Portfolio Optimization with Combinatorial and Downside Return Constraints

Miguel A. Lejeune

Abstract We study a probabilistic portfolio optimization model in which trading restrictions modeled with combinatorial constraints are accounted for. We provide several deterministic reformulations equivalent to this stochastic programming problem and discuss their computational efficiency. The reformulated problem takes the form of a mixed-integer nonlinear problem and is solved with an exact outer approximation algorithm. This latter is based on the early recognition of the problem structure and permits a hierarchical organization of the computations. Computational tests show the contribution of the proposed algorithm that outperforms the Cplex 12.4 solver in terms of computational time and quality of the obtained solutions.

Keywords Stochastic portfolio optimization • Probabilistic Markowitz • Combinatorial trading constraints • Stochastic programming • Outer approximation algorithm

1 Introduction

Much effort has been devoted in the last decade to the extension of the mean–variance portfolio optimization model proposed by Markowitz [17] in the late 1950s (see [9] for a recent review). The mean–variance approach trades off expected returns against variance with this latter used as the risk measure reflecting the volatility of the market.

Consider n assets with position given by the vector w. Let $\mu \in \mathcal{R}^n$ be the n-dimensional mean return. The Markowitz model assumes that the expected returns

---

M.A. Lejeune (✉)
Department of Decision Sciences, George Washington University, 2201 G Street, Suite 415, NW Washington, DC, USA
e-mail: mlejeune@gwu.edu

L.F. Zuluaga and T. Terlaky (eds.), *Modeling and Optimization: Theory and Applications*, Springer Proceedings in Mathematics & Statistics 84, DOI 10.1007/978-1-4614-8987-0_2,

and the components of the positive definite variance–covariance return matrix $\Sigma \in \mathcal{R}^{n\times n}$ are known. Each of the three variants of the mean–variance model is formulated as a convex nonlinear optimization problem. One of them involves the construction of a portfolio with minimal portfolio variance $w^T \Sigma w$ provided that a prescribed return level R is attained and can be formulated as the following mathematical program:

$$\mathbf{MV}: \quad \min \quad w^T \Sigma w \tag{1}$$

$$\text{subject to} \quad \mu^T w \geq R \tag{2}$$

$$w \in \mathcal{X}. \tag{3}$$

The notation $\mathcal{X}$ refers to the polytope defining the set of admissible portfolios

$$\mathcal{X} = \left\{ w \in \mathbb{R}^n : e^T w = 1,\ w \geq 0 \right\}, \tag{4}$$

and e is an all-one vector or appropriate dimension. Optimal portfolios located on the efficient frontier are those exposing the investor to the minimum possible risk (i.e., variance) and providing a specified return level R in (2). The first constraint in (4) ensures that the totality of the capital is invested in the n assets while the other constraints prevents short selling.

An issue commonly associated with the Markowitz model is the estimation risk which refers to the sensitivity of the mean–variance optimal portfolios with respect to the single-point estimate of the parameters (expected returns and variance–covariance matrix) of the model. Multiple empirical studies have shown that minor perturbations in the estimates of these parameters can lead to a drastically different optimal portfolio. In practice, investors prefer to trade off some return for more safety and construct a portfolio that performs well under diverse market conditions. Models based on stochastic dominance [5] and programming [21], robust optimization [6], or robust statistics [22] have been designed in order to obtain allocation policies that are less affected by the side effects of the estimation risk. In this study, we focus on the probabilistic asset allocation model **PMV** introduced by Bonami and Lejeune [2] that takes into account the difficulty to estimate asset returns and the risk to rely on a one-point estimate. Accordingly, the **PMV** model accounts for the incomplete knowledge of the return behavior and defines the vector of asset returns as a vector ξ of random variables. The model **PMV** takes the form of a stochastic programming problem with random technology matrix in which the decision variables w are multiplied by stochastic coefficients ξ that are not (necessarily) independent

$$\begin{aligned} \mathbf{PMV}: \quad \min \quad & w^T \Sigma w \\ \text{subject to} \quad & \mathbb{P}(\xi^T w \geq R) \geq p \qquad (5) \\ & w \in \mathcal{X}. \end{aligned}$$

The symbol $\mathbb{P}$ refers to a probability measure while p is the specified probability level typically defined on $[0.7, 1)$. The asset allocation model **PMV** implements a downside risk measure that probabilistically prevents the return of the portfolio to fall below the return level R in (5) and can be viewed as a probabilistic version of the Markowitz model. The risk measure is closely related to Roy's safety-first risk metric [20] and to Kataoka's model [11].

Besides the estimation risk concern, the Markowitz model is often amended by portfolio managers to integrate specific institutional features and trading criteria [6]. We describe in this paper some of the most common practical investment requirements that are formulated with combinatorial constraints. Those include the presence of nonproportional transaction costs, the requirements to invest a minimal amount in any selected asset (buy-in threshold constraint), to invest in a minimum number of asset classes or industrial sectors (diversification constraint), to restrict the number of positions hold (cardinality constraint), and to buy shares by lots of batches (round lot constraint). As we shall see in Sect. 3, these practical considerations require the use of integer decision variables and the introduction of combinatorial constraints in model **PMV**, transforming it into a stochastic integer problem and further compounding its numerical solution.

Our contributions are twofold. On the modeling side, we provide a series of deterministic and convex formulations equivalent to the probabilistic portfolio optimization problem and analyze their features (Sect. 2). The analysis is based on earlier results presented in [2, 7]. We also review some of the most common trading restrictions and describe how they can be formulated with combinatorial constraints (Sect. 3). On the algorithmic side, we develop a variant of the exact outer approximation algorithm proposed in [13] (see Sect. 4). Computational tests indicate the computational benefits of using the proposed algorithm (Sect. 5).

2 Model Reformulations

Under the formulation **PMV**, the probabilistic portfolio optimization problem cannot be handled by any optimization solvers. We shall now review a number of conic reformulations (see [2,7]) that are greatly beneficial for the numerical solution of problem **PMV**.

2.1 *Nonlinear Reformulation*

Let $\psi = \frac{\xi^T w - \mu^T w}{\sqrt{w^T \Sigma w}}$ be a random variable with mean 0 and variance 1 representing the normalized portfolio return. Further, we denote by $F_{(w)}$ the cumulative probability distribution of ψ, by $F^{-1}_{(w)}$ its inverse, and by $F^{-1}_{(w)}(1-p)$ its $(1-p)$ −quantile. The subscript w indicates that the exact form of the probability distribution F depends on the holdings w of the portfolio.

Theorem 1 (Kataoka [11]). *The nonlinear optimization problem* **C1**

$$\begin{aligned} \textbf{C1}: \quad & \min && w^T \Sigma w \\ & \text{subject to} && \mu^T w + F^{-1}_{(w)}(1-p)\sqrt{w^T \Sigma w} \geq R \\ &&& w \in \mathcal{X} \end{aligned} \tag{6}$$

is equivalent to **PMV**.

The next subsection is based on Sect. 2.2.1 in [2] and studies under which conditions the feasible set defined by the nonlinear constraint (6) is convex, thereby making **C1** a convex optimization problem.

2.2 Convexity

(a) Symmetric Probability Distributions

Theorem 2 studies the convexity of the feasible set defined by (6) for symmetric probability distributions.

Theorem 2 (Bonami and Lejeune [2]). *If $p \in [0.5, 1)$ and if the probability distribution of $\xi^T w$ is symmetric, the constraint $\mu^T w + F^{-1}_{(w)}(1-p)\sqrt{w^T \Sigma w} \geq R$ equivalent to (5) is a second-order cone constraint.*

Thus, problem **C1** minimizes a convex quadratic function over a second-order cone and some linear constraints and is therefore a convex problem.

(b) Positively Skewed Probability Distributions

Skewness is typically used to measure the asymmetry of a probability distribution. A probability distribution is said to be right-skewed or to have positive skewness (resp., left-skewed or negative skewness) if the right, upper value (resp., left, lower value) tail is longer or fatter than the left, lower value (resp., right, upper value).

Definition 1. Let m be the median of the probability distribution F of an r-variate random vector ψ with $E[\psi] = 0$. The distribution F has positive skewness if

$$\mathbb{P}(0 \geq \psi) \geq \mathbb{P}(m \geq \psi) \Leftrightarrow F^{-1}(\alpha) \leq 0,\ \alpha \leq 0.5\ .$$

Theorem 3 (Bonami and Lejeune [2]). *If $p \in [0.5, 1)$ and if the probability distribution of $\xi^T w$ has positive skewness, then $\mu^T w + F^{-1}_{(w)}(1-p)\sqrt{w^T \Sigma w} \geq R$ equivalent to (5) is a second-order cone constraint.*

2.3 Compact Reformulations of Variance

This section is based on the results proposed by Filomena and Lejeune [7] and provides two alternate ways to calculate the variance of the portfolio. To ease the notation, we shall thereafter refer to $F^{-1}_{(w)}(1-p) = -\kappa$.

2.3.1 Convex Reformulations with Cholesky Decomposition

We shall first use the Cholesky decomposition to derive the convex programming problems **C2** and **E2** equivalent to **C1**. The Cholesky decomposition involves the calculation of a lower triangular matrix C such that $\Sigma = CC^T$, with Σ positive definite. Substituting CC^T for Σ in the objective function and in (6) and introducing the auxiliary nonnegative decision variable h (10), we obtain the convex optimization problem **C2**:

$$\mathbf{C2}: \quad \min \quad \|C^T w\|_2^2 \tag{7}$$

$$\text{subject to} \quad (3)$$

$$\mu^T w - R \geq h \tag{8}$$

$$\kappa \, \|C^T w\|_2 \leq h, \tag{9}$$

$$h \geq 0 \tag{10}$$

where $\|x\|_2$ denotes the Euclidean norm of the vector x. The epigraph formulation of problem **C2** follows and has the canonical form of a second-order cone programming problem.

Theorem 4 (Filomena and Lejeune [7]). *Problem* **E2**

$$\mathbf{E2}: \quad \min \quad h \tag{11}$$

$$\text{subject to} \quad (3); (8); (9); (10)$$

is equivalent to problem **C2**.

2.4 Convex Reformulations with Period-Separable Formulation of Variance

The model proposed in this section uses Proposition 1 to reformulate the functions involving the variance of the portfolio as the Euclidian norm of a multidimensional vector.

Proposition 1 (Filomena and Lejeune [7]). *Let M be the number of data points used for estimation purposes. Let r_{jt} be the observed return of j at t and define the auxiliary variables:*
$b_t = \sum_{j=1}^{n} \left(r_{jt} - \mu_j\right) w_j \ , \ t = 1, \ldots, M$. *The variance of the portfolio return becomes*

$$w^T \Sigma w = \frac{1}{M} \|b\|_2^2 \, .$$

The variance of the portfolio is now written in a separable form and includes M squared terms b_t representing each the part of the variance associated with period $t, t = 1, \ldots, M$.

We introduce M decision variables b_t unrestricted in sign and apply Proposition 1 for the reformulation of the probabilistic model as the convex optimization problem **C2**:

$$\mathbf{C2}: \quad \min \quad \frac{1}{M}\|b\|_2^2 \tag{12}$$

$$\text{subject to} \quad (3); (8); (10)$$

$$\frac{\kappa}{\sqrt{M}}\ \|b\|_2 \leq h \tag{13}$$

$$b_t - \sum_{j=1}^{n} \left(r_{jt} - \mu_j\right) w_j = 0\ ,\ t = 1, \ldots, M. \tag{14}$$

As for **C1**, we provide the epigraph formulation **E2** of problem **C2**:

$$\mathbf{E2}: \quad \min \quad h \tag{15}$$

$$\text{subject to} \quad (3); (8); (10); (13); (14)$$

The modeling of the variance in problems **C0**, **C1**, and **E1** requires the a priori estimate of $\frac{n(n+1)}{2}$ covariance terms and can generate model specification issues, such as the obtaining of a variance–covariance matrix that is not positive semi-definite [4]. The computation of the variance proposed in this section does not have such limitations and does not make any assumption on the form or rank of Σ. Furthermore, the number of quadratic terms in problems **C2** and **E2** is much smaller than in the formulations **C0**, **C1**, and **E1** for large-scale portfolio optimization problems when the number n of assets exceed by far the number M of observations.

2.5 Equivalent Formulations and Inner Approximations Based on Probability Inequalities

In this section, we study for which classes of probability distributions the above problem is a second-order cone optimization problem (i.e., thus convex and solvable in polynomial time). Results extracted from [2, 12] are presented here and indicate that it is not always possible to derive an exact closed-form formulation of the second-order cone problem for each probability distribution. The exact value of the quantile $F^{-1}_{(w)}(1 - p)$ can be derived for some probability distributions (e.g., normal, student, uniform distribution on an ellipsoid). Often, the probability distribution of the portfolio return is only partially known and the exact value of its quantiles can only be approximated. The Cantelli [2], Chebyshev [2], and Camp–Meidell [12] probability inequalities can be used for approximating the value of the quantile, which allows for the formulation of a convex inner approximation of the probabilistic optimization problem **PMV**.

Theorem 5 (Bonami and Lejeune [2]). *Let the first two moments of the probability distribution of the portfolio return be finite. The second-order cone constraint*

$$\mu^T w - \sqrt{\frac{p}{1-p}}\sqrt{w^T \Sigma w} \geq R$$

is a valid inner approximation of the probabilistic constraint $\mathbb{P}\left(\xi^T w \geq R\right) \geq p$ (5).

Theorem 6 (Lejeune [12]). *Let the probability distribution of the portfolio return have finite first and second moments. The second-order cone constraints*

$$\mu^T w - \sqrt{\frac{1}{2(1-p)}}\sqrt{w^T \Sigma w} \geq R \tag{16}$$

and

$$\mu^T w - \sqrt{\frac{2}{9(1-p)}}\sqrt{w^T \Sigma w} \geq R. \tag{17}$$

are valid inner approximations of the probabilistic constraint (5) *when the probability distribution of* $\xi^T w$ *is*

- *Symmetric [for* (16)*]*
- *Symmetric and unimodal [for* (17)*]*

The constraints (16) and (17) are respectively derived from the one-sided Chebyshev inequality for symmetric probability distribution:

$$\mathbb{P}(X - \mu \geq a) \leq \frac{\sigma^2}{2a^2}, \tag{18}$$

and from the Camp–Meidell's inequality

$$P(X - \mu \geq a) \leq \frac{2\sigma^2}{9a^2}. \tag{19}$$

The expression $\min[a, b]$ refers to the minimum value of a and b. The three inner approximations of (5) proposed above are second-order cone constraints and define a convex feasible area.

3 Combinatorial Constraints for Asset Allocation

Asset allocation policies are often subjected to specific constraints not present in the standard formulation of the mean–variance model. In this section, we describe some of the most common restrictions faced by asset allocation managers and modeled with combinatorial constraints (see, e.g., [2, 10, 18]). The introduction of integer variables forces to solve problem **C0**, **C1**, **E1**, **C2**, or **E2** multiple times (once at each node of a branch-and-bound tree). Filomena and Lejeune's previous computational study [7] has shown that the formulation **E2** can be solved the fastest and all subsequent formulations will be based on **E2**. We refer to [6] for a presentation of other trading restrictions (e.g., proportional transaction costs, liquidity, tracking and turnover constraints) whose consideration does not require the introduction of integer decision variables.

3.1 Diversification Constraints

For diversification purposes, asset managers are sometimes required to hold a representative position (i.e., at least equal to a specified level $s_{\min}$) in a minimal number $l_{\min}$ of asset classes, industrial sectors, and/or securities. Let l be the number of industrial sectors in which positions can be hold. Every security j is associated to a particular sector k. The sets $S_k, k = 1, \ldots, l$ of assets affiliated with a sector k form a partition of $\{1, \ldots, n\}$. Let ζ_k be a binary variable taking value 1 if the portfolio includes a position in sector k and taking value 0 otherwise. The feasible set defined by the diversification constraint reads

$$G_D = \left\{(w, \zeta) \in \mathcal{R}^n_+ \times \{0,1\}^l : s_{\min}\, \zeta_k \leq \sum_{j \in S_k} w_j \leq s_{\min} + (1 - s_{\min})\ \zeta_k, k = 1, \ldots l,\ e^T \zeta \geq l_{\min}\right\}. \tag{20}$$

The last constraint in (20) is a cardinality constraint that requires to hold positions in at least $l_{\min}$ sectors.

3.2 Cardinality Constraints

Besides enforcing diversification goals, cardinality constraints are also needed when the asset manager is tasked (e.g., index tracking fund) to track a market benchmark and reproduce its behavior with a limited number of securities. Under such circumstances, the cardinality constraints define an upper bound $u_{\max}$ on the number of positions that can be held. Let $\delta_j \in \{0, 1\}, j = 1$ denote a binary variable taking value 1 if one invests in asset j and taking value 0 otherwise. The feasible set defined by such cardinality constraints is given by

$$G_C = \left\{(w, \delta) \in \mathcal{R}_+^n \times \{0,1\}^n : w \le \delta \ , \ e^T\delta \le u_{\max}\right\}. \tag{21}$$

Cardinality constraints limit the number of assets included in the portfolio, thereby controlling the transaction costs. However, they are not sufficient to exclude the occurrence of small trades and are therefore often juxtaposed to buy-in threshold constraints.

3.3 Buy-In Threshold Constraints

Small positions have typically little impact on the performance of the portfolio, but have weak liquidity and can be costly in terms of brokerage fees, bid-ask spreads, monitoring costs, etc. Buy-in threshold constraints prevent the holding of a position if one does not invest at least $w_{\min}$ which represents the minimal allowable position size. The feasible set defined by the buy-in threshold constraints reads

$$G_B = \left\{(w, \delta) \in \mathcal{R}_+^n \times \{0,1\}^n : w \le \delta \ , \ w_{\min}\, \delta \le w\right\}. \tag{22}$$

3.4 Transaction Round Lot Constraints

Institutional investors have often to transact securities in large lots or batches of M_j (e.g., 100, 500) shares. The formulation of the round lot trading requirements requires the introduction of general integer decision variables here represented by the n-dimensional vector γ. Let p_j denote the face value of stock j and V be the available capital. The feasible set defined by the round lot transaction constraints is

$$G_L = \left\{(w, \gamma) \in \mathcal{R}_+^n \times \mathcal{Z}^n : w_j = \frac{p_j \gamma_j M_j}{V} \ , \ j = 1, \ldots, n\right\}, \tag{23}$$

and imposes that the number $\gamma_j M_j$ of shares of any asset j in the portfolio is a multiple of M. The impact of round lot transaction constraints on the structure of the portfolio can be very marked when the asset prices are large relative to the size of the trade. Transaction round lot constraints force to buy shares in large lots, thereby eliminating the risk of holding small positions.

3.5 Fixed Transaction Cost Constraints

Transaction costs caused, for example, by brokerage fees, liquidity costs, fund loans, and tax [15] are omitted in the standard mean–variance formulations. However, they can substantially reduce the returns of a portfolio. Accounting for them is likely to reduce the volume of trading and rebalancing operations. While sometimes modeled as proportional to the amount traded [7, 19] to represent the gap between bid and ask prices, transaction costs often involve a fixed component invariant with the transacted amount [8, 15]. We shall present a formulation of transaction costs which includes a proportional and a fixed component. We refer the reader to [16] for a comprehensive overview of functional forms used for transaction costs.

Let w^+ and w^- denote the portfolio rebalancing quantities: w_j^+ is the portion of capital used to purchase j and w_j^- is the portion of capital obtained by selling shares of asset j when constructing a new portfolio or rebalancing an existing one. We denote by y_j the net position in security j after rebalancing the portfolio:

$$w = w^0 + w^+ - w^- \tag{24}$$

$$w^- \leq w^0 \tag{25}$$

$$w, w^+, w^- \geq 0 \tag{26}$$

The set of equalities (24) are balance constraints that ensure that the rebalanced position w_j in asset j is equal to the initial position w_j^0 increased (resp., decreased) by the purchased (resp., sold) shares w_j^+ (resp. w_j^-) of j. The set of constraints (25) preclude short selling when rebalancing. The proportional linear transaction cost associated to security j is c_j and the fixed transaction cost is d_j equal to $\frac{f}{V}$, where f is the fixed cost to invest in an asset and V is the value of the portfolio. The probabilistic portfolio optimization model **PMV** with fixed and proportional transaction costs takes the following form:

$$\min \; h \tag{27}$$

$$\text{subject to} \quad \sum_{j=1}^{n} \Big(\mu_j w_j - c_j (w_j^+ + w_j^-) - d_j \delta_j \Big) - R \geq h \tag{28}$$

$$\sum_{j=1}^{n} \Big(w_j + c_j (w_j^+ + w_j^-) + d_j \delta_j \Big) = 1 \tag{29}$$

$$w^+ + w^- \leq \delta \tag{30}$$

$$\delta \in \{0, 1\}^n \tag{31}$$

$$(10); (13); (14); (24); (25); (26).$$

Constraint (28) requires the net portfolio return (after deduction of the transaction costs) to exceed the specified return level R. The budget constraint (29) accounts for the wealth invested in each asset and the fixed and proportional transaction costs. Constraints (30) impose each binary variable δ_j in (31) to take value 1 if the investor holds a position in the corresponding asset j. Note that w^0 is here a vector of known nonnegative parameters. If one builds a new portfolio, then each component of w^0 and w^- are equal to 0.

Note that one can also add some buy-in threshold constraints so that the net position w_j after rebalancing as well as the amounts (w_j^+ and w_j^-) traded be significant and at least equal to the minimum prescribed:

$$w \leq \delta \tag{32}$$

$$w^+ \leq \delta^+ \tag{33}$$

$$w^- \leq \delta^- \tag{34}$$

$$w_{\min}\, \delta \leq w_j \tag{35}$$

$$w_{\min}\, \delta^+ \leq w_j^+ \tag{36}$$

$$w_{\min}\, \delta^- \leq w_j^- \tag{37}$$

$$\delta^-, \delta^+ \in \{0, 1\}^n \tag{38}$$

As an alternative to transaction cost constraints, practitioners sometimes employ turnover constraints that can be linearized and can limit the turnover, and thus the transaction costs, on each individual security or over the entire portfolio.

4 Outer Approximation Algorithm

The introduction of one or more of the combinatorial constraints presented in Sects. 3.1–3.5 in one of the deterministic formulations equivalent to **PMV** gives a mixed-integer nonlinear programming (MINLP) formulation. Its continuous relaxation is convex and takes the form of a second-order cone programming problem. We illustrate the method with problem **PMVC** that includes diversification, cardinality, buy-in threshold, and concentration constraints. Let G_{DCL} be the mixed-integer set defined as $G_{DCL} = G_D \bigcap G_C \bigcap G_L$. Problem **PMVC** reads

$$\textbf{PMVC}: \quad \min \; h$$

$$\text{subject to} \quad (8); (10); (13); (14)$$

$$w \leq w_{\max} \tag{39}$$

$$\sum_{j \in S_k} w_j \leq L_{\max},\; l = 1, \ldots, l \tag{40}$$

$$(w, \delta, \zeta) \in \mathcal{X} \bigcap G_{DCL}. \tag{41}$$

Constraints (39) and (40) are concentration constraints and ensure that a position in an asset (39) or in a sector (40) does not exceed a specified limit $u_{\max}$ or $l_{\max}$.

4.1 Algorithm Description

We shall now present an outer approximation algorithm to solve the reformulation of the probabilistic portfolio optimization problem **PMVC** with combinatorial constraints. The proposed algorithm is a derivative of the solution method proposed by [13] and used to construct risk-averse enhanced index funds. The outer approximation algorithm provides a hierarchical organization of the computations in which the set of binary-restricted variables is expanded at each iteration and converges to the exact solution within a finite number of iterations.

At each iteration t, we remove the integrality conditions on some of the binary variables δ and reformulate the cardinality constraint in G_C (21) to obtain an outer approximation of problem **PMVC**. The set $\{1, \ldots, n\}$ of binary variables is partitioned into the subsets $S_1^{(t)}$ and $S_2^{(t)}$, and the integrality restrictions are relaxed on the variables in $S_2^{(t)}$. The partitioning is designed so that the set of variables with integrality restrictions is small enough to make the outer approximation problem easy to solve and large enough to contain, with a high probability, the assets included in the optimal portfolio. The solution of the outer approximation problem gives a lower bound on the optimal solution of the original problem **PMVC** and is used to define the stopping criterion. If the optimal solution of the approximation problem is not feasible for **PMVC**, we tighten the approximation problem for the next iteration $(t + 1)$ by expanding the set of variables on which the binary restrictions are enforced: $S_1^{(t)} \subseteq S_1^{(t+1)}$. Hence, we obtain a series of increasingly tighter outer approximations, which ensures the finite convergence of the algorithm. Note that Coleman et al. [3] have also proposed an outer approximation approach for the solution of deterministic asset allocation problems including a cardinality constraint. While the proposed algorithm solves a finite number of mixed-integer nonlinear programming problems containing a small set of binary variables, Coleman et al. [3] solve a series of continuous outer approximation problems. Our outer approximation formulations are obtained by relaxing the integrality restrictions on some binary variables, whereas Coleman et al. [3] derive continuous outer approximations that proxy the function counting the number of assets in the constructed fund. The next subsections detail the structure of the algorithm. The reader is referred to [1] for a recent and comprehensive review of the available MINLP solvers and algorithms used to solve MINLP problems.

4.2 Algorithm Structure

The proposed exact outer approximation algorithm involves an initialization (Sect. 4.2.1) and an iterative (Sect. 4.2.2) phases.

4.2.1 Initialization

The successive outer approximation problems are obtained by relaxing the integrality requirements on a subset of the binary variables δ. Let us denote by G^r_{CL} the continuous relaxation of the MIP set $G_{CL} = G_C \bigcap G_L$. The initial outer approximation problem $\mathbf{OA}^{(0)}$:

$$\begin{aligned} \mathbf{OA}^{(0)}: \quad & \min \quad h \\ & \text{subject to} \quad (8); (10); (13); (14); (39); (40) \\ & \qquad (w, \zeta) \in \mathcal{X} \bigcap G_D \qquad (42) \\ & \qquad (w, \delta) \in \mathcal{X} \bigcap G^r_{CL}. \qquad (43) \end{aligned}$$

is a partial continuous relaxation of **PMVC**. Indeed, we maintain the integrality restrictions on the vector ζ, which is of small dimension equal to the number of considered industrial sectors. The notation $\delta^{*(t)}$ refers to the value taken by δ in the optimal solution of the outer approximation problem $\mathbf{OA}^{(t)}$. The same notation style will be used for all decision variables.

If the optimal values δ in $\mathbf{OA}^{(0)}$ are integer, then the optimal solution $(w^{*(0)}, \delta^{*(0)}, \zeta^{*(0)})$ of $\mathbf{OA}^{(0)}$ is optimal for **PMVC** too, and we stop. Otherwise, we start the iterative process and partition the set of securities into:

$$S_1^{(0)} = \{j : w_j^{*(0)} > 0, \ j = 1, \ldots, n\} \qquad (44)$$

$$S_2^{(0)} = \{j : w_j^{*(0)} = 0, \ j = 1, \ldots, n\}. \qquad (45)$$

The integrality restrictions on the binary variables δ_j, $j \in S_1^{(0)}$ are maintained, while they are removed on those in $S_2^{(0)}$. The idea is to maintain the integrality restrictions on the variables associated with the assets in which the investor would hold positions if the cardinality constraint $e^T\delta \leq u_{\max}$ in G_C was absent. Prior experiments (see also [13]) revealed that a security j with variable w_j taking a positive (resp., null) value in the optimal solution of the relaxed problem is likely (resp., unlikely) to be in the optimal portfolio. Stated differently, the optimal solution of $\mathbf{OA}^{(0)}$ provides key information about the structure of the optimal portfolio.

We shall now use the initial outer approximation problem $\mathbf{OA}^{(0)}$ to derive an optimality cut, also called objective level cut [14], for problem **PMVC**. We construct a portfolio by only using the assets included in the set $S_1^{(0)}$ whose composition (44) is determined by the optimal solution of $\mathbf{OA}^{(0)}$. This portfolio is obtained by the nonlinear optimization problem **IA**:

$$\begin{aligned} \mathbf{IA}: \quad & \min \quad h \\ & \text{subject to} \quad (8); (10); (13); (14); (39); (40); (41) \\ & \qquad w_j = \delta_j = 0, \ j \in S_2^{(0)}. \qquad (46) \end{aligned}$$

Problem **IA** is easier to solve than **PMVC**, since the continuous relaxation of **IA** is also a second-order cone programming problem, but **IA** has many of its decision variables fixed (46). The asset universe of problem **IA** is restricted to the assets included in $S_1^{(0)}$ and is a subset of the original asset universe used for problem **PMVC**. Proposition 2 follows immediately.

Proposition 2. *The nonlinear optimization problem* **IA** *is an inner approximation of problem* **PMVC**. *Let* $\mathbf{u}^*_{\mathbf{IA}}$ *be the optimal value of* **IA**. *The inequality*

$$h \le \mathbf{u}^*_{\mathbf{IA}} \tag{47}$$

is an objective level cut for **PMVC**.

Evidently, every feasible solution for **IA** is feasible for **PMVC** and the optimal value $\mathbf{u}^*_{\mathbf{IA}}$ of **IA** is an upper bound on the optimal value of **PMVC**. Therefore, the cut (47) can eliminate feasible, yet nonoptimal solutions for **PMVC**. As demonstrated in [13], the introduction of such a cut is highly beneficial in speeding up the finding of the optimal solution.

4.2.2 Iterative Process

The optimal solution of the outer approximation problem $\mathbf{OA}^{(t-1)}$ is used to update the composition of the sets $S_1^{(t)}$ and $S_2^{(t)}$ at the current iteration t:

$$S_1^{(t)} = S_1^{(t-1)} \bigcup \left\{ j : w_j^{*(t-1)} > 0, j \in S_2^{(t-1)} \right\} \tag{48}$$

$$S_2^{(t)} = \{1, 2, \ldots, n\} \setminus S_1^{(t)}. \tag{49}$$

The formulation of the current outer approximation problem $\mathbf{OA}^{(t)}$ is based on the set updating process (48) and (49). Problem $\mathbf{OA}^{(t)}$ is such that (i) a binary variable δ_j is associated with each asset j in $S_1^{(t)}$ (see (51) and (56)); (ii) there is no binary variable individually assigned to any of the assets in $S_2^{(t)}$. Instead, we associate a binary variable δ_S with the group of assets in $S_2^{(t)}$ [see (52) and (57)]: δ_S takes value 1 if any variable w_j, $j \in S_2^{(t)}$ is strictly positive.

$$\mathbf{OA}^{(t)}: \quad \min \quad h \tag{50}$$

$$\text{subject to} \quad (3); (8); (10); (13); (14); (39); (40); (42)$$

$$w_j \le \delta_j, \; j \in S_1^{(t)} \tag{51}$$

$$\sum_{j \in S_2^{(t)}} w_j \le \delta_S \tag{52}$$

$$\sum_{j \in S_1^{(t)}} \delta_j + \delta_S \le u_{\max} \tag{53}$$

$$w_{\min}\delta_j \le w_j, \; j \in S_1^{(t)} \tag{54}$$

$$w_{\min}\delta_S \le \sum_{j \in S_2^{(t)}} w_j \tag{55}$$

$$\delta_j \in \{0, 1\}, \; j \in S_1^{(t)} \tag{56}$$

$$\delta_S \in \{0, 1\}. \tag{57}$$

The approximation problem $\mathbf{OA}^{(t)}$ contains $(|S_1^{(t)}| + 1 + l)$ binary variables, while the original problem contains $(n+l)$ of them. In the next $(t+1)$ iteration, the assets j for which $w_j^{(t)*} > 0$ are moved to $S_1^{(t+1)}$ and a new binary variable is introduced for each of them.

As demonstrated in [13], the outer approximation algorithm outlined above exhibits the following properties.

Proposition 3. *Problem* $(\mathbf{OA})^{(t)}$ *is an outer approximation of problem* **PMVC**.

This can be easily seen by observing that the set $\{(w, \delta, \delta_S) : (51)–(57)\}$ is a relaxation of the set G_{CL}.

Proposition 4. *The optimal solution* $(w^{*(t)}, \delta^{*(t)}, \delta_S^{*(t)}, \zeta^{*(t)})$ *of* $\mathbf{OA}^{(t)}$ *is optimal for* **PMVC** *if*

$$|Q| \le u_{\max}\,, \quad \textit{with} \quad Q = \{j : w_j^{*(t)} > 0, j = 1, \ldots, n\}. \tag{58}$$

The above condition (58) is used as the stopping criterion for the algorithmic process.

Proposition 5. *The outer approximation algorithm generates a series of increasingly tighter outer approximations.*

It follows immediately from the updating process of the sets $S_1^{(t)}$ and $S_2^{(t)}$ with (48) and (49). At t, all the integer variables in $S_1^{(t-1)}$ defined as binary at $(t-1)$ remain defined as binary, while at least one of the integer variables in $S_2^{(t-1)}$ on which the integrality restriction was relaxed at $(t-1)$ is binary at t.

Proposition 6. *The outer approximation algorithm has the finite convergence property. It terminates after at most (n-$u_{\max}$-1) iterations.*

If $|S_1^{(0)}| \le u_{\max}$, the solution is optimal for **PMVC**. If $|S_1^{(0)}| > u_{\max}$, another iteration is needed, in which the integrality restrictions are restored on at least one of the $\delta_j, j \in S_2^{(0)}$ that were so far defined as continuous. Thus, it is clear that the algorithm finds the optimal solution of **PMVC** in at most (n-$u_{\max}$-1) iterations. Each iteration consists in the solution of a second-order cone optimization problem with binary variables that optimization solvers such as Cplex or Gurobi can solve with algorithms based on interior point methods and branch-and-bound techniques.

Proposition 5 indicates that each new outer approximation problem has a larger number of binary variables and is likely more challenging to solve. For the algorithm to be efficient, it is crucial that the number of iterations remains small. This underlines the importance of the selection of the variables for which integrality restrictions are enforced. The pseudo-code of the outer approximation algorithm follows.

Pseudo-Code of Outer Approximation Algorithm

Initialization:
$t := 0$;
Solve the continuous relaxation $\mathbf{OA}^{(0)}$ of **PMVC**. Let
$(w^{*(0)}, \delta^{*(0)}, \zeta^{*(0)}) := argmin(\mathbf{OA}^{(0)})$;
Define $S_1^{(0)}$ and $S_2^{(0)}$ according to (44) and (45);
if $|S_1^{(0)}| \leq u_{max}$ **then**
  $(w^{*(0)}, \delta^{*(0)}, \zeta^{*(0)})$ is optimal for **PMVC**
end
else
  Construct and solve inner approximation problem **IA**;
  Generate optimality cut (47) and insert it in $\mathbf{OA}^{(0)}$;
  Iterative Process:
  repeat
    $t := t + 1$;
    Let $(w^{*(t-1)}, \delta^{*(t-1)}, \delta_S^{*(t-1)}, \zeta^{*(t-1)}) := argmin(\mathbf{OA}^{(t-1)})$;
    Update $S_1^{(t)}$ and $S_2^{(t)}$ as in (48) and (49), respectively ;
    Solve $\mathbf{OA}^{(t)}$;
  until $|Q| \leq u_{max}$;
end

---

5 Computational Results

5.1 Testbed

For the period spanning from January 1999 to December 2010, we have collected the monthly returns of more than 2,000 stocks from the CRSP US Monthly Stock Database available through the Wharton Research Data Service. Each stock is traded on NYSE and NASDAQ and does not have any missing observation for its

Table 1 Computational results

	ATP		NP		NF		AIG		AOG		NB		NFA	
n	CP	OA	CP	OA	CP	OA	CP	OA	CP	OA	CP	OA	CP	OA
250	208.5	187.6	8	8	8	8	0	0	0	0	8	8	2	7
500	1,660.9	1,270.1	6	6	7	8	0.150 %	0.023 %	0.008 %	0	7	8	2	6
750	2,387.6	1,998.5	5	6	6	8	0.205 %	0.068 %	0.024 %	0	6	8	0	8
1,000	3,222.6	2,797.3	2	4	3	7	0.621 %	0.157 %	0.165 %	0.004 %	4	8	1	7

return data over the considered period. We use the data item called "monthly price alternate" that provides the last available stock price of the month and accounts for splits and dividends. We have built four families of problem instances that differ in the size n (i.e., $n = 250$, 500, 750, and 1,000) of the market universe. For each family, we have generated eight problem instances. For each of them, the securities included in the asset universe have been selected randomly. We have arbitrarily fixed the number $u_{\max}$ of assets that can be in the portfolio to 20.

To evaluate the computational benefits of the proposed algorithm, we have benchmarked the proposed algorithmic OA with the default branch-and-bound algorithm (and default options) of the Cplex 12.4 solver, referred to as CP. Each problem instance has been modeled with the AMPL modeling language and solved with both the OA and CP algorithms on a 64-bit Dell Optiplex 990 Workstation with Quad Core Intel Processor i7-2600 3.40GHz CPU and 16GB of RAM. We shall now use the recorded information to analyze the quality of the solution obtained with and the computational tractability of the proposed hierarchical outer approximation algorithm.

5.2 Algorithmic Efficiency

For each family of problem instance, Table 1 displays the average time ATP (in CPU seconds) to prove optimality, the number NP of instances for which optimality is proven, the number NF of instances for which the optimal solution is reached, the average integrality AIG and optimality AOG gaps, the number NB of instances for which the algorithm obtains the best solution, and the number NFA of instances for which the algorithm is the fastest to find the best solution. We define the optimality gap OG (resp., integrality gap IG) as the normalized difference between the best solution z^b found in our hour and the known optimal solution z^* (resp., best lower bound L):

$$OG = \frac{z^b - z^*}{z^b} \quad \text{and} \quad IG = \frac{z^b - L}{z^b} .$$

Since we solve a minimization problem,and the lower bound is the solution of a relaxation of this problem, we have that $IG \geq OG$. The complete results are given in Table 2 in Appendix. One hour of CPU time is allowed for the solution of each problem. If the optimal solution is not reached within one hour, we set the time equal to 3600 CPU seconds. If the time differential between the two algorithms is less than one second, we consider them equally fast. An entry with 0 in Tables 1 and 2 means that the gap is below 10^{-6}. In order to obtain some of the results displayed in Table 1, the knowledge of the optimal solution for each instance is needed. We obtained it by letting the algorithms running for more than an hour when required.

It appears that the outer approximation algorithm is, on average, at least 10% faster than `CP` for each family of problem instances. The `OA` algorithm is the fastest to reach the best solution for 87.5% of the problem instances. The `OA` algorithm also outperforms `CP` in terms on the number of instances for which (i) optimality is proven, and (ii) the optimal solution is reached. It is therefore not surprising that the average optimality and integrality gaps are smaller with `OA` than with `CP`. Columns 12 and 13 show that `OA` finds a better solution than `CP` for 21.875% of the instances and, in particular, for 50% of the largest instances with 1,000 assets.

Clearly, the proposed outer approximation algorithm permits a more efficient and faster solution of the analyzed instances. It is due to its ability to detect early the structure of the problem and, more precisely, to identify the assets that are likely to be included in the optimal portfolio. Indeed, the outer approximation algorithm requires two or more iterations for only two of the thirty-two problem instances.

6 Conclusion

We have studied a probabilistic portfolio optimization problem in which trading restrictions modeled with combinatorial constraints are accounted for. We have provided five deterministic reformulations equivalent to this stochastic programming optimization problem and have discussed their computational efficiency. We have used an exact outer approximation algorithm to solve multiple instances of the reformulated MINLP problem. The algorithm is based on the early recognition of the problem structure and permits a hierarchical organization of the computations. Computational tests show the contribution of the propose algorithm which outperforms the Cplex 12.4 solver in terms of computational times and quality of the obtained solutions. The algorithmic procedure can be easily extended to other asset allocation models with combinatorial constraints.

Appendix

Table 2 Computational results for thirty-two problem instances

	Time in CPU seconds		Integrality gap		Optimality gap	
n	CP	OA	CP	OA	CP	OA
250-1	246.4	259.3	0	0	0	0
250-2	158.2	157.6	0	0	0	0
250-3	331.3	274.1	0	0	0	0
250-4	123	118	0	0	0	0
250-5	111.2	89.6	0	0	0	0
250-6	198.6	130.9	0	0	0	0
250-7	220.7	189.3	0	0	0	0
250-8	278.5	282.3	0	0	0	0
500-1	612.7	207.6	0	0	0	0
500-2	1,128.6	441.8	0	0	0	0
500-3	2,456.3	737.9	0	0	0	0
500-4	546.9	214.9	0	0	0	0
500-5	3,600	3,600	0.369%	0.087%	0	0
500-6	897.2	904.1	0	0	0	0
500-7	3,600	3,600	0.833%	0.097%	0.066%	0
500-8	445.3	454.1	0	0	0	0
750-1	3,600	2,987.2	0.572%	0	0.140%	0
750-2	1,065.3	1,008.3	0	0	0	0
750-3	1,048.9	642.8	0	0	0	0
750-4	1,400.3	781.1	0	0	0	0
750-5	2,759.8	1,902.3	0	0	0	0
750-6	3,600	3,600	0.768%	0.286%	0.050%	0
750-7	3,600	3,600	0.298%	0.255%	0	0
750-8	2,026.3	1,466.2	0	0	0	0
1,000-1	3,600	2,719.1	0.881%	0	0.368%	0
1,000-2	3,600	3,600	0.572%	0.292%	0	0
1,000-3	3,600	3,600	0.897%	0.343%	0.223%	0
1,000-4	2,458.3	2,473.2	0	0	0	0
1,000-5	3,600	1,597.1	0.832%	0	0.259%	0
1,000-6	3,600	3,600	0.891%	0.361%	0.396%	0
1,000-7	3,600	3,600	0.897%	0.259%	0.076%	0.031%
1,000-8	1,722.3	1,189.1	0	0	0	0

References

1. Bonami, P., Kılınç, M., Linderoth, J.: Algorithms and software for convex mixed integer nonlinear programs. IMA Vol. Math. Appl. **154**, 1–39 (2012)
2. Bonami, P., Lejeune, M.A.: An exact solution approach for portfolio optimization problems under stochastic and integer constraints. Oper. Res. **57**(3), 650–670 (2009)

3. Coleman, T.F., Li, Y., Henniger, J.: Minimizing tracking error while restricting the number of assets. J. Risk **8**, 33–56 (2006)
4. Cornuejóls, G., Tütüncü, R.: Optimization Methods in Finance. University Press, Cambridge (2007)
5. Dentcheva, D., Ruszczyński, A.: Portfolio optimization with stochastic dominance constraints. J. Bank. Financ. **30**(2), 433–451 (2006)
6. Fabozzi, F.J., Kolm, P.N., Pachamanova, D., Focardi, S.M.: Robust Portfolio Optimization and Management. Wiley, Hoboken, New Jersey, USA (2007)
7. Filomena, T., Lejeune, M.A.: Stochastic portfolio optimization with proportional transaction costs: convex reformulations and computational experiments. Oper. Res Lett. **40**(1), 207–212 (2012)
8. Filomena, T., Lejeune, M.A.: Warm-start heuristic for stochastic portfolio optimization with fixed and proportional transaction costs. J. Optim. Theorey Appl. http://link.springer.com/article/10.1007%2Fs10957-013-0348-y (2013)
9. Guerard, J.B. Jr.: Handbook of Portfolio Construction - Contemporary Applications of Markowitz Techniques. Springer (2010)
10. Jobst, N.J., Horniman, M.D., Lucas, C.A., Mitra, G.: Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints. Quant. Financ. **1**, 1–13 (2001)
11. Kataoka, S.: A stochastic programming model. Econometrica **31**(1–2), 181–196 (1963)
12. Lejeune, M.A.: A VaR Black-Litterman Model for the construction of absolute return fund-of-funds. Quant. Financ. **11**, 1489–1501 (2011)
13. Lejeune, M.A., Samatli-Paç, G.: Construction of risk-averse enhanced index funds. INFORMS J. Comput. http://joc.journal.informs.org/content/early/2012/11/26/ijoc.1120.0533.abstract (2013)
14. Li, D., Sun, X.L., Wang, J.: Optimal lot solution to cardinality constrained mean-variance formulation for portfolio selection. Math. Financ. **16**(1), 83–101 (2006)
15. Lobo, M.S., Fazel, M., Boyd, S.: Portfolio optimization with linear and fixed transaction costs. Ann. Oper. Res. **152**, 341–365 (2007)
16. Maringer, D.: Transaction costs and integer constraints. In: Amman, H.M., Rustem B., (eds.) Portfolio Management with Heuristic Optimization - Advances in Computational Management Science. Springer, Dordrecht, The Netherlands **8**, 77–99 (2005)
17. Markowitz, H.M.: Portfolio selection. J. Financ. **7**, 77–91 (1952)
18. Mitra, G., Ellison, F., Scowcroft, A. Quadratic programming for portfolio planning: Insights into algorithmic and computation issues. Part II: processing of portfolio planning models with discrete constraints. J. Asset Manag. **8**, 249–258 (2007)
19. Muthuraman, K., Sunil, K.: Multidimensional portfolio optimization with proportional transaction costs. Math. Financ. **16**, 301–335 (2006)
20. Roy, A.D.: Safety first and the holding assets. Econometrica **20**, 431–449 (1952)
21. Schultz, R., Tiedemann, S.: Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse. SIAM J. Optim. **14**, 115–138 (2006)
22. Welsch, R.E., Zhou, X.: Application of robust statistics to asset allocation models. Revstat **5**, 97–114 (2007)

http://www.springer.com/978-1-4614-8986-3

Modeling and Optimization: Theory and Applications
Selected Contributions from the MOPTA 2012
Conference
Zuluaga, L.F.; Terlaky, T. (Eds.)
2013, VII, 136 p. 10 illus., Hardcover
ISBN: 978-1-4614-8986-3