): Proximity control in bundle methods for convex nondifferentiable optimization, **Mathematical Programming 46**, 105-122.
- HANDSCHIN, E.; NEISE, F.; NEUMANN, H.; SCHULTZ, R. (2005): Optimal operation of dispersed generation under uncertainty using mathematical programming, **Preprint 594-2005, Dep. Of Mathematics, University of Duisburg-Essen, and presented at PSCC 05, Liege/Belgium, 2005.**
- CPLEX CALLABLE LIBRARY 8.1 (2003): **ILOG, 2003.**

Received June 2006
Revised January 2007