Rutcor
Research
Report

Algorithms for handling CVaR-constraints in dynamic stochastic programming models with applications to finance

Csaba I. Fábián[a]   Anna Veszprémi[b]

[a]Department of Operations Research, Loránd Eötvös University of Budapest, and Kecskemét College. Postal address: P.O.Box 120, Budapest, H-1518 Hungary. E-mail: fabian@cs.elte.hu. This research has been supported by OTKA, Hungarian National Fund for Scientific Research, project No. 47340.

[b]Department of Algorithms and their Applications, Loránd Eötvös University of Budapest.

ALGORITHMS FOR HANDLING CVAR-CONSTRAINTS IN DYNAMIC STOCHASTIC PROGRAMMING MODELS WITH APPLICATIONS TO FINANCE

Csaba I. Fábián   Anna Veszprémi

Abstract. We propose dual decomposition and solution schemes for multistage CVaR-constrained problems. These schemes meet the need for handling multiple CVaR-constraints for different time frames and at different confidence levels. Hence they allow shaping distributions according to the decision maker's preferences.

With minor modifications, the proposed schemes can be used to decompose further types of risk constraints in dynamic portfolio management problems. We consider integrated chance constraints, second-order stochastic dominance constraints, and constraints involving a special value-of-information risk measure. We also suggest application to further financial problems. We propose a dynamic risk-constrained optimization model for option pricing. Moreover we propose special mid-term constraints for use in asset-liability management.

1   Introduction

Value-at-Risk (VaR) is a widely accepted risk measure. Using stochastic programming terminology, VaR constraints can be formulated as chance constraints or probabilistic constraints. These constraints were introduced and studied by Charnes et al. (1958) and by Prékopa (1973). Portfolio optimization problems with VaR objectives or constraints are generally hard to solve due to non-convexity. A characteristic cause of difficulty is that the technological coefficients are random since they represent returns of the different assets.

An alternative risk measure, namely *Conditional Value-at-Risk* (CVaR), has been proposed by Rockafellar and Uryasev (2000). They derived a representation of CVaR as the optimum of a special minimization problem. This representation makes CVaR tractable in optimization problems. An overview of VaR- and CVaR-optimization models and methods can be found in Prékopa (2003) and Kall and Mayer (2005).

Andersson, Mausser, Rosen, and Uryasev (2001) examined single-stage models for credit risk optimization. In one model type they minimized CVaR under a constraint on expected return, in another they constructed the risk/return efficient frontier. They demonstrated that diverse risk measures (including VaR) can be effectively controlled through CVaR optimization.

Krokhmal, Palmquist, and Uryasev (2002) and Rockafellar and Uryasev (2002) introduced the idea of using CVaR in dynamic models. Krokhmal, Palmquist, and Uryasev (2002) is also the first paper dealing with CVaR constraints. The authors observe that portfolio optimization with multiple CVaR-constraints for different time frames and at different confidence levels allows the shaping of distributions according to the decision maker's preferences. Claessens and Kreuser (2004) utilize these ideas in an asset/liability management tool.

Topaloglou (2004) and Topaloglou, Vladimirou, Zenios (2006) developed elaborate multistage financial models. The objective was minimization of end-of-horizon CVaR under a constraint on expected return.

Roman, Mitra, and Darby-Dowman (2006) proposed and studied portfolio optimization models in which variance is minimized under constraints on expected yield and CVaR. This approach represents a compromise between regulators' requirements for short tails and classical fund managers' requirements for small variance. The authors construct an approximation of the efficient frontier. In making the final choice, the decision maker plays a key role.

According to our knowledge, the solution methods applied in the above mentioned projects do not exploit the special structure of CVaR-optimization problems. Research in this direction started recently:

Künzi-Bay and Mayer (2006) proposed a polyhedral representation of CVaR, and on the basis of this, developed a special method for the minimization of CVaR in one-stage stochastic problems. They implemented the method and their experimental results show the clear superiority of their approach over general-purpose methods.

Ahmed (2006) examined the complexity of mean/risk stochastic programming under different risk measures. He proved that the problem is tractable with Quantile-deviation as

risk measure, and hence with CVaR as a special case. He proposed a decomposition scheme and a parametric cutting-plane algorithm to generate the efficient frontier. His approach focuses on classic two-stage models having the decision/observation/decision pattern.

Fábián (2007) proposed a decomposition framework for two-stage CVaR minimization and CVaR constrained problems having the decision/observation/decision/observation pattern. The decomposition scheme is based on the Künzi-Bay – Mayer polyhedral representation. The solution methods are special Level-type methods, variants of the Level Method of Lemaréchal, Nemirovskii, and Nesterov (1995). This scheme is effective only for the handling of a limited number of CVaR-constraints. (The number of the CVaR-constraints should be significantly less then the number of the first-stage scenarios.)

In this paper we propose a dual decomposition scheme and a solution method that can handle problems having a large number of CVaR-constraints. This scheme meets the need for handling simultaneous CVaR-constraints for different time frames and at different confidence levels.

The process consists of two steps. In the first step we construct an approximation of the efficient frontier. This helps the decision maker in calibrating the right-hand-side parameters of the CVaR constraints. In course of the first step we only need near-optimal dual solutions.

Once the parameters are set to the satisfaction of the decision maker, we pass to the second step, and find a primal optimal solution. Like Roman et al., we include variance terms in the objective function. Besides the economic advantages observed by Roman et al., the variance terms yield a technical advantage: They make the objective function strictly convex, and hence enable solution of the primal problem through a dual approach.

With minor modifications, the proposed scheme can be used to decompose further risk constraints in dynamic portfolio management problems. We describe adaptations to

– Integrated Chance Constraints. A generic ICC was formulated by Klein Haneveld et al. (2005).

– Second-order Stochastic Dominance constraints. A single-stage portfolio optimization model involving an SSD constraint was proposed by Dentcheva and Ruszczyński (2006). We propose multistage generalization of the SSD-constrained model.

– constraints involving a special value-of-information risk measure proposed by Pflug (2006).

We suggest application to further financial problems:

– Option pricing. Ryabchenko et al. (2004) propose a grid-based replication model for the pricing of a European option in an incomplete market. We start from this method as a basis but propose a traditional scenario-tree approach with risk minimization.

– Asset-liability management models with special mid-term constraints were examined by Drijver et al. (2002) and Klein Haneveld et al. (2005). We propose another type of mid-term constraints that serves similar purpose but is easier to handle.

The paper is organized as follows: In Section 1.1 we cite the representation result of Rockafellar and Uryasev. In Section 1.2 we compare embedding and cutting-plane representations of CVaR. In 1.3 we present static prototype models.

In Section 2 we present decomposition and solution schemes for the minimization of CVaR in dynamic stochastic models. These are generalizations of the two-stage schemes proposed by Fábián (2007).

In Section 3 we present a dual decomposition scheme that can handle a large number of CVaR-constraints.

The results of the paper are summarized and potential fields of application are suggested in Section 4.

1.1   Conditional value-at-risk

Let us consider a one-period financial investment.

w denotes the total wealth at the end of the examined period. This is a random variable.

$w^{\mathcal{B}}$ denotes a benchmark for end-of-period wealth (i.e., the wealth that we intend to accumulate by the end of the examined period). We assume it is a parameter that has been set by the decision maker.

Then the loss relative to the benchmark can be expressed as: $w^{\mathcal{B}} - w$. Given a probability α, a heuristic definition of the risk measures is the following.

α-*Value-at-Risk* (VaR) answers the question: what is the maximum loss with the confidence level $\alpha * 100\%$?

α-*Conditional-Value-at-Risk* (CVaR) is the (conditional) mean value of the worst $(1 - \alpha) * 100\%$ losses.

Rockafellar and Uryasev (2000) proved that α-VaR and α-CVaR can be computed through the solution of the following problem:

$$\min_{z \in \mathbb{R}} \ z + \frac{1}{1 - \alpha} \mathrm{E} \left([w^{\mathcal{B}} - w - z]_+ \right). \tag{1}$$

α-VaR is the optimal value of z in case the problem has a unique solution. Generally, the set of the optimal solutions of (1) is a closed interval whose left endpoint is α-VaR.

In this paper we assume that w has a discrete distribution. Let the realizations be $w^{(1)}, \ldots, w^{(N)}$ with probabilities $p_1, \ldots, p_N$, respectively. Problem (1) takes the form

$$\min_{z \in \mathbb{R}} \ z + \frac{1}{1 - \alpha} \sum_{i=1}^{N} p_i \left[w^{\mathcal{B}} - w^{(i)} - z \right]_+. \tag{2}$$

Rockafellar and Uryasev (2002) observe that problem (2) can be solved by just sorting the values $w^{(i)}$.

1.2 CVaR minimization: embedding and cutting-plane representations

The end-of-period wealth w will be the yield of a *portfolio* selected by the decision maker at the beginning of the examined period. Assume there are n assets.

$\boldsymbol{x} \in \mathbb{R}^n$ represents a portfolio (i.e., the amounts of money invested in the different assets).

The sum of the components of $\boldsymbol{x}$ should be equal to the initial capital that we denote by w_0. We will formalize this constraint as $\mathbf{1} \cdot \boldsymbol{x} = w_0$. By $\mathbf{1}$ we will denote a vector having 1 in each components ; and the scalar product of vectors will be denoted by a dot.

We may impose further constraints on the portfolios (e.g., we may prescribe proportions between different positions or impose lower/upper bounds on certain positions). Let $X \subset \mathbb{R}^n$ denote the set of the feasible portfolios.

$\boldsymbol{r}$ denotes the returns for the different assets. This is an n-dimensional random vector.

Assume there are N realizations, $\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(N)}$, occurring with probabilities $p_1, \ldots, p_N$, respectively.

The end-of-period total wealth w is computed by the scalar product $\boldsymbol{r} \cdot \boldsymbol{x}$. This is a random variable with realizations $w^{(i)} = \boldsymbol{r}^{(i)} \cdot \boldsymbol{x}$.

In this paper we assume that no short positions are allowed, i.e., $\boldsymbol{x} \geq \mathbf{0}$ is imposed among the constraints of X. We assume moreover that X is determined by a *homogeneous* system of linear inequalities, i.e., it prescribes bounds on proportions between different positions. We also assume that the returns are non-negative, and that the returns of different assets are linearly independent. (These are technical assumptions required by the convergence proof of the proposed method. The decomposition scheme works without these assumptions, and the method can be implemented under milder assumptions.)

Suppose we need a feasible portfolio that minimizes risk. Using (2), this problem can be written as

$$\min \ z + \frac{1}{1-\alpha} \sum_{i=1}^{N} p_i \left[w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z \right]_+ \qquad \text{such that} \quad z \in \mathbb{R}, \ \ \boldsymbol{x} \in X, \ \mathbf{1} \cdot \boldsymbol{x} = w_0. \qquad (3)$$

Remark 1 *It is easily seen that the setting of the parameter $w^{\mathcal{B}}$ does not affect the optimal value of the variables $\boldsymbol{x}$ in the above problem. (Changes in $w^{\mathcal{B}}$ are compensated by equal changes in the optimal value of z.) From a purely mathematical point of view, we can set $w^{\mathcal{B}} = 0$.*

Rockafellar and Uryasev (2000) propose transforming (3) into a linear programming problem by introducing new variables y_i to represent $\left[w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z \right]_+$ $(i = 1, \ldots, N)$:

$$\min \ z + \frac{1}{1-\alpha} \sum_{i=1}^{N} p_i y_i \qquad \text{such that} \quad z \in \mathbb{R}, \ \ \boldsymbol{x} \in X, \ \mathbf{1} \cdot \boldsymbol{x} = w_0,$$
$$y_i \geq 0, \quad y_i \geq w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z, \quad (i = 1, \ldots, N). \qquad (4)$$

From a geometric point of view, the feasible set of (3) is embedded into a higher dimensional space. We will say that the linear programming problem (4) is a *cutting-plane representation* for the convex problem (3).

Künzi-Bay and Mayer (2006) consider (4) as a two-stage problem: z is the first-stage variable, and y_i are the second-stage variables. Based on this approach, they derive the following equivalent formulation of (3) :

$$\min \ z + \tfrac{1}{1-\alpha}v \qquad \text{such that} \quad z, v \in \mathbb{R}, \ \ \boldsymbol{x} \in X, \ \boldsymbol{1} \cdot \boldsymbol{x} = w_0,$$

$$\sum_{i \in \mathcal{I}} p_i \left(w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z \right) \ \leq v \quad (\, \mathcal{I} \subset \{1, \dots, N\} \,). \tag{5}$$

A subset $\mathcal{I} \subset \{1, \dots, N\}$ represents the possibility that $w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z$ is positive for $i \in \mathcal{I}$; and non-positive for $i \notin \mathcal{I}$. The constraint belonging to $\mathcal{I} = \emptyset$ imposes non-negativity of v.

We will say that the linear programming problem (5) is a *cutting-plane representation* for the convex problem (3). There is an analogy between the cutting-plane representation of CVaR proposed by Künzi-Bay and Mayer ; and the cutting-plane representation of *integrated chance constraints* proposed by Klein Haneveld and Van der Vlerk (2006). Independent test results of Klein Haneveld and Van der Vlerk, and of Künzi-Bay and Mayer show that in case of large problems, cutting-plane representations are more effectively solved than embedding representations.

1.3   A static prototype model

We consider a mean/risk portfolio-optimization model, where risk is measured by Conditional-Value-at-Risk $\mathrm{CVaR}(w)$. The customary objective is

$$\max \ \mathrm{E}(w) \ - \lambda \, \mathrm{CVaR}(w), \tag{6}$$

where the parameter $\lambda > 0$ measures *risk aversion* of the decision maker. (We assume a known fixed λ.)

The Künzi-Bay – Mayer formulation (5) yields the following polyhedral representation of this problem:

$$\max \ \mathrm{E}(\boldsymbol{r}) \cdot \boldsymbol{x} \ - \lambda \left(z + \tfrac{1}{1-\alpha}v \right)$$

$$\text{such that} \ \ \boldsymbol{x} \in X, \ \boldsymbol{1} \cdot \boldsymbol{x} = w_0, \quad z, v \in \mathbb{R},$$

$$\sum_{i \in \mathcal{I}} p_i \left(w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z \right) \ \leq v \quad (\, \mathcal{I} \subset \{1, \dots, N\} \,). \tag{7}$$

Künzi-Bay and Mayer (2006) propose a special cutting-plane method for such problems: The constraint belonging to a subset $\mathcal{I}$ is interpreted as a cut. A model problem is constructed that includes only certain cuts. The model problem is solved ; and a new cut is generated taking into account the optimal solution. Namely, the cut that is deepest at the optimal

solution, i.e., a linear support function of the CVaR function at the optimal solution. It is easily constructed in the present case.

This process is iterated while the depth of the currently added cut is significant. Künzi-Bay and Mayer observe that from a purely mathematical point of view, their method can be considered as a version of the general method of Klein Haneveld and Van der Vlerk (2006), as specialized for CVaR minimization.

Künzi-Bay and Mayer implemented their method, and solved several CVaR problems with their specialized solver called CVaRMin. Their experimental results clearly demonstrate the superiority of CVaRMin over general-purpose linear or stochastic problem solvers. For the largest test problems, CVaRMin was by at least one order of magnitude faster then either of the other solvers involved.

Roman, Mitra, and Darby-Dowman (2006) proposed and studied portfolio optimization models in which variance is minimized under constraints on expected yield and CVaR. The authors tested the model on real-life data. Several levels of expected yield were considered; and for each level, five portfolios were selected from the mean/variance/CVaR efficient frontier: the minimum variance portfolio, the minimum CVaR portfolio, and three intermediate ones. The selected portfolios were then tested. Both in-sample and out-of-sample analysis shows that the performance of intermediate portfolios are superior.

After Roman et al. we include a variance term in the objective:

$$\max \ \mathrm{E}(w) \ - \varrho \, \mathrm{V}(w) \ - \lambda \, \mathrm{CVaR}(w), \tag{8}$$

where ϱ is a positive constant, and $\mathrm{V}(w) := \mathrm{E}\left(\left[w - \mathrm{E}(w)\right]^2\right)$ is the variance of w. Besides the economic advantages observed by Roman et al., the variance terms yield a technical advantage: They make the objective function strictly convex under the present assumption of independent asset returns.

It is well known and easy to prove that

$$\mathrm{V}(w) = \min_{w^{\mathcal{E}} \in \mathbb{R}} \ \mathrm{E}\left(\left[w - w^{\mathcal{E}}\right]^2\right),$$

and the optimal value of the variable $w^{\mathcal{E}}$ is $\mathrm{E}(w)$. Hence the problem can written in the form

$$\max \ \mathrm{E}(\boldsymbol{r}) \cdot \boldsymbol{x} \ - \varrho \, \mathrm{E}\left(\left[\boldsymbol{r} \cdot \boldsymbol{x} - w^{\mathcal{E}}\right]^2\right) \ - \lambda \left(z + \tfrac{1}{1-\alpha} v\right)$$

$$\text{such that } \ \boldsymbol{x} \in X, \ \mathbf{1} \cdot \boldsymbol{x} = w_0, \quad w^{\mathcal{E}}, z, v \in \mathbb{R}, \tag{9}$$

$$\sum_{i \in \mathcal{I}} p_i \left(w^{\mathcal{B}} - \boldsymbol{r}^{(i)} \cdot \boldsymbol{x} - z\right) \ \leq v \quad (\ \mathcal{I} \subset \{1, \ldots, N\} \ ).$$

Let us introduce the $(n+1) \times (n+1)$ symmetric matrix $\overline{R} := \mathrm{E}\left((\boldsymbol{r}, -1)(\boldsymbol{r}, -1)^T\right)$, where $(\boldsymbol{r}, -1)$ denotes the $(n+1)$-vector obtained by concatenating random $\boldsymbol{r}$ vector and the number

-1. The quadratic term in the objective of (9) can be written as $\mathrm{E}\left(\left[\boldsymbol{r}\cdot\boldsymbol{x}-w^{\mathcal{E}}\right]^2\right)=$ $\left(\boldsymbol{x},w^{\mathcal{E}}\right)^T\overline{R}\left(\boldsymbol{x},w^{\mathcal{E}}\right)$. This is obviously a convex fucntion of $\left(\boldsymbol{x},w^{\mathcal{E}}\right)$, and a support function at $\left(\hat{\boldsymbol{x}},\hat{w}^{\mathcal{E}}\right)$ is

$$l\left(\boldsymbol{x},w^{\mathcal{E}}\right):=\left(\hat{\boldsymbol{x}},\hat{w}^{\mathcal{E}}\right)^T\overline{R}\left(\hat{\boldsymbol{x}},\hat{w}^{\mathcal{E}}\right)+2\left(\hat{\boldsymbol{x}},\hat{w}^{\mathcal{E}}\right)^T\overline{R}\left(\boldsymbol{x}-\hat{\boldsymbol{x}},\ w^{\mathcal{E}}-\hat{w}^{\mathcal{E}}\right).$$

Problem (9) can be solved by either a cutting-plane method or a bundle-type method ; support functions to the objective function can be constructed by aggregating respective support functions to the variance and CVaR terms.

   For the solution of the prototype problem we propose using the Level Method of Lemaréchal, Nemirovskii, and Nesterov (1995). This is an iterative method. A *cutting-plane model* of the objective function is maintained: this is the upper cover of the linear support functions drawn at the known iterates. Level sets of this model function are used for regularization. The next iterate is obtained by projecting the current iterate onto a certain level set of the current model function.

   Lemaréchal et al. report on successful application of the method to a variety of problems, including problems having objective functions like (8) ; sums of piecewise linear functions and of quadratic functions. Lemaréchal et al. observe that the method was comparable to the best known methods.

2   A dynamic prototype model

Topaloglou (2004) and Topaloglou, Vladimirou, Zenios (2006) developed elaborate multistage financial models. The objective is minimizing end-of-horizon CVaR under a constraint on expected end-of-horizon return. (These models take into account transaction costs and hedging possibilities.) The authors constructed the respective efficient frontiers of one- and two-stage problems. The two-stage risk/return profile is clearly dominating, and the difference grows with increasing targets of expected return. Moreover, the authors made backtesting with real market data, solving consecutive problems in a rolling horizon manner. (The single-stage model had a horizon of one month, and the two-stage model two months. In the two-stage case, the optimal first-stage portfolios were evaluated in the rolling horizon scheme.) The performance of the two-stage model proved superior in each examined point of time. The differences are more evident when higher target returns are imposed, forcing the selection of riskier portfolios. (Examining optimal solutions reveals that the two-stage model produces more diversified portfolios.)

   From an algorithmic point of view, these financial problems fit the prototype to be described in this section. Hence the special decomposition framework and solution method we propose can also be applied to these financial problems.

   In order to keep notation as simple as possible we describe a three-stage model that easily generalizes to further stages. Initial capital will be denoted by w_0. This is a given parameter. For time periods $t=1,2,3$, we will use the following notation:

$\boldsymbol{x}_t$ represents the portfolio selected at the beginning of the t-th time period. This is a decision vector. Feasible portfolios are represented by $\boldsymbol{x}_t \in X$, $\mathbf{1} \cdot \boldsymbol{x}_t = w_{t-1}$.

$\boldsymbol{r}_t$ denotes the random return occurring in course of the t-th time period.

$w_t := \boldsymbol{r}_t \cdot \boldsymbol{x}_t$ denotes the wealth at the end of the t-th time period.

Description of the random parameters.   Assume that the (discrete finite) joint distribution of the random vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}_3$ is known:

The realizations of $\boldsymbol{r}_1$ are $\boldsymbol{r}_1^{(i)}$ $(1 \leq i \leq N_1)$, occurring with respective probabilities $p_1^{(i)}$.

For $1 \leq i \leq N_1$, the realizations of $\boldsymbol{r}_2$ given that $\boldsymbol{r}_1 = \boldsymbol{r}_1^{(i)}$

are $\boldsymbol{r}_2^{(ij)}$ $(1 \leq j \leq N_2^{(i)})$, occurring with respective probabilities $p_2^{(ij)}$.

For $1 \leq i \leq N_1$, $1 \leq j \leq N_2^{(i)}$, the realizations of $\boldsymbol{r}_3$ given that $\boldsymbol{r}_1 = \boldsymbol{r}_1^{(i)}$ and $\boldsymbol{r}_2 = \boldsymbol{r}_2^{(ij)}$

are $\boldsymbol{r}_3^{(ijk)}$ $(1 \leq k \leq N_3^{(ij)})$, occurring with respective probabilities $p_3^{(ijk)}$.

We introduce the following index sets to identify events and scenarios:

$$\mathcal{S}_1 := \{1, \ldots, N_1\}, \quad \mathcal{S}_2 := \left\{ (i,j) \,\middle|\, i \in \mathcal{S}_1,\ 1 \leq j \leq N_2^{(i)} \right\}, \quad \mathcal{S}_3 := \left\{ (i,j,k) \,\middle|\, (i,j) \in \mathcal{S}_2,\ 1 \leq k \leq N_3^{(ij)} \right\}$$

E.g., $(i,j,k) \in \mathcal{S}_3$ is the scenario that the first-, second-, and third-period returns $\boldsymbol{r}_1^{(i)}$, $\boldsymbol{r}_2^{(ij)}$, and $\boldsymbol{r}_3^{(ijk)}$ realize. The probability of this scenario will be denoted by

$$p_{(ijk)} := p_1^{(i)} p_2^{(ij)} p_3^{(ijk)}.$$

Like in the static case, we assume that the returns are non-negative, and that the returns of different assets are linearly independent.

2.1   Problem formulation

We try to find a balance between end-of-horizon expected return, variance, and Conditional-Value-at-Risk. Denoting end-of-horizon wealth by $w := w_3$, the problem can be formulated as (8). We assume known fixed parameters ϱ and λ.

Using the Künzi-Bay – Mayer cutting-plane representation of CVaR, the problem can be written as

$$\max \quad \sum_{(i,j,k)\in\mathcal{S}_3} p_{(ijk)}\, \boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3^{(ij)} \quad -\varrho \sum_{(i,j,k)\in\mathcal{S}_3} p_{(ijk)} \left[\boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3^{(ij)} - w^{\mathcal{E}} \right]^2 \quad - \lambda \left(z + \tfrac{1}{1-\alpha}v \right)$$

such that $\quad \boldsymbol{x}_1 \in X, \quad \mathbf{1} \cdot \boldsymbol{x}_1 = w_0,$

$$\boldsymbol{x}_2^{(i)} \in X, \quad \mathbf{1} \cdot \boldsymbol{x}_2^{(i)} \;=\; \boldsymbol{r}_1^{(i)} \cdot \boldsymbol{x}_1 \qquad (\, i \in \mathcal{S}_1 \,),$$

$$\boldsymbol{x}_3^{(ij)} \in X, \quad \mathbf{1} \cdot \boldsymbol{x}_3^{(ij)} \;=\; \boldsymbol{r}_2^{(ij)} \cdot \boldsymbol{x}_2^{(i)} \qquad (\, (i,j) \in \mathcal{S}_2 \,),$$

$$w^{\mathcal{E}},\, z, v \in \mathbb{R}, \qquad \sum_{(i,j,k)\in\mathcal{R}} p_{(ijk)} \left(w^{\mathcal{B}} - \boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3^{(ij)} - z \right) \;\le\; v \quad (\, \mathcal{R} \subset \mathcal{S}_3 \,).$$

$$(10)$$

Decomposition.   The first-stage problem will be:

$$\max \quad \sum_{i=1}^{N_1} p_1^{(i)}\, \mathcal{D}^{(i)} \left(\, \boldsymbol{r}_1^{(i)} \cdot \boldsymbol{x}_1,\; w^{\mathcal{E}}, z \,\right) \quad - \lambda z$$

$$(11)$$

such that $\quad w^{\mathcal{E}}, z \in \mathbb{R}, \quad \boldsymbol{x}_1 \in X, \; \mathbf{1} \cdot \boldsymbol{x}_1 = w_0,$

where the functions $\mathcal{D}^{(i)} : \mathbb{R}^3 \to \mathbb{R} \quad (i \in \mathcal{S}_1)$ are defined by the second-stage problem:

$$\mathcal{D}^{(i)}(\omega, \varpi, \zeta) := \quad \max \quad \sum_{j=1}^{N_2^{(i)}} p_2^{(ij)}\, \mathcal{D}^{(ij)} \left(\, \boldsymbol{r}_2^{(ij)} \cdot \boldsymbol{x}_2,\; \varpi, \zeta \,\right)$$

$$(12)$$

such that $\quad \boldsymbol{x}_2 \in X, \quad \mathbf{1} \cdot \boldsymbol{x}_2 = \omega,$

where the functions $\mathcal{D}^{(ij)} : \mathbb{R}^3 \to \mathbb{R} \quad ((i,j) \in \mathcal{S}_3)$ are defined by the third-stage problem:

$$\mathcal{D}^{(ij)}(\omega, \varpi, \zeta) := \quad \max \quad \sum_{k=1}^{N_3^{(ij)}} p_3^{(ijk)}\, \boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3 \quad -\varrho \sum_{k=1}^{N_3^{(ij)}} p_3^{(ijk)} \left[\boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3 - \varpi \right]^2 \quad - \lambda \tfrac{1}{1-\alpha}v$$

such that $\quad \boldsymbol{x}_3 \in X, \quad \mathbf{1} \cdot \boldsymbol{x}_3 = \omega, \qquad v \in \mathbb{R},$

$$\sum_{k\in\mathcal{K}} p_3^{(ijk)} \left(w^{\mathcal{B}} - \boldsymbol{r}_3^{(ijk)} \cdot \boldsymbol{x}_3 - \zeta \right) \;\le\; v \quad \left(\mathcal{K} \subset \left\{ 1, \ldots, N_3^{(ij)} \right\} \right).$$

$$(13)$$

Proposition 2 *The three-stage problem (11 - 13) is equivalent to the polyhedral representation problem (10).*

As a heuristic proof, let us observe that in each $\mathcal{D}$ function, the parameter ω represents initial wealth for the remaining period of time; and ζ represents a 'tolerable' loss level tentatively set in the first-stage problem. The optimal value of the variable v is the (conditional) expectation of losses above the 'tolerable' level. A formal proof for the ($\varrho = 0$) two-stage case can be found in Fábián (2007). This proof generalizes to multiple stages in a straightforward manner.

2.2   Solution method

Fábián (2007) describes a solution method for the two-stage problem that easily generalizes for multi-stage case. We give a brief sketch of the procedure, proceeding in a bottom-up manner:

Solution of a third-stage problem.   We assume a fixed $(i, j) \in \mathcal{S}_2$, given parameter values $\hat{\omega}, \hat{\varpi}, \hat{\zeta}$, and given stopping tolerance $\epsilon_3 > 0$. The problem can be solved by the Level Method that builds successive cutting-plane models of the problem.

The process stops when an ϵ_3-optimal solution is found. The final model problem is a linear programming problem that has the parameters $\omega = \hat{\omega}, \varpi = \hat{\varpi}, \zeta = \hat{\zeta}$ in the right-hand side. Using optimal dual variables of the final model problem, we can construct a linear function $\mathcal{L}^{(ij)}(\omega, \varpi, \zeta)$ such that

$$\mathcal{L}^{(ij)}(\omega, \varpi, \zeta) \geq \mathcal{D}^{(ij)}(\omega, \varpi, \zeta) \quad \left((\omega, \varpi, \zeta) \in \mathbb{R}^3 \right) \quad \text{and} \quad \mathcal{L}^{(ij)}(\hat{\omega}, \hat{\varpi}, \hat{\zeta}) \leq \mathcal{D}^{(ij)}(\hat{\omega}, \hat{\varpi}, \hat{\zeta}) + \epsilon_3.$$
(14)

We will call such a linear function an ϵ_3-support function to $\mathcal{D}^{(ij)}(\omega, \varpi, \zeta)$ at the point $(\hat{\omega}, \hat{\varpi}, \hat{\zeta})$.

Solution of a second-stage problem.   We assume a fixed $i \in \mathcal{S}_1$. Given the parameter values $\hat{\omega}, \hat{\varpi}, \hat{\zeta}$, and a tolerance $\epsilon_2 > 0$, we can use an approximate method to find an ϵ_2-optimal solution for the problem (12: $\omega = \hat{\omega}, \varpi = \hat{\varpi}, \zeta = \hat{\zeta}$). We recommend an inexact version of the Level Method. The inexact version was proposed by Fábián (2000). The Inexact Level Method is a bundle-type method that successively builds a cutting-plane model of the objective function. It needs an *oracle* to provide objective function data, in the form of ϵ_3-support functions. The accuracy tolerance ϵ_3 of the cuts is gradually decreased as the optimum is approached. (At each step, we have an estimate of how closely the optimum has been approached. The successive cut is generated with an accuracy tolerance derived from that estimate.)

Given an iterate $\hat{\boldsymbol{x}}_2$, suppose we need an ϵ_3-support function at the point $\left(\boldsymbol{r}_2^{(ij)} \cdot \hat{\boldsymbol{x}}_2, \ \hat{\varpi}, \hat{\zeta} \right)$. The oracle includes the solution of the third-stage problems (i, j) for $j = 1, \ldots, N_2^{(i)}$, with parameter values $\boldsymbol{r}_2^{(ij)} \cdot \hat{\boldsymbol{x}}_2, \ \hat{\varpi}, \hat{\zeta}$. Let us construct the ϵ_3-support functions $\mathcal{L}^{(ij)}$ as in (14). The aggregate function $\sum_j p_2^{(ij)} \mathcal{L}^{(ij)}$ is obviously an ϵ_3-support function to $\mathcal{D}^{(i)}$ at the point $\boldsymbol{r}_2^{(ij)} \cdot \hat{\boldsymbol{x}}_2, \ \hat{\varpi}, \hat{\zeta}$.

The process stops when an ϵ_2-optimal solution is found to the second-stage problem. The final cutting-plane model problem is a linear programming problem that has the parameters $\omega = \hat{\omega}, \varpi = \hat{\varpi}, \zeta = \hat{\zeta}$ in the right-hand side. Using optimal dual variables of the final model problem, we can construct an ϵ_2-support function to $\mathcal{D}^{(i)}(\omega, \varpi, \zeta)$ at the point $(\hat{\omega}, \hat{\varpi}, \hat{\zeta})$.

Solution of the first-stage problem. Suppose the decision maker prescribes the accuracy $\epsilon_1 > 0$. The Inexact Level Method can be used to find an ϵ_1-optimal solution of the first-stage problem.

In a former project, Fábián and Szőke (2006) successfully used Inexact Level-type methods for the solution of general two-stage stochastic programming problems. A stochastic programming problem with medium-sized first-stage and second-stage problems and almost-complete-recourse was solved with a growing number of scenarios. 6-digit optimal solution of the $600,000$-scenario problem was found in 13 minutes on a regular desktop computer. (The first-stage problem had 188 variables and 62 constraints ; and each of the $600,000$ second-stage problems had 272 variables and 104 constraints.)

Fábián and Szőke found that for a fixed stopping tolerance ϵ, the required number of steps does not depend on the number N of the scenarios. They also solved problems with increasing accuracy, and found that the number of steps required to find an ϵ-optimal solution, grew in proportion with $\log 1/\epsilon$. (This suggest a much better practical behavior than the theoretical efficiency estimate of the Level Method.)

In the present case, the second-stage problems have a special structure that is effectively exploited by the method of Künzi-Bay and Mayer.

A definitive validation of the numerical efficiency of the Level Decomposition scheme adapted to the present specific problems will require numerical experiments which will be the subject of subsequent research. However, the computational evidence published by Künzi-Bay and Mayer (2006) on the one hand, and by Fábián and Szőke (2006) on the other hand, support our confidence in the efficiency of the scheme.

3   Handling CVaR-constraints

Jobst and Zenios (2001) describe an experiment: they solved one-stage portfolio selection problems with time periods of one year. The objective was the minimization of a risk measure. As risk measure, they used MAD in one problem, and CVaR in another problem. They then simulated the returns of the optimal portfolios, at the time points 3, 6, 9, and 12 months after the beginning of the time period.

With the MAD-optimal portfolio, they found that catastrophic losses are probable after the first 3 months. But with the CVaR-optimal portfolio, worst losses were limited to around 10 % of the total wealth throughout. However, an interesting phenomenon occurred: the experimental distribution of the 9-month losses had a tail much heavier than the experimental distribution of the end-of-period losses had.

From the above phenomenon we conclude that the prototype model needs a correction: Beside controlling final risk, we also need controlling the risk at the ends of intermediate periods. One way is penalizing $\text{CVaR}_\alpha(w_t)$, i.e., α-Conditional-Value-at-Risk of the wealth at the end of the t-th period.

Let T denote the number of the time periods, and $\alpha_1, \ldots, \alpha_\tau$ a fixed finite set of positive confidence levels. We can include the new terms in the objective function (8) of the prototype model. The augmented objective will be:

$$\min \ -\text{E}(w_T) \ + \varrho \, \text{V}(w_T) \ + \sum_{t,\alpha} \lambda_{t\alpha} \text{CVaR}_\alpha(w_t), \tag{15}$$

where summation goes for $t = 1, \ldots, T$ and $\alpha \in \{\alpha_1, \ldots, \alpha_\tau\}$. The risk aversion parameters $\lambda_{t\alpha} > 0$ need to be determined by the decision maker.

If the decision maker can not determine the values of the parameter-vector $\boldsymbol{\lambda} = (\ldots, \lambda_{t\alpha}, \ldots)^T$, then we can help him by building an approximation of the efficient frontier. This is a $(T\tau+1)$-dimensional concave surface in the $(\ldots, Risk_{t\alpha}, \ldots, ExpectedYield - \varrho Variance)$ coordinate system. Points from this surface can be found by minimizing the objective function (15) with different settings of the parameter-vector $\boldsymbol{\lambda}$.

Alternatively, the decision maker may be able to prescribe risk constraints in the form $\text{CVaR}_\alpha(w_t) \leq \gamma_{t\alpha}$ $(t = 1, \ldots, T, \ \alpha \in \{\alpha_1, \ldots, \alpha_\tau\})$. The motivation for using risk constraints is that decision makers interpret and quantify right-hand sides easier than penalties in the objective function. Krokhmal, Palmquist, and Uryasev (2002) observe that optimization with multiple CVaR-constraints for different time frames and at different confidence levels allows the shaping of distributions according to the decision maker's preferences.

Roman, Mitra, and Darby-Dowman (2006) propose minimizing variance under constraints on expected yield and CVaR. They argue that such constraints are intuitively more appealing than variance constraints ; and minimizing a strictly convex objective is more convenient from a computational point view. In order to keep formulas as simple as possible, we are going to retain both expected yield and variance in the objective function. The decomposition scheme and the solution method we are going to describe easily adapt to handle constraints on expected yield.

We are going to formulate the problem in CVaR-constrained form. Virtual decision variables will $\boldsymbol{x}_1^{(i)}$ $(i \in \mathcal{S}_1)$, $\boldsymbol{x}_2^{(ij)}$ $((i,j) \in \mathcal{S}_2)$, $\ldots$, and they will be concatenated to form the long decision vector $\boldsymbol{x}$. The yields $w_1^{(i)}$ $(i \in \mathcal{S}_1)$, $w_2^{(ij)}$ $((i,j) \in \mathcal{S}_2)$, $\ldots$, and they will be concatenated to form the long yield vector $\boldsymbol{w}$. (Components of $\boldsymbol{w}$ directly depend on appropriate sub-vectors of $\boldsymbol{x}$.) Formally, the variables will be $\boldsymbol{\chi} := (\boldsymbol{x}, \boldsymbol{w})$. Let Ξ denote the feasible domain. (This is a convex polyhedron that expresses feasibility constraints on portfolios, and balances of wealth.)

We also make a technical modification in the objective function. Instead the term $\varrho \, \text{V}(w_T)$ measuring variance of the final yield, we introduce the function $\varphi(\boldsymbol{x})$. The latter will represent a positively weighted sum of variances of yields belonging to different time windows.

We formulate the problem in the form

$$
\begin{aligned}
\min \ &-\mathrm{E}(w_T) + \varphi(\boldsymbol{x}) \\
\text{such that} \ &(\boldsymbol{x}, \boldsymbol{w}) \in \Xi, \\
&\mathrm{CVaR}_\alpha(w_t) \le \gamma_{t\alpha} \quad (t = 1, \dots, T, \ \alpha \in \{\alpha_1, \dots, \alpha_\tau\}).
\end{aligned}
\tag{16}
$$

We assume that the function $\varphi(\boldsymbol{x})$ is strictly convex.

Remark 3 *Strict convexity can be ensured by appropriate selection of the variance terms included in $\varphi(\boldsymbol{x})$. Consider, e.g., the following construction: At each non-terminal node of the scenario tree, the single-stage variance of next period's yield is computed relative to the knowledge available at that node. The covariance matrices of the return vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2 | \boldsymbol{r}_1 = \boldsymbol{r}_1^{(i)}$ $(i \in \mathcal{S}_1)$, $\dots$ are positive definite, due to our independence assumptions. It follows that the variances $V(w_1)$, $V\left(w_2 | \boldsymbol{r}_1 = \boldsymbol{r}_1^{(i)}\right)$ $(i \in \mathcal{S}_1)$, $\dots$ are strictly convex functions of the variables $\boldsymbol{x}_1$, $\boldsymbol{x}_2^{(i)}$ $(i \in \mathcal{S}_1)$, $\dots$, respectively.*

Fábián (2007) proposed a decomposition framework for two-stage CVaR minimization and CVaR constrained problems. The decomposition scheme is based on the Künzi-Bay – Mayer polyhedral representation. The solution methods are special Level-type methods, variants of the Level Method of Lemaréchal, Nemirovskii, and Nesterov (1995). This scheme is effective only for the handling of a limited number of CVaR-constraints. (The number of the CVaR-constraints should be significantly less then the number of the first-stage scenarios.)

We propose a dual decomposition framework that is capable of handling a large number of simultaneous CVaR constraints in (16). The parameters γ in (16) are to be set by the decision maker. We assume that the parameters are set in such a manner that the problem is feasible, and the Slater condition holds. The parameters can be calibrated by exploring the efficient frontier. The decision maker ought to take into consideration the slope of the efficient frontier also. (Slope is interpreted as risk aversion.)

In subsection 3.1 we describe a method that estimates respective optimal objective values for given values of the right-hand sides $\boldsymbol{\gamma}$. This method can be used to explore the efficient frontier, and thus help the decision maker in setting the right-hand side parameters. In subsection 3.2 we describe a method for finding an appropriate primal solution.

3.1 Estimating optimal objective values

Relaxing CVaR-constraints in (16), we get the Lagrangian

$$
L(\boldsymbol{\chi}, \boldsymbol{\lambda}) := -\mathrm{E}(w_T) + \varphi(\boldsymbol{x}) + \sum_{t,\alpha} \lambda_{t\alpha} \left(\mathrm{CVaR}_\alpha(w_t) - \gamma_{t\alpha}\right).
\tag{17}
$$

Here $\boldsymbol{\chi} := (\boldsymbol{x}, \boldsymbol{w})$ denotes the vector of the formal variables of (16). On the other hand, $\boldsymbol{\lambda} := (\dots, \lambda_{t\alpha}, \dots)^T$ denotes the non-negative vector of Lagrange multipliers.

The dual objective function is then

$$D(\boldsymbol{\lambda}) := \min_{\boldsymbol{\chi} \in \Xi} L(\boldsymbol{\chi}, \boldsymbol{\lambda}) \qquad (\boldsymbol{\lambda} \geq \mathbf{0}), \tag{18}$$

where Ξ denotes the feasible polyhedron of (16).

Given fixed values $\hat{\boldsymbol{\lambda}} \geq \mathbf{0}$ of the multipliers, the minimization problem of (18) fits the prototype discussed in section 2. The decomposition scheme and the solution method of section 2 can be adapted to the present problem in a straightforward manner. (CVaR-terms belonging to different time periods can be decomposed independently of each other.) Formally, let $\epsilon > 0$ be fixed. We can find a vector $\hat{\boldsymbol{\chi}}$ that is an ϵ-optimal solution in the sense

$$\hat{\boldsymbol{\chi}} \in \Xi \quad \text{and} \quad L(\hat{\boldsymbol{\chi}}, \hat{\boldsymbol{\lambda}}) \leq L(\boldsymbol{\chi}, \hat{\boldsymbol{\lambda}}) + \epsilon \quad \text{holds for every } \boldsymbol{\chi} \in \Xi. \tag{19}$$

Let $\hat{\boldsymbol{x}}$ denote the x-components of $\hat{\boldsymbol{\chi}}$, and let $\hat{\boldsymbol{w}}$ denote the w-components. Then the linear function

$$l(\boldsymbol{\lambda}) := L(\hat{\boldsymbol{\chi}}, \boldsymbol{\lambda}) = -\mathrm{E}(\hat{w}_T) + \varphi(\hat{\boldsymbol{x}}) + \sum_{t,\alpha} \lambda_{t\alpha} \left(\mathrm{CVaR}_\alpha(\hat{w}_t) - \gamma_{t\alpha} \right) \tag{20}$$

is an ϵ-support function to the concave $D(\boldsymbol{\lambda})$ at $\hat{\boldsymbol{\lambda}}$. Indeed, $l(\boldsymbol{\lambda}) \geq D(\boldsymbol{\lambda})$ $(\boldsymbol{\lambda} \geq \mathbf{0})$ follows from the definition of the dual function and the inclusion in (19). Moreover, $l(\hat{\boldsymbol{\lambda}}) \leq D(\hat{\boldsymbol{\lambda}}) + \epsilon$ follows from the inequality in (19).

Instead of the CVaR-constrained problem (16), we are going to solve the dual problem

$$\max_{\boldsymbol{\lambda} \geq \mathbf{0}} D(\boldsymbol{\lambda}). \tag{21}$$

For the solution of the dual problem we propose using the Inexact Level Method mentioned in the previous section. Given a stopping tolerance $\delta > 0$, we can find a δ-optimal solution.

Remark 4 *Convergence proof of the Inexact Level Method requires compactness of the feasible domain ; and existence of a common upper bound on the slopes of the linear ϵ-support functions (20) generated in course of the process.*

The latter requirement is satisfied by finding a common upper bound on the possible $CVaR_\alpha(w_t)$ terms in (20). This is easily done using our assumptions on the feasible domain X. (See remark 2 in Fábián (2007) for a formal computation.)

The feasible domain of the dual problem (21) is not bounded formally. But the slope of the efficient frontier is interpreted as risk aversion of the decision maker. Hence we may assume that the decision maker will select such right-hand sides that for each (t, α), the $Risk_{t\alpha}$-direction slope of the efficient frontier be less than $\overline{\lambda}$, with a reasonably large constant $\overline{\lambda} > 0$. Thus the virtual feasible domain is $\overline{\lambda}\mathbf{1} \geq \boldsymbol{\lambda} \geq \mathbf{0}$, and that satisfies the compactness requirement.

3.2   Finding primal solution

Approximations described in the previous section help the decision maker in exploring the efficient frontier. Let γ denote the vector of right-hand sides finally selected. (We assume that γ is such that the constrained problem (16) is feasible and the Slater condition holds.)

Let $\lambda^\star$ be an optimal solution of the dual problem (21). Moreover, let $\chi^\star$ be the optimal solution of the minimization problem in (18 : $\lambda = \lambda^\star$). Due to strict convexity of the objective function, the latter problem has a unique optimal solution. Hence the unique $\chi^\star$ will be an optimal solution of the CVaR-constrained problem (16).

4   Conclusions and prospect

In this paper we proposed a dual decomposition scheme and a solution method that can handle problems having a large number of CVaR-constraints. This scheme meets the need for handling simultaneous CVaR-constraints for different time frames and at different confidence levels.

The process consists of two steps. In the first step we construct an approximation of the efficient frontier. This helps the decision maker in calibrating the right-hand-side parameters of the CVaR constraints. In course of the first step we only need near-optimal dual solutions. The dual problem is decomposed and solved by a scheme proposed by Fábián (2007). We use the Künzi-Bay – Mayer (2006) cutting-plane representations of CVaR.

Once the parameters are set to the satisfaction of the decision maker, we pass to the second step, and find a primal optimal solution. Like Roman, Mitra, and Darby-Dowman (2006), we include variance terms in the objective function. Besides the economic advantages observed by Roman et al., the variance terms yield a technical advantage: They make the objective function strictly convex, and hence enable solution of the primal problem through a dual approach.

A definitive validation of the numerical efficiency of the Level Decomposition scheme adapted to the present specific problems will require numerical experiments which will be the subject of subsequent research. However, the computational evidence published by Künzi-Bay and Mayer (2006) on the one hand, and by Fábián and Szőke (2006) on the other hand, support our confidence in the efficiency of the scheme.

4.1   Decomposition of further risk constraints

With minor modifications, the proposed scheme can be used to decompose further risk constraints in portfolio management problems:

Integrated Chance Constraints   were introduced by Klein Haneveld (1986) for static models. Drijver et al. (2002) and Klein Haneveld et al. (2005) use single-stage ICCs in dynamic models. Fábián and Szőke (2006) proposed two-stage ICCs to formulate non-traditional stochastic constraints.

Klein Haneveld and van der Vlerk (2006) proposed a cutting-plane representation of generic ICCs. As we mentioned in Section 1.2, the Künzi-Bay – Mayer cutting-plane representation is the CVaR analogue of this. Based on the cutting-plane representation, Klein Haneveld and van der Vlerk developed a solution method for simple recourse problems. They implemented the method and for simple recourse problems, their special solver proved orders of magnitude faster than existing complete recourse solvers.

A generic ICC prescribes a bound for a term $E([w^{\mathcal{B}} - w]_+)$. In portfolio management context, $w^{\mathcal{B}}$ is a user-defined parameter that represents the wealth we intend to accumulate by a certain point of time and w is a random variable that represents our yield at this point of time. Let us consider a multi-stage problem with the objective

$$\max \ E(w) \ - \varrho \, V(w) \ - \lambda \, \text{Risk}(w),$$

like we did in section 2. But for the Risk term, let us substitute $E([w^{\mathcal{B}} - w]_+)$ instead of $\text{CVaR}(w)$. The resulting stochastic programming problem decomposes in a straightforward manner. (If we fix the z variables to 0 in the cutting-plane representations of section 2, then the Künzi-Bay – Mayer cuts transform into Klein Haneveld – van der Vlerk cuts.) Hence the dual decomposition scheme of section 3 easily adapts to ICCs having the form $E([w^{\mathcal{B}} - w]_+) \leq \gamma$. ($\gamma$ being a user-defined tolerance.)

Second-order Stochastic Dominance establishes a partial ordering between random variables. Let w and $\widehat{w}$ denote the respective random yields of two different portfolios. If $E(U(w)) \geq E(U(\widehat{w}))$ holds for any monotonic and concave (integrable) utility function U, then w is said to dominate $\widehat{w}$. This is equivalent to $E([w^{\mathcal{B}} - w]_+) \leq E([w^{\mathcal{B}} - \widehat{w}]_+)$ holding for any $w^{\mathcal{B}} \in \mathbb{R}$. (Consistency between SSD and CVaR is established in Ogryczak and Ruszczyński (2002).)

Dentcheva and Ruszczyński (2006) propose a single-stage portfolio optimization model involving an SSD constraint. The motivation is the following: assume we want to emulate a certain stock index $\widehat{w}$. This is achieved by constructing a portfolio whose yield w dominates $\widehat{w}$.

In general, an SSD constraint can be expressed by a continuum of ICCs. If, however, $\widehat{w}$ has a finite discrete distribution with realizations $\widehat{w}^{(s)}$ ($s = 1, \ldots, S$), then the SSD constraint is equivalent to the finite set of ICCs

$$E([\widehat{w}^{(s)} - w]_+) \ \leq \ E([\widehat{w}^{(s)} - \widehat{w}]_+) \qquad (s = 1, \ldots, S).$$

Assuming discrete finite distributions, Dentcheva and Ruszczyński formulate the SSD-constrained problem in linear programming form. They use the embedding representation, introducing new variables for positive parts. The resulting LP problems have a specific structure. For such specific problems, the authors develop a duality theory in which the dual objects are utility functions. Based on this duality theory, they construct a dual problem that consists of the minimization of a weighted sum of polyhedral convex functions. Domains, function values, subgradients are easily computable. The authors adapted the Regularized Decomposition method of Ruszczyński (1986) to these special dual problems.

As an alternative, we propose a direct cutting-plane approach. Since the SSD constraint is equivalent to a finite set of ICCs, the previous paragraph is applicable. Hence the Klein Haneveld – van der Vlerk cuts and the dual decomposition scheme of section 3 easily adapts to the SSD-constrained problem.

Moreover, we propose a multistage generalization of the SSD-constrained model. To extend the notation introduced in section 2, let $\widehat{w}_1, \widehat{w}_2, \widehat{w}_3$ denote the stock index at the respective ends of the first, second, and third periods. We want our first-period yield w_1 to dominate $\widehat{w}_1$, our second-period yield w_2 to dominate $\widehat{w}_2$, and our third-period yield w_3 to dominate $\widehat{w}_3$. The resulting problem can be solved by the dual decomposition scheme of section 3.

A value-of-information risk measure   for multiperiod income processes was introduced by Pflug (2006). If the random process is represented by a scenario tree, then the value-of-information risk measure is computed as a weighted sum of *conditional-value-at-risk deviations*. The conditional-value-at-risk deviation of a random income Y is defined as $\mathrm{CVaRD}(Y) := \mathrm{E}(Y) + \mathrm{CVaR}(Y)$. At each non-terminal node of the tree, the single-stage CVaRD of next period's income is computed relative to the knowledge available at that node. (Confidence level is determined by deterministic parameters.)

Pflug defines the random income process in case of a special multiperiod portfolio optimization problem. He proposes minimizing the value-of-information risk measure under a constraint on expected end-of-horizon wealth. He constructs a linear programming problem. The number of the decision variables and constraints in this problem is a multiple of the number of nodes in the scenario tree. Hence in a realistic application this LP problem will be very large.

We propose relaxing the constraint on expected end-of-horizon wealth and including it in the objective function. The solution method of section 2 easily adapts to the resulting problem. Alternatively, we propose maximizing expected end-of-horizon wealth under a constraint on the value-of-information risk measure. The dual decomposition scheme of section 3 can be adapted to the resulting problem.

4.2   Further fields of potential application

Option pricing. Ryabchenko et al. (2004) propose an ingenious grid-based replication model for the pricing of a European option in an incomplete market. Replication models represent option price as the value of a portfolio consisting of the underlying stock and a risk-free asset. The replicating portfolio is dynamically rebalanced as the value of the underlying stock varies. (Background information and references can be found in the cited paper.)

Ryabchenko et al. approximate optimal rebalancing strategy through a grid in the (time × underlying stock price) plane. They use a set of sample paths to model underlying stock behavior. Variables in this model belong to grid points, hence the number of the variables is independent of the number of the sample paths. Sample paths only affect the objective function and a single constraint, and these contain only expectations. The objective is

the minimization of the average of squared approximation error. The mentioned constraint prescribes that the approximation error should average out to 0. Further constraints express distribution-independent monotonicity and convexity characteristics of option prices. The paper reports remarkably good test results obtained with only 100 - 200 sample paths.

The present decomposition scheme enables a traditional approach: Let underlying stock behavior be modeled by a scenario tree. We construct a self-financing portfolio, i.e., no re-financing is allowed at decision points. We assume an incomplete market, hence the final net result of the option seller is random. We impose a CVaR constraint on final net result (or multiple CVaR constraints with different confidence levels). Moreover we prescribe that expected final net result should cover administration costs of the seller. (Transaction costs can also be taken into consideration.) We minimize the required initial investment under these conditions.

The models of sections 2 and 3 need but slight modifications: Option price will be represented by initial investment w_0, hence it will be a decision variable. Option value at maturity will be represented by benchmark wealth $w^{\mathcal{B}}$, hence it will be a random parameter. There will be two assets ($n = 2$), only one of them having random returns. Instead of maximizing end-of-horizon portfolio yield $\mathrm{E}(w)$, we will impose a constraint on expected final net result $\mathrm{E}(w - w^{\mathcal{B}})$, and minimize w_0.

Asset-Liability Management. Drijver et al. (2002) and Klein Haneveld et al. (2005) model the operation of a company-owned pension fund. In this case, liabilities consist of future pension payments. Funding comes from revenues of investments and regular contributions by active members. The funding ratio is A/L, where A is the value of the assets, and L is the value of liabilities. Seen over a number of years, the funding ratio may fall short occasionally, but if this happens too often or if the shortage is too large, then the owner company is required to make a remedial contribution to the fund.

The cited authors represent random parameters by a scenario tree and build a multistage stochastic programming model. Time periods are years, and decisions are made at the points of time $t = 0, 1, \ldots, T$. At each non-terminal decision node, a short-term risk constraint is imposed. It is an integrated chance constraint on subsequent year's funding ratio in the form of $\mathrm{E}([\alpha L_{t+1} - A_{t+1}]_+) \leq \beta$, where α, β are user-defined parameters, and the expectation is composed relative to the knowledge available at the relevant node. Moreover, in the regulation prescribing that 'funding ratio should not fall short too often', the term 'too often' is interpreted as 'in two consecutive years'. Hence at each (non-root) node an additional, mid-term constraint is imposed: if previous year's funding ratio fall short, then any shortage in current year's funding ratio has to be corrected by a remedial contribution at the present node. These mid-term constraints are expressed by binary variables.

As an alternative to these mid-term constraints, we propose a direct generalization of the above described short-term constraints: Given a decision node that belongs to the point of time $t < T - 1$, we impose a risk constraint on the funding ratio of year $t + 2$. If we use integrated chance constraints, then this alternative mid-term constraint takes the form $\mathrm{E}([\alpha L_{t+2} - A_{t+2}]_+) \leq \beta$, where the expectation is composed relative to the knowledge

available at the relevant node.

These alternative mid-term constraints are clearly less direct than those proposed in the cited papers. But the heuristic meaning is similar: at each decision point, special care must be taken of short- and mid-term consequences. From computational point of view, the alternative formulation requires no binary variables, and the alternative mid-term constraints fit into the proposed decomposition scheme.

References

[1] AHMED, S. (2006). Convexity and decomposition of mean-risk stochastic programs. *Mathematical Programming* **106**, 433-446.

[2] ANDERSSON, F., H. MAUSSER, D.ROSEN, and S. URYASEV (2001). Credit risk optimization with Conditional Value-at-Risk criterion. *Mathematical Programming,* Series B **89**, 273-291.

[3] CHARNES, A., W.W. COOPER, and G.H. SYMONDS (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science* **4**, 235-263.

[4] CLAESSENS, S. and J. KREUSER (2004). A framework for strategic foreign reserves risk management. In: *Risk management for Central Bank foreign reserves* (editors: C. Bernadell, P. Cardon, J. Coche, F.X. Diebold, and S. Manganelli), 47-74. European Central Bank, Frankfurt, Germany.

[5] DENTCHEVA, D. and A. RUSZCZYŃSKI (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance* **30**, 433-451.

[6] DRIJVER, S.J., W.K. KLEIN HANEVELD, and M.H. VAN DER VLERK (2002) ALM model for pension funds: numerical results for a prototype model. *Research Report* 02A44, Research School SOM, University of Groningen, the Netherlands.

[7] FÁBIÁN, C.I. (2000). Bundle-type methods for inexact data. *Central European Journal of Operations Research* **8** (special issue, T. Csendes and T. Rapcsák, eds.), 35-55. (Also available through http://www.cs.elte.hu/∼fabian.)

[8] FÁBIÁN, C.I. (2007). Handling CVaR objectives and constraints in two-stage stochastic models. *European Journal of Operational Research* (special issue on Continuous Optimization in Industry, T. Illés, M. Lopez, J. Vörös, T. Terlaky, G-W. Weber, eds.), accepted, DOI: 10.1016/j.ejor.2007.02.052. (Also available through http://www.cs.elte.hu/∼fabian.)

[9] FÁBIÁN, C.I. and Z. SZŐKE (2006). Solving two-stage stochastic programming problems with level decomposition, *Computational Management Science* to appear, DOI: 10.1007/s10287-006-0026-8. (Also available through http://www.cs.elte.hu/∼fabian.)

[10] JOBST, N.J. and S. ZENIOS (2001). The tail that wags the dog: Integrating credit risk in asset portfolios. *Working Paper* #01-05. HERMES, European Center of Excellence on Computational Finance and Economics, University of Cyprus, Nicosia.

[11] KALL, P. and J. MAYER. (2005). *Stochastic Linear Programming: Models, Theory, and Computation.* Springer, International Series in Operations Research and Management Science.

[12] KLEIN HANEVELD, W.K. (1986). Duality in Stochastic Linear and Dynamic Programming. *Lecture Notes in Economics and Math. Systems* **274** Springer-Verlag, New York.

[13] KLEIN HANEVELD, W.K., M.H. STREUTKER, and M.H. VAN DER VLERK (2005). An ALM model for pension funds using integrated chance constraints. *Research Report* 05A03, Research School SOM, University of Groningen, the Netherlands.

[14] KLEIN HANEVELD, W.K. and M.H. VAN DER VLERK (2006). Integrated chance constraints: reduced forms and an algorithm. *Computational Management Science* **3**, 245-269. First published as *SOM Research Report* 02A33, University of Groningen, 2002.

[15] KROKHMAL, P., J. PALMQUIST, and S. URYASEV (2002). Portfolio optimization with Conditional Value-at-Risk objective and constraints. *Journal of Risk* **4**, 11-27.

[16] KÜNZI-BAY, A. and J. MAYER (2006). Computational aspects of minimizing conditional value-at-risk. *Computational Management Science* **3**, 3-27.

[17] LEMARÉCHAL, C., A. NEMIROVSKII, and YU. NESTEROV. (1995). New variants of bundle methods. *Mathematical Programming* **69**, 111-147.

[18] OGRYCZAK, W. and A. RUSZCZYŃSKI (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization* **13**, 60-78.

[19] PFLUG, G.CH. (2006). A value-of-information approach to measuring risk in multi-period economic activity. *Journal of Banking & Finance* **30**, 695-715.

[20] PRÉKOPA, A. (1973). Contributions to the theory of stochastic programming. *Mathematical Programming* **4**, 202-221.

[21] PRÉKOPA, A. (2003). Probabilistic programming. In: *Stochastic Programming, Handbooks in Operations Research and Management Science* **10** (A. Ruszczyński and A. Shapiro, eds.), Elsevier, Amsterdam, 267-351.

[22] ROCKAFELLAR, R.T. and S. URYASEV (2000). Optimization of conditional value-at-risk. *Journal of Risk* **2**, 21-41.

[23] ROCKAFELLAR, R.T. and S. URYASEV (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* **26**, 1443-1471.

[24] ROMAN, D., G. MITRA, and K. DARBY-DOWMAN (2006). Mean-risk models using two risk measures: a multi-objective approach. *Technical Report* 51/06. CARISMA: The Centre for the Analysis of Risk and Optimisation Modelling Applications, School of Information Systems, Computing and Mathematics, Brunel University, London.

[25] RUSZCZYŃSKI, A. (1986). A Regularized Decomposition Method for Minimizing the Sum of Polyhedral Functions. *Mathematical Programming* **35**, 309-333.

[26] RYABCHENKO, V., S. SARYKALIN, and S. URYASEV (2004). Pricing European options by numerical replication: quadratic programming with constraints. *Asia-Pacific Financial Markets* **11**, 301-333.

[27] TOPALOGLOU, N. (2004). A stochastic programming framework for international portfolio management. *Ph.D. dissertation* (advisors: H. Vladimirou and S. Zenios). School of Economics and Management, University of Cyprus.

[28] TOPALOGLOU, N., H. VLADIMIROU, and S. ZENIOS (2006). A dynamic stochastic programming model for international portfolio management. *European Journal of Operational Research*, in press.