

IBM Research Report

An Exact Solution Approach for Portfolio Optimization Problems under Stochastic and Integer Constraints

P. Bonami

IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

M. A. Lejeune

Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213



Research Division

Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

An Exact Solution Approach for Portfolio Optimization Problems under Stochastic and Integer Constraints

P. Bonami*, M.A. Lejeune †

Abstract

In this paper, we study extensions of the classical Markowitz' mean-variance portfolio optimization model. First, we consider that the expected asset returns are stochastic by introducing a probabilistic constraint imposing that the expected return of the constructed portfolio must exceed a prescribed return level with a high confidence level. We study the deterministic equivalents of these models. In particular, we define under which types of probability distributions the deterministic equivalents are second-order cone programs, and give exact or approximate closed-form formulations. Second, we account for real-world trading constraints, such as the need to diversify the investments in a number of industrial sectors, the non-profitability of holding small positions and the constraint of buying stocks by lots, modeled with integer variables. To solve the resulting problems, we propose an *exact* solution approach in which the uncertainty in the estimate of the expected returns and the integer trading restrictions are *simultaneously* considered. The proposed algorithmic approach rests on a non-linear branch-and-bound algorithm which features two new branching rules. The first one is a static rule, called *idiosyncratic risk* branching, while the second one is dynamic and called *portfolio risk* branching. The proposed branching rules are implemented and tested using the open-source framework of the solver Bonmin. The comparison of the computational results obtained with standard MINLP solvers and with the proposed approach shows the effectiveness of this latter which permits to solve to optimality problems with up to 200 assets in a reasonable amount of time. *Keywords:* Programming: stochastic, integer: nonlinear, branch-and-bound, Finance: portfolio; Probability: distributions

1 Introduction

Since Markowitz' groundbreaking work in portfolio selection [22], portfolio optimization has been receiving sustained attention from both asset liability professionals and academics. All such studies define a portfolio optimization criterion such as mean-variance, mean absolute deviation, value-at-risk, conditional value-at-risk, stochastic dominance of first and second order, etc. In this paper, we use the mean-variance approach that studies how risk-averse investors can construct optimal portfolios taking into consideration the trade-off between market volatility and expected returns. Out of a universe of r risky assets and one non-risky asset characterized by a known return μ_0 that usually reflects the interest rate on the money market, an efficient frontier of optimal portfolios can be constructed. Portfolios on the efficient frontier offers the maximum possible expected return for a given level of risk.

*IBM TJ Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598, USA; pbonami@us.ibm.com

†Tepper School of Business, Carnegie Mellon University, Pittsburgh PA 15213, USA; mlejeune@andrew.cmu.edu

The original Markowitz' model assumes that the expected returns $\mu \in \mathcal{R}^r$ of the risky assets and the variance-covariance matrix $\Sigma \in \mathcal{R}^{r \times r}$ of the returns are known. One of the several formulations of the mean-variance portfolio selection problems involves the construction of a portfolio with minimal risk provided that a prescribed return level R is attained. This model is formulated by the following mathematical program:

$$\begin{aligned} & \min w^T \Sigma w \\ & \text{subject to } \mu_0 w_0 + \mu^T w \geq R \\ & \sum_{j=0}^r w_j = 1 \\ & w \in \mathcal{R}^{r+1} \end{aligned} \quad (1)$$

In the problem above, the decision variables $w_j, j = 1, \dots, r$ represent the proportion of capital invested in the risky asset j while w_0 is the fraction of capital invested in the money market. The objective function aims at minimizing the variance of the portfolio $w^T \Sigma w$, and the constraint

$$w_0 + \sum_{j=1}^r w_j = 1 \quad (2)$$

enforces that the sum of the investments is equal to 1. Clearly, the investor can allocate part of the available capital K to the money market w_0 .

In the last decade, there has been much effort devoted to extending Markowitz' work and making his modern portfolio theory more practical. In this study, we propose models that account for two limitations associated with the mean-variance approach, namely (i) the randomness in the parameters describing the model and (ii) some of the trading restrictions of stock markets.

The classical mean-variance framework relies on the perfect knowledge of the expected returns of the assets and their variance-covariance matrix. However, these returns are unobservable and unknown. Even obtaining accurate estimates of them is very complicated. Indeed, many possible sources of errors (e.g., impossibility to obtain a sufficient number of data samples, instability of data, differing personal views of decision makers on the future returns [24], etc.) affect their estimation, and lead to what Bawa et al. [3] call *estimation risk* in portfolio selection. This estimation risk has been shown to be the source of very erroneous decisions, for, as pointed in [7, 10], the composition of the optimal portfolio is very sensitive to the mean and the covariance matrix of the asset returns, and minor perturbations in the moments of the random returns can result in the construction of very different portfolios. Decision-makers would often rather trade-off some return for a more secure portfolio that performs well under a wide set of realizations of the random variables. The need for constructing portfolios that are much less impacted by inaccuracies in the estimation of the mean and the variance of the return is therefore clear.

The focus here is on the uncertainty associated with the estimation of the expected returns. It is indeed a widespread belief among portfolio managers, and its was shown in [9], that the the portfolio estimation risk is mainly due to errors in the estimation of the expected return and not so much to errors in the estimation of the variance-covariance matrix [7]. In this paper, we assume that the expected return

is stochastic and characterized by a probability distribution, and we require that the expected return of the portfolio is larger than a given level with a high confidence level. We show that the associated problem takes the form of a *probabilistically constrained problem with random technology matrix* [18, 25] that can be reformulated as a nonlinear optimization problem (not necessarily convex). We define under which conditions and for which classes of probability distributions the deterministic equivalent problem is convex and takes the form of a second-order cone problem. We examine in which cases an exact closed-formulation can be derived. If a closed-form formulation cannot be obtained, we provide convex approximations that are obtained by using Chebychev's inequality [23] and whose tightness depends on the properties of the probability distribution. This convexity analysis of the model gives insights about the applicability and the computational tractability of the proposed model. In related studies, Costa and Paiva [11], Tütüncü and Koenig [31] and Goldfarb and Iyengar [15] have also studied the mean-variance framework in a *robust* context, assuming that the expected return is stochastic. They characterize the parameters involved in the mean and the variance-covariance matrix with specific types (polytopic, box, ellipsoidal) of uncertainty, and build semi-definite or second-order cone programs. In [12], a risk-averse approach is used for the value-at-risk formulation of the optimization problem, in which only partial information about the probability distribution is known.

The need to account for stock market specifics exacerbates the complexity of the portfolio selection problem. Real-life trading restrictions, such as the minimum amount to invest in an asset, the requirements to buy assets in large lots, or the purchase of assets in a minimal number of industrial sectors, are not considered in the classical mean-variance models. In the present study, we consider those requirements that are respectively called *buy-in threshold*, *round lot*, and *diversification* trading constraints. The modeling of such constraints involves the introduction of integer variables and further challenges the computational tractability of the associated problems [10, 29]. In the next paragraph, we proceed to a review of the literature in which the construction of optimal portfolios satisfying such integer constraints is addressed.

Bienstock [5] considers variants of the Markowitz model which features a cardinality constraint and buy-in threshold constraints. He shows that the problem is NP-complete when a cardinality constraint on the number of asset in the portfolio is present. An exact solution framework by branch-and-cut is developed for which computational results on the exact solution of problems with up to 3300 assets are reported. In [17], an exact branch-and-bound solution approach is proposed for problems subject to buy-in threshold, cardinality and round lot constraints. Frangioni and Gentile [14] also consider buy-in threshold constraints, and develop a new family of cutting planes to handle them. Computational results for problems with up to 300 assets are reported. Using mean absolute deviation as optimization criterion, Konno and Yamamoto [19] consider cardinality and fixed transaction cost constraints and solve problems in which up to 54 assets can be included in the portfolio. It is important to remark that all the studies above do not account for the uncertainty in the problem parameters.

To the best of our knowledge, this study is the first one to propose an *exact* solution approach for portfolio optimization problems in which uncertainty in the estimate of the expected return and real-life

market restriction modeled with integer constraints are *simultaneously* considered. The combination of integrality and of a probabilistic constraints makes such problems very difficult to solve. Such problems belong to the family of Mixed Integer Non-Linear Programs (MINLP) for which only very few solvers are available. In this paper, we use the computational framework offered by the open-source mixed-integer non-linear solver Bonmin [6]. We propose a non-linear branch-and-bound algorithmic approach, and we develop two new branching rules, called *idiosyncratic risk* and *portfolio risk* branching rules. Extended computational experiments on problems containing up to 200 assets clearly show the effectiveness and utility of the two new branching rules. The reader will note that, although the results reported in the paper are obtained for one of the variants of the probabilistic Markowitz' model (i.e., risk minimization subject to the attainment of a predefined return level), the proposed solution approach can be easily extended to the other variants.

The paper is organized as follows. In the first part of Section 2, we describe the characteristics of the constraint enforcing that the portfolio return exceeds with a probability p a given prescribed return level. We present the problem formulation and its deterministic equivalent, we study under which condition it is convex, and we propose exact or approximate closed-form formulations of the deterministic equivalent problem. The second part of Section 2 is devoted to the formulation of the integer constraints and models associated with three types of trading restrictions. Section 3 describes the proposed solution approach. Section 4 reports and comments the computational results. Section 5 provides concluding remarks and suggests extensions to the proposed study.

2 Problem formulation and properties

2.1 Probabilistic portfolio optimization model

The proposed portfolio optimization model takes the form of a probabilistically constrained optimization model with random technology matrix. We refer the reader to the seminal papers of Kataoka [18] and van de Panne and Poppe [25] for a first study of such stochastic programming models in applications pertaining to the transportation and diet problems, and to [8, 16, 26, 33] for more recent studies.

We denote by ξ the random vector of return of the r risky assets; ξ has an r -variate distribution with mean

$$\mu = (\mu_1, \mu_2, \dots, \mu_r)^T, \quad \mu_j = E(\xi_j), j = 1, \dots, r,$$

and variance-covariance matrix

$$\Sigma = E[(\xi - \mu)(\xi - \mu)^T].$$

The probabilistic constraint

$$\mathcal{P}\left(\mu_0 w_0 + \sum_{j=1}^r \xi_j w_j \geq R\right) \geq p \tag{3}$$

in which the coefficients ξ multiplying the decision variables w are stochastic and not (necessarily) independent, guarantees that the expected return of the portfolio $\mu_0 w_0 + \sum_{j=1}^r \xi_j w_j$ is above the prescribed

minimal level of return R with a high probability p , typically defined on $[0.7, 1)$.

The stochastic version of Markowitz' mean-variance portfolio optimization problem [22] reads:

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mathcal{P}\left(\mu_0 w_0 + \sum_{j=1}^r \xi_j w_j \geq R\right) \geq p \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& w \in \mathcal{R}_+^{r+1}
\end{aligned} \tag{4}$$

The decision variables are given by the $r + 1$ -dimensional vector w of portfolio positions. We recall that w_0 is the proportion of the available capital K invested in the money market with fixed return μ_0 , w_j , $j = 1, \dots, r$ is the proportion of the capital K invested in the risky asset j , ξ_j is the stochastic return of asset j , Σ is the variance-covariance matrix of the returns, and the objective function $w^T \Sigma w$ represents the variance of the portfolio. In our model, we assume that the variables are positive, not allowing short-selling positions. This constraint can be removed without affecting the nature of the problem.

2.2 Deterministic equivalent

We shall first show that the deterministic equivalent of the probabilistic portfolio optimization model is a nonlinear programming optimization problem.

Defining by $\psi = \frac{\xi^T w - \mu^T w}{\sqrt{w^T \Sigma w}}$ the normalized (i.e., mean 0 and variance 1) random variable representing the expected portfolio return, it follows that

$$\mathcal{P}(\xi^T w \geq R) = \mathcal{P}\left(\psi \geq \frac{R - \mu^T w}{\sqrt{w^T \Sigma w}}\right) = 1 - F\left(\frac{R - \mu^T w}{\sqrt{w^T \Sigma w}}\right), \tag{5}$$

where F is the cumulative probability distribution of the (normalized) portfolio return and F^{-1} is its inverse. Therefore, the probabilistic constraint (3) becomes

$$\begin{aligned}
& 1 - F\left(\frac{R - \mu^T w}{\sqrt{w^T \Sigma w}}\right) \geq p \\
& \Leftrightarrow F\left(\frac{R - \mu^T w}{\sqrt{w^T \Sigma w}}\right) \leq 1 - p \\
& \Leftrightarrow \mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R
\end{aligned} \tag{6}$$

where $F^{-1}(1 - p)$ is the $(1 - p)$ -quantile of F .

The deterministic equivalent of (4) is the following nonlinear optimization problem:

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& w \in \mathcal{R}_+^{r+1}
\end{aligned} \tag{7}$$

In the next-subsections, we shall study under which conditions, i.e. for which classes of probability distributions the above problem is a second-order cone optimization problem (and is therefore convex, and solvable in polynomial time). We shall see that it is not possible to derive an exact closed-form formulation of the second-order cone problem for each probability distribution. We shall, therefore, using Chebychev's inequality, derive closed-form approximations of the second-order cone problem that are valid for some families of probability distributions.

2.2.1 Convexity results

a) Symmetric probability distributions

The probability distribution F of a random variable ξ is symmetric around its mean μ if $P(\xi \geq \mu + b) = P(\xi \leq \mu - b)$, $b \in \mathcal{R}$, and is centrally symmetric if $P(\xi \geq b) = P(\xi \leq -b)$. We provide a more formal definition below.

Definition 2.1 *A probability distribution of an r -variate random vector is centrally symmetric if its density function f is such that $f(A) = f(-A)$ for all Borel sets $A \subseteq \mathcal{R}^r$.*

Theorem 2.2 *If $p \in [0.5, 1)$ and if the probability distribution of $\xi^T w$ is symmetric, the deterministic equivalent $\mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R$ of the probabilistic constraint $\mathcal{P}(\xi^T w \geq R) \geq p$ is a second-order cone constraint.*

Proof. The matrix of variance-covariance Σ is positive semidefinite, and thus the function $\sqrt{w^T \Sigma w}$ is convex. To show that $\mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R$ is a second-order cone constraint whose feasible set is convex, it is enough to prove that the function $\mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w}$ is concave, which is the case if $F^{-1}(1 - p)$ is smaller than or equal to 0.

Since the probability distribution of ξ is symmetric, the probability distribution F of the normalized random variable ψ is centrally symmetric. It follows that $F(0) = 0.5$ (or, equivalently, that $F^{-1}(0.5) = 0$). This, combined with the fact that any cumulative distribution function is increasing, implies that $F^{-1}(1 - p)$, $p \in [0.5, 1)$ is at most equal to 0, which was set out to prove. \square

Clearly, problem (7) minimizes a convex quadratic function over a second-order cone and some linear constraints, and is therefore a convex, second-order cone problem.

b) Positively skewed probability distributions

The skewness is a measure of the asymmetry of a probability distribution of a real-valued random variable [1], and is computed as

$$skew(\xi) = \frac{E[\xi - \mu]^3}{\sigma^3},$$

where μ and σ are respectively the mean and standard deviation of ξ .

The probability distribution F of a random variable is said to be right-skewed or to have positive skewness (left-skewed or negative skewness, respectively) if the right, upper value (left, lower value,

resp.) tail is longer or fatter than the left, lower value (right, upper value, resp.), or, stated differently, if its median m is strictly smaller (larger, resp.) than its mean μ .

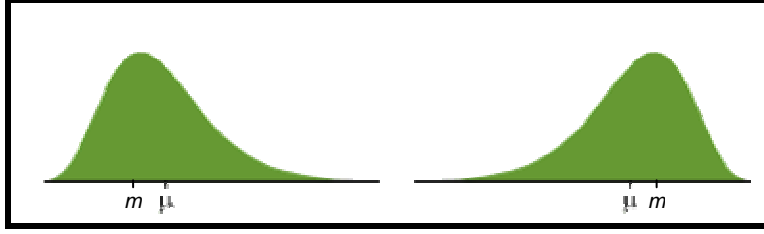


Figure 1: Skewed probability distributions

The two graphs above illustrate the notion of skewness. Both probability distribution functions have the same expectation and variance. The one on the left is positively skewed, while this on the right is negatively skewed.

Definition 2.3 *The probability distribution of an r -variate random vector ξ has positive skewness if*

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

where $E[\psi] = E[\xi - \mu] = 0$ and $F(m) = P(m \geq \psi) = 0.5$.

Theorem 2.4 *If $p \in [0.5, 1)$ and if the probability distribution of $\xi^T w$ has positive skewness, the deterministic equivalent $\mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R$ of the probabilistic constraint $\mathcal{P}(\xi^T w \geq R) \geq p$ is a second-order cone constraint.*

Proof. As mentioned above, $\mu^T w + F^{-1}(1 - p)\sqrt{w^T \Sigma w} \geq R$ is a second-order cone constraint if $F^{-1}(1 - p)(p \geq 0.5)$ is smaller than or equal to 0.

This follows immediately from Definition 2.3:

$$0 > F^{-1}(1 - p), 1 - p \leq 0.5$$

for the probability distribution F of the normalized random variable ψ has positive skewness. □

The exact value of the quantile $F^{-1}(1 - p)$ can be derived for some probability distributions. If we assume the returns of the risky assets to be normally distributed, then the normalized portfolio return ψ has a standard normal cumulative distribution function

$$\phi(\psi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\psi} e^{-t^2/2} dt,$$

and the numerical value of its quantile $\phi^{-1}(p)$ is known. The same applies if the normalized portfolio return is uniformly distributed in an ellipsoid $\Omega = \{\omega = Qz : \|z\| \leq 1\}$ with $\|z\|$ denoting the Euclidian norm of z .

2.2.2 Quantile approximation

The exact value of the $(1-p)$ -quantile $F^{-1}(1-p)$ cannot be derived for each probability distribution F which therefore impedes the derivation of the exact deterministic equivalent of the probabilistic constraint (3) in (4). In this section, using variants of Chebychev's inequality, we derive convex approximations of (3) for different classes of probability distributions. Such approximations are popular in the robust optimization literature [4], and differ in terms of their conservativeness.

Theorem 2.5 *The second-order cone constraint*

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

is a valid approximation of the probabilistic constraint

$$\mathcal{P}(\xi^T w \geq R) \geq p \quad (8)$$

when the portfolio return follows any probability distribution characterized by its first two moments μ and σ^2 .

Proof. Let us consider the random variable Y such that $Y^T w = (2\mu^T - \xi^T)w$: $Y^T w$ has the same mean and variance than $\xi^T w$.

Applying Chebychev's inequality, we obtain

$$\mathcal{P}(Y^T w - \mu^T w > \mu^T w - R) \leq \begin{cases} \frac{1}{1 + \frac{(\mu^T w - R)^2}{w^T \Sigma w}} = \frac{w^T \Sigma w}{w^T \Sigma w + (\mu^T w - R)^2} & \text{if } \mu^T w \geq R \\ 1 & \text{otherwise} \end{cases} . \quad (9)$$

Clearly,

$$\mathcal{P}(Y^T w - \mu^T w > \mu^T w - R) = \mathcal{P}(\mu^T w - Y^T w < R - \mu^T w) = \mathcal{P}(\xi^T w - \mu^T w < R - \mu^T w) .$$

This, combined with (9), successively implies that

$$\begin{aligned} \mathcal{P}(\xi^T w - \mu^T w < R - \mu^T w) &\leq \frac{w^T \Sigma w}{w^T \Sigma w + (\mu^T w - R)^2} \\ 1 - \mathcal{P}(\xi^T w - \mu^T w \geq R - \mu^T w) &\leq \frac{w^T \Sigma w}{w^T \Sigma w + (\mu^T w - R)^2} \\ \mathcal{P}(\xi^T w - \mu^T w \geq R - \mu^T w) &\geq 1 - \frac{w^T \Sigma w}{w^T \Sigma w + (\mu^T w - R)^2} \end{aligned} \quad (10)$$

Therefore,

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

is sufficient for constraint (8) to hold. The expression above can be successively rewritten as:

$$\begin{aligned} (1-p) (w^T \Sigma w + (\mu^T w - R)^2) &\geq w^T \Sigma w \\ (1-p) (\mu^T w - R)^2 &\geq p w^T \Sigma w \\ \mu^T w - \sqrt{\frac{p}{1-p}} \sqrt{w^T \Sigma w} &\geq R \end{aligned}$$

which was set out to prove. □

A tighter approximation can be obtained if the probability distribution is symmetric.

Theorem 2.6 *The second-order cone constraint*

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

is a valid approximation of the probabilistic constraint

$$\mathcal{P} (\xi^T w \geq R) \geq p$$

when the portfolio return has a symmetric probability distribution.

Proof. Chebychev's inequality for symmetric probability distributions is formulated as follows:

$$\mathcal{P} (\xi^T w - \mu^T w > \mu^T w - R) \leq \begin{cases} 0.5 \cdot \min \left[1, \frac{w^T \Sigma w}{(\mu^T w - R)^2} \right] & \text{if } \mu^T w \geq R \\ 1 & \text{otherwise} \end{cases}, \quad (11)$$

where the expression $\min[a, b]$ returns the minimum value of a and b

Consequently, we have that

$$1 - \frac{1}{2} \frac{w^T \Sigma w}{(\mu^T w - R)^2} \leq P (\xi^T w - \mu^T w \leq \mu^T w - R),$$

and, using the same variable substitution approach as above, we obtain

$$P (\xi^T w - \mu^T w \geq R - \mu^T w) \geq 1 - \frac{1}{2} \frac{w^T \Sigma w}{(\mu^T w - R)^2}. \quad (12)$$

Therefore,

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

is a sufficient condition for $\mathcal{P} (\xi^T w \geq R) \geq p$ to hold true. Consequently,

$$\begin{aligned} 2 (\mu^T w - R)^2 &\geq \frac{w^T \Sigma w}{1-p} \\ \mu^T w - \sqrt{\frac{1}{2(1-p)}} \sqrt{w^T \Sigma w} &\geq R \end{aligned}$$

which was set out to prove. □

2.3 Integrality constraints for stock market restrictions

We now propose extensions of problem (7) in order to take into account real-life stock market restrictions. These are modeled through the introduction of integer decision variables in (7), and pertain to prevention from holding small positions (Section 2.3.1), to the requirement of purchasing shares by batch of a certain size (Section 2.3.2), and to the investment in a predefined minimal number of industrial sectors (Section 2.3.3).

This leads to the formulation of integer convex probabilistic problems whose deterministic equivalents are second-order cone mixed-integer problems whose general formulation is given below:

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w + F^{-1}(1-p)\sqrt{w^T \Sigma w} \geq R, \\
& \quad g_j(w, y) \leq 0, \quad j = 1, \dots, m \\
& \quad w \in \mathcal{R}_+^{r+1}, \\
& \quad y \in \mathcal{Z}_+.
\end{aligned} \tag{13}$$

Problem (13) minimizes the volatility of the portfolio over a convex feasible set determined by the second-order cone constraint on the expected return and m deterministic constraints $g_j(w, y) \leq 0$. The decision variables y are integer-valued.

2.3.1 Buy-in threshold constraints

In this section, we introduce constraints that prevent investors from holding very small active positions. The rationale for this hinges on the fact that such small positions have very limited impact on the total performance of the portfolio [29], but trigger some tracking and monitoring costs. Certain portfolio selection models, such as the Markowitz model, are known for occasionally returning an optimal portfolio containing very small investments in a (large) number of securities. Such a portfolio is in practice very difficult to justify due to the costs of establishing and maintaining it (brokerage fees, bid-ask spreads, etc.), and the usually poor liquidity of small positions. In order to avoid this, constraints preventing from holding an active position representing strictly less than a prescribed proportion w_{min} of the available capital are useful. To model such constraints, we introduce r extra binary variables $\delta_j \in \{0, 1\}$, $j = 1, \dots, r$ taking value 1 if the investor detains shares of asset j (i.e., $w_j > 0$):

$$w_j \leq \delta_j, \quad j = 1, \dots, r. \tag{14}$$

Small investments are avoided by introducing the following constraints:

$$w_{min} \delta_j \leq w_j, \quad j = 1, \dots, r. \tag{15}$$

With these additional variables and constraints, problem (13) becomes

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w + F^{-1}(1-p)\sqrt{w^T \Sigma w} \geq R \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& w_j \leq \delta_j, \quad j = 1, \dots, r \\
& w_{\min} \delta_j \leq w_j, \quad j = 1, \dots, r \\
& \delta \in \{0, 1\}^r \\
& w \in \mathcal{R}_+^{r+1}
\end{aligned} \tag{16}$$

2.3.2 Round lot purchasing constraints

Large institutional investors usually purchase large (i.e., even lot) blocks of individual financial assets. This is primarily because such blocks are more easily traded than smaller (i.e., odd lot) holdings, but also for liquidity reasons, i.e., to avoid the risk of getting stuck with a small, poorly liquid holding of a financial asset. Another reason to buy stocks by lots of large quantity is that, often, brokers require a premium for odd lot trades because they may have to split an even lot which would leave them with the remaining odd lot part. This is what motivates the construction of portfolio optimization models including *round lot* constraints that require the purchase of shares by *batches* or *lots* of M stocks.

To each risky asset j , we associate a general integer variable γ_j , and a round lot constraint

$$x_j = \gamma_j M, \quad j = 1, \dots, r \tag{17}$$

imposing that the number x_j of shares of asset j in the portfolio is a multiple of M . Denoting by p_j the face value of stock j and by K the available capital, it follows that $x_j = \frac{w_j K}{p_j}$. We can therefore reformulate (17) as

$$w_j = \frac{p_j \gamma_j M}{K}, \quad j = 1, \dots, r.$$

Problem (13) becomes a second-order cone problem with general integer decision variables

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w + F^{-1}(1-p)\sqrt{w^T \Sigma w} \geq R \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& w_j = \frac{p_j \gamma_j M}{K}, \quad j = 1, \dots, r \\
& \gamma \in \mathcal{Z}_+^r \\
& w \in \mathcal{R}_+^{r+1}
\end{aligned} \tag{18}$$

2.3.3 Diversification constraints

Many institutional investors have limitations on the allowable exposure to risky investments. Very often, such limits are defined by an upper bound on the maximum percentage of the portfolio value that may be invested in certain categories of financial assets, and/or by the requirement to invest in a predefined minimum number of asset categories or industrial sectors. In this section, we consider constraints that force the investor to diversify its portfolio by purchasing assets in at least L_{min} different economic sectors. Every asset j is linked with an economy sector k , so that the sets $S_k, k = 1, \dots, L$ of assets affiliated with a sector k form an exact partition of $\{1, \dots, r\}$. We associate a binary variable $\zeta_k \in \{0, 1\}$ with each economic sector k : ζ_k is equal to 1 if and only if the investment in sector k ($\sum_{j \in S_k} w_j$) is above a minimum pre-defined level s_{min} :

$$s_{min} \zeta_k \leq \sum_{j \in S_k} w_j \leq s_{min} + (1 - s_{min}) \zeta_k .$$

In addition to the constraint above we must add a *cardinality* constraint

$$\sum_{k=1}^L \zeta_k \geq L_{min}$$

to satisfy the diversification requirement.

The diversification condition requires to detain "representative" positions in at least L_{min} sectors. Note that the constraints above do not consider a very small position in a sector k (i.e., $\leq s_{min}$) as contributing to the diversification of the portfolio. The probabilistic Markowitz model with diversification constraint reads:

$$\begin{aligned} & \min w^T \Sigma w \\ & \text{subject to } \mu^T w + F^{-1}(1-p)\sqrt{w^T \Sigma w} \geq R \\ & w_0 + \sum_{j=1}^r w_j = 1 \\ & s_{min} \zeta_k \leq \sum_{i \in S_k} w_i \leq s_{min} + (1 - s_{min}) \zeta_k, k = 1, \dots, L \\ & \sum_{k=1}^L \zeta_k \geq L_{min} \\ & \zeta \in \{0, 1\}^L \\ & w \in \mathcal{Z}_+^{r+1} \end{aligned} \tag{19}$$

3 Solution Method

In this paper, we develop an *exact* Mixed-Integer Non-Linear Programming (MINLP) solution method for portfolio optimization problems subject to the joint enforcement of probabilistic constraint on the expected portfolio return and integer constraints representative of trading mechanisms. More precisely,

we rely on a non-linear branch-and-bound algorithm that we complement with new branching rules, namely the *idiosyncratic risk* and *portfolio risk* branching rules.

The proposed solution approach is implemented within the open-source solver Bonmin [6, 2] (available under the Common Public License) designed to solve to optimality general MINLP of form

$$\begin{aligned} & \min f(x) \\ & \text{subject to } g_i(x) \leq 0, \forall i = 1, \dots, m \\ & \quad x_i \in \mathcal{Z}, \forall i \in I \\ & \quad x \in \mathbb{R}^n, \end{aligned}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g_i : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are at least once continuously differentiable convex functions. Several algorithm can be chosen in Bonmin for solving MINLP and the reader is referred to [6] for a detailed description of the algorithms and their implementation. In our case, preliminary tests indicated that the branch-and-bound was the best method available in Bonmin for solving portfolio optimization problems. In the following, we give a brief reminder of the classical branch-and-bound algorithm used in Bonmin and then describe two new branching rules which are suitable for the considered portfolio optimization problems.

3.1 Non-linear branch-and-bound algorithm

The non-linear branch-and-bound algorithm solves problems of the form (20) by performing an implicit enumeration through a tree search. The algorithm starts by solving the continuous relaxation of (20), where all integrality requirements have been removed, using the interior point solver Ipopt [32]. We denote by (w^*, y^*) the optimal solution of this continuous relaxation. If y^* is integer valued, then (w^*, y^*) is the optimal solution and the problem is solved. Otherwise, at least one of the integer variables (y_i) has a non integer value. Such a variable is chosen for *branching*: two sub-problems (or nodes) are created where the upper and lower bounds on y_i are set to $\lfloor y_i^* \rfloor$ and $\lceil y_i^* \rceil$, respectively, and the two sub-problems are put in a list of open nodes.

Then, at each subsequent iteration of the algorithm, a sub-problem is chosen from the list of open nodes, and the continuous relaxation of the current node is solved providing a lower bound. The enumeration at the current node can be stopped, or stated differently, the node is said to be fathomed or pruned, if any of the three following conditions happen:

- the continuous relaxation is infeasible (pruning by infeasibility);
- the optimal solution of the continuous relaxation is not better than the value of the best integer feasible solution found so far (pruning by bounds);
- the optimal solution of the continuous relaxation is integer feasible (pruning by optimality).

If the optimal solution of the continuous relaxation solution (w^*, y^*) cannot be pruned, then at least one of the integer variables (y_i) has a non integer value ($y_i^* \notin \mathcal{Z}$) in the optimal solution. One of the

integer infeasible variables y_i is then chosen for *branching*, and two new sub-problems are thus added to the list of open nodes. By iterating the process a search tree is created and the algorithm continues until the list of open sub-problems is empty.

One of the key ingredients of the branch-and-bound procedure is the choice of the variable to branch on. The most classical rule is to choose the variable which has the largest fractional part, but this rule is often not very efficient. In this paper, we present two new rules specifically adapted to the portfolio optimization problems presented in Section 2. These two rules are respectively called *idiosyncratic risk* and *portfolio risk* branching and are described in the next sections.

3.2 Static branching rule: Idiosyncratic Risk Branching

The idiosyncratic risk branching rule is a static branching rule in which branching priorities are determined *a priori* (i.e., before the optimization is started).

For each integer decision variable, the branching priority is given by an integer π_i . At each node, the variable chosen for branching is the one, among the integer constrained variables with fractional value in the optimal solution of the current continuous relaxation, which has the highest priority. In case of a tie (i.e., when several candidate variables for branching have the same priority), the variable selected for branching is the one among those with highest priority which is the most fractional in the continuous relaxation.

It is important to recall that, in the optimization problems with buy-in threshold constraints (16) and with round lot constraints (18), there is a mapping between assets and integer decision variables: to each asset j corresponds a unique integer decision variable δ_j in (18) and γ_j in (16). In the context of mean-variance portfolio optimization problems, we propose to give the highest priority to the integer decision variable associated with the asset whose return has the greatest variance. We refer thereafter to this branching procedure as the *idiosyncratic risk branching* procedure. The intuition behind these priorities is that the asset with the largest variance is the one which has the most significant impact on the overall risk of the portfolio. Therefore, if the variance is the largest, the two sub-problems resulting from the branching are more likely to have an optimal value differing substantially from that of the parent node.

For the problems with diversification constraints, each integer decision variable is associated with a specific industrial sector. To each binary variable, and therefore to each sector, we assign a branching priority which is an increasing function of the sum of the variances of each asset stock related to the considered sector.

3.3 Dynamic branching rule: Portfolio Risk Branching

The portfolio risk branching rule is a dynamic branching rule, in which the branching priorities change at each node and tributary of the structure of the portfolio at the current node. Clearly, the branching variable is determined by relying upon a *dynamic, integrated* risk approach. The dynamics of the branching rule stems from the revision of the branching priorities at each node in the search tree, while its integrated

risk approach derives from the fact that the branching priorities are a function of the specific contribution of each variable (asset) to the overall risk of the portfolio.

The dynamic feature is relevant since, in the course of the optimization process, a new optimal portfolio can potentially be constructed at each node in the branch-and-bound tree. Therefore, an iterative (at each node) evaluation of the contribution of each variable to the variance of the portfolio is desirable. As it will be detailed in the next subsections, it is possible to establish a direct correspondence between an integer decision variable and an asset. At each node in the branch-and-bound tree, we consider each integer variable whose optimal value in the current continuous relaxation is not integer feasible. For each such variable, we evaluate how the restoration of the integrality condition impacts (increases) the variance of the current portfolio. The variable whose integer feasibility restoration has the largest impact on the variance receives the highest priority, and is the one with respect to which we branch.

To carry out this evaluation, we approximate the problem at hand by a more simple disjunctive program with quadratic objective function and linear equality constraints which takes into account the integrality of only one variable:

$$\begin{aligned}
& \min f(w) = w^T \Sigma w \\
& \text{subject to } Aw = b, \\
& (w_i \leq \underline{w}_i) \vee (w_i \geq \bar{w}_i), \quad i \in 1, \dots, r \\
& w \in \mathcal{R}^r
\end{aligned} \tag{20}$$

Clearly, the problem above, and therefore the evaluation of the impact of the integer feasibility restoration, are obtained by omitting the non-linear term in the portfolio return constraint and relaxing the bounds on the variables.

In the next subsections, we give a precise description of how this approximation is obtained for each variant of the probabilistic Markowitz problem. Prior to this, we explain how the branching rule is applied in the general setting of (20).

Let w^* be the (continuous) optimal solution of (20), and let $\mathcal{L}_\lambda(w)$ be the Lagrangian function:

$$\mathcal{L}_\lambda(w) = f(w) + \lambda^T (Aw - b). \tag{21}$$

We estimate the change in the objective value of (20) through the Lagrangian function. A movement of $\delta \in \mathcal{R}^r$ from w^* induces the following change in (21):

$$\begin{aligned}
\mathcal{L}_\lambda(w^* + \delta) - \mathcal{L}_\lambda(w^*) &= (w^* + \delta)^T \Sigma (w^* + \delta) - w^{*T} \Sigma w^* + \lambda^T (A\delta) \\
&= \delta^T \Sigma \delta + (2w^{*T} \Sigma + \lambda^T A) \delta.
\end{aligned}$$

Since w^* is optimal, it satisfies the KKT conditions:

$$\begin{aligned}
2w^{*T} \Sigma + \lambda^T A &= 0 \\
\lambda(Aw^* - b) &= 0
\end{aligned}$$

which implies that $\mathcal{L}_\lambda(w^* + \delta) - \mathcal{L}_\lambda(w^*) = \delta^T \Sigma \delta$.

Let us consider a variable w_i with value w_i^* , such that $w_i^* \in [\underline{w}_i, \bar{w}_i]$. Branching on w_i creates two nodes: in each of them we add one of the constraints $w_i \leq \underline{w}_i$ and $w_i \geq \bar{w}_i$. Using the procedure described above, we estimate the change in the Lagrangian of (20) by computing the two estimates δ_i^- and δ_i^+ defined by

$$\begin{aligned}\delta_i^- &= (w_i^* - \underline{w}_i) e_i^T \Sigma (w_i^* - \underline{w}_i) e_i = (w_i^* - \underline{w}_i)^2 \sigma_{ii} \\ \delta_i^+ &= (\bar{w}_i - w_i^*) e_i^T \Sigma (\bar{w}_i - w_i^*) e_i = (\bar{w}_i - w_i^*)^2 \sigma_{ii}\end{aligned}\tag{22}$$

where e_i is a vector whose components are all equal to 0 but the i -th one which is equal to 1.

By analogy to mixed-integer programming [21], we then combine these two estimates to obtain the *score* of variable w_i by taking a linear combination of the minimum and the maximum of the two [21]:

$$\delta_i = L \min(\delta_i^-, \delta_i^+) + U \max(\delta_i^-, \delta_i^+).\tag{23}$$

We set the values of L to 1 and U to 2.

We calculate δ_i for all integer variables with fractional values in the optimal solution of the continuous relaxation, and we select as branching variable the one which has the highest score:

$$\hat{i} = \arg \max_{\{i: w_i^* \in (\underline{w}_i, \bar{w}_i)\}} \delta_i.$$

The quality of the branching scheme depends on the quality of the relaxation (20) with respect to the original problem. For the problems handled in this paper, it is easy to build such relaxations, and the computational experiments indicate that they are of good quality.

3.3.1 Problem with buy-in threshold constraints

In this section, we discuss the implementation of the dynamic portfolio risk branching rule in problem (16) in which the constraints (14) and (15) define the minimum proportion of available wealth K that must be invested in any active position.

In this case, we use the following formulation:

$$\begin{aligned}\min \quad & w^T \Sigma w \\ \text{subject to} \quad & \mu^T w + F^{-1}(1-p) \sqrt{w^T \Sigma w} \geq R \\ & w_0 + \sum_{j=1}^r w_j = 1 \\ & (w_i \leq 0) \vee (w_i \geq w_{\min}) \\ & w \in \mathcal{R}_+^{r+1}\end{aligned}\tag{24}$$

Note that this formulation is strictly equivalent to (16): imposing the condition $w_i \leq 0$ is equivalent to setting γ_i to 0 in (16) and imposing $w_i \geq w_{\min}$ is equivalent to setting γ_i to 1. The continuous relaxation is obtained by removing the disjunctive constraints.

The selection of the branching variable is performed by applying the scheme described in Section 3.3 to the following relaxation

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w = R \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& (w_i \leq 0) \vee (w_i \geq w_{\min}) \\
& w \in \mathcal{R}^{r+1}
\end{aligned}$$

of problem (24). The relaxation is obtained by transforming the portfolio return constraint into an equality constraint from which the non-linear component is dropped, and by removing the non-negativity constraints.

3.3.2 Problem with round lot constraints

The constraint $\gamma_i = \frac{K}{Mp_i} w_i$ establishes a direct correspondence between the continuous variables w_j and the integer ones γ_j in portfolio optimization problems with round lot constraints (18). Therefore, for a particular value of w^* , we use the following relaxation

$$\begin{aligned}
& \min w^T \Sigma w \\
& \text{subject to } \mu^T w = R \\
& w_0 + \sum_{j=1}^r w_j = 1 \\
& (w_i \leq \left\lfloor \frac{K}{Mp_i} w_i^* \right\rfloor) \vee (w_i \leq \left\lceil \frac{K}{Mp_i} w_i^* \right\rceil)
\end{aligned} \tag{25}$$

to select the branching variable in portfolio optimization problems with round lot constraints (18).

4 Computational results

4.1 Test problems

To build the test bed for our approach, we use the daily return data of more than 600 stocks that have been part of Standard&Poor's 500 index between 1990 and 2004. The data accounts for the splits that the considered stocks have undergone in the period indicated above. Based on the time series, we calculate the estimates of the geometric mean of the returns and their variance-covariance matrix.

Using those data, we build 36 portfolio optimization instances of various sizes (12 problems with 50 assets, 12 with 100 and 12 with 200) by randomly selecting the assets included in those problems. For each problem instance, we formulate three models corresponding to the trading constraints (buy-in threshold, round lot purchase, and diversification) considered in this paper. To model the problems with

diversification constraints, we use the Global Industry Classification Standard (GICS) [30] developed by S&P and Morgan Stanley Capital International to identify the industrial sector to which each company belongs. The GICS structure consists of 10 Sectors, 24 Industry Groups, 67 Industries and 147 Sub-Industries. The present study allocates each company to one of the 67 industries. The data come from the CRSP database and were obtained using the Wharton Research Database Service.

In each problem instance, the prescribed return level R is set equal to 7%, the fixed return of the money market is equal to 2% and the prescribed reliability level p , by which the investor wants the expected portfolio return to exceed the prescribed return level, is set to 85%. The asset returns are assumed to follow a normal distribution. The problem instances are modeled by using the Ampl modeling language.

In our experiments, we compare MINLP_BB[20] and the default branch-and-bound algorithm of Bonmin[6] to our specialized branch-and-bound algorithms implemented within the Bonmin framework. MINLP_BB uses a branch-and-bound method and solves the continuous relaxations with a sequential quadratic trust region algorithm called filterSQP[13]. Some of main differences between MINLP_BB and Bonmin are:

- MINLP_BB uses an active set method for solving the continuous relaxation while Bonmin uses an interior point algorithm,
- MINLP_PP uses the depth-first search strategy for choosing the next node to process in the tree search (i.e., it selects the deepest node for processing next) while Bonmin, by default, uses best-bound (i.e., the next node to be processed is chosen as the one whose parent provides the smallest lower bound).

All tests were performed on an IBM IntellistationZ Pro with an Intel Xeon 3.2GHz CPU, 2 gigabytes of RAM and running Linux Fedora Core 3.

4.2 Evaluation of solution approaches

4.2.1 Model with buy-in threshold constraints

In this section, we analyze the computational results obtained for the problem instances containing buy-in threshold constraints. The experiments have been conducted by setting the minimum fraction of wealth (w_{\min}) to be invested in an asset (should the investor decide to include that asset in his portfolio) equal to 2%, 3% and 5% for the instances with 50, 100 and 200 stocks, respectively.

Table 5 reports the results obtained with the four solution approaches listed below on the 36 problem instances with buy-in threshold constraints:

- Bonmin’s branch-and-bound algorithm with branching performed on the most fractional integer variable (i.e., the default branching rule in Bonmin),
- Bonmin’s branch-and-bound algorithm with the idiosyncratic risk branching rule (Section 3.2),

- Bonmin’s branch-and-bound algorithm with the portfolio risk branching rule (Section 3.3.1),
- MINLP_BB’s branch-and-bound algorithm.

The above solution approaches will thereafter be referred to as *MF*, *IR*, *PR*, and *MBB*, respectively.

For each ”combination” of problem instance and solution approach, Table 5 reports

- the quality of the best obtained solution (columns 2, 5, 8, 11). We use the acronym ”NS” to indicate that no feasible integer was found. We report the value of the mixed-integer optimality gap when the best integer solution found is not optimal, and use the symbol ”*” when the optimal solution is found;
- the computing time (in CPU seconds) needed to find the optimal solution (columns 3, 6, 9, 12). If this latter cannot be found within the allowed computing time (3 hours), the entry in the table reads ”>10800”;
- the number of explored nodes in the branch-and-bound tree (columns 4, 7, 10, 13);

First, we comment on the accuracy of the solutions found. It is well known that the structure of the variance-covariance matrix of returns often leads to numerical difficulties [5]. While we can not exactly establishing the optimality of the obtained solutions, outside of the tolerances of the solvers, we can compare the values of the optimal solutions obtained with Bonmin and with MINLP_BB; we recall that both solvers are based on very different continuous nonlinear programming methods. We observe that the relative difference between the optimal solutions found by Bonmin and MINLP_BB are in the order of 10^{-4} except for problem 050_1 where it is $8.56 * 10^{-3}$ (note that the solution found by the three variants of Bonmin are always identical for these instances as well as for all the other instances in the paper). It is also worth pointing out that the solution claimed by Bonmin has always a better objective value than the one claimed by MINLP_BB on these problems.

The instances with 50 and 100 do not really allow us to discriminate the four solution approaches in terms of the quality of the solution. Indeed, Figure 2 shows that the optimal solution is found by each solution approach for each 50-stock and 100-stock problem instance.

For the most complex problems containing 200 stocks, the solution approaches *IR* and *PR* utilizing the two new proposed branching rules clearly dominate *MP* and *MBB*. The former two approaches solve each instance to optimality, while the latter two solve only 25% of those instances to optimality. It is also worth noting that *MP* does not find any integer feasible solution when it cannot find the optimal one, while *MBB* always finds an integer feasible solution, and has an average optimality gap of 5.29%.

Figures 3 and 4 display the average computing time for each combination of solution approach and size of problem instance.

In Figure 4, the left-sided graph shows the average time computed over 200-stock instances, while the right-sided one shows the average time computed over the only instances that could be solved to optimality by every solution approach. It is clear that the solution approaches *IR* and *PR* relying respectively

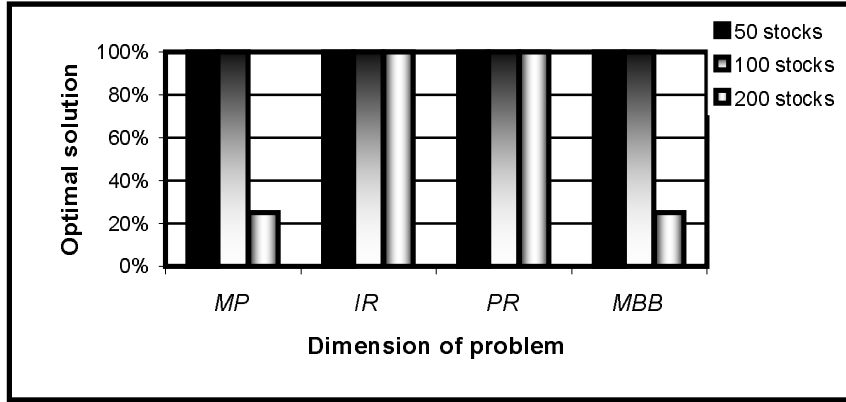


Figure 2: Quality of solution for problems with buy-in constraints

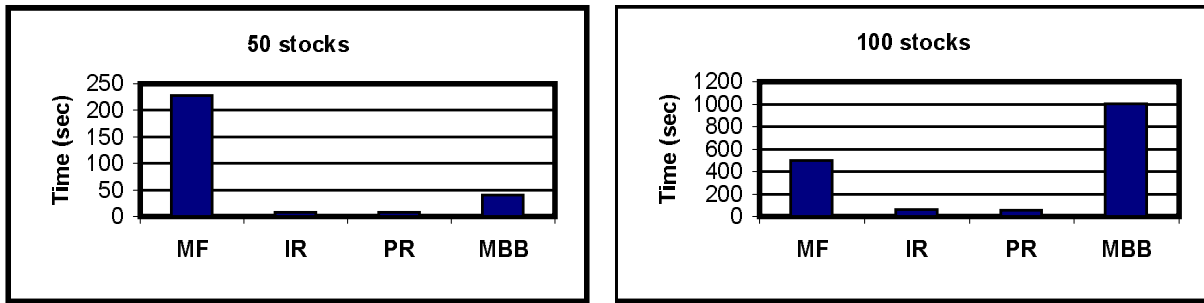


Figure 3: Average computing time for 50-stock and 100-stock instances with buy-in constraints

on the idiosyncratic and portfolio risk branching rules are, regardless of the size of the problem, much faster than *MP* and *MBB*. The *PR* solution approach is slightly faster than *IR*, and is on average more than 5 (respectively, 17 and 25) times faster than *MBB* on 50-stock (respectively, 100- and 200-stock) instances.

Figure 5 shows the evolution of the average computing time (for all instances on the left-hand side, for instances solved to optimality on the right-hand side). We can see that *PR* and *IR* scale very well: the rhythm at which their average computing time increases is very reasonable, therefore indicating their applicability to problems of larger size. This must be contrasted to the *MF* and *MBB* approaches for which the computing time seems to increase exponentially in the number of assets.

4.2.2 Model with round lot constraints

Table 5 reports the computational results for the 36 problem instances with round lot constraints and in which the investor is constrained to buy shares by multiples of M set equal to 100 in our experiments. Table 5 provides the same outputs (optimality gap, CPU time, number of nodes) and uses the same notations as those of Table 5. The following four integer solution methods have been tested:

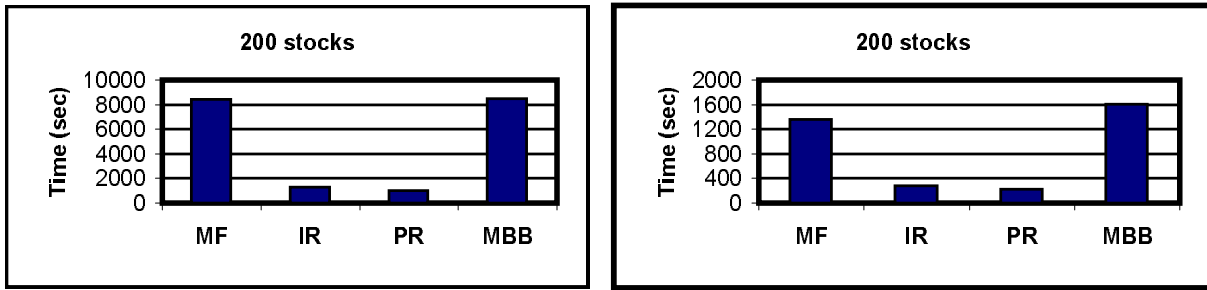


Figure 4: Average computing times for 200-stock instances with buy-in constraints

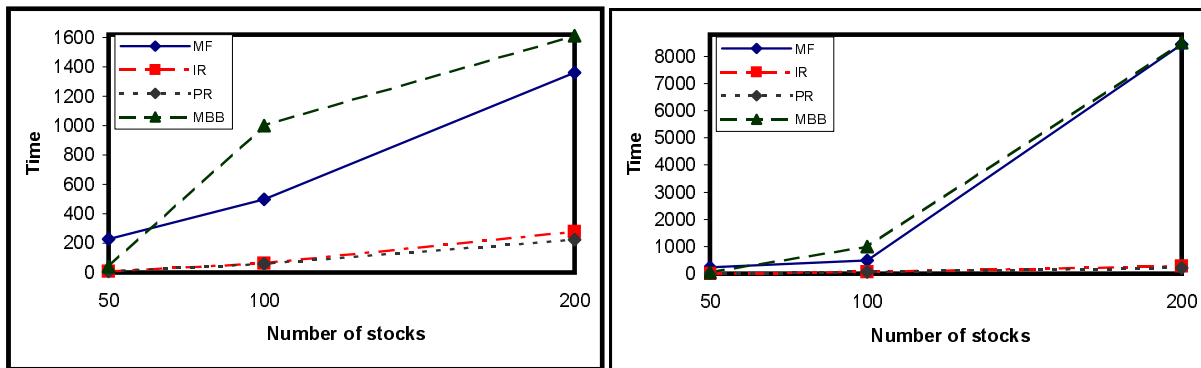


Figure 5: Buy-in constraints: computing time as a function of dimensionality

- Bonmin’s branch-and-bound algorithm with branching on the most fractional integer variable,
- Bonmin’s branch-and-bound algorithm with the idiosyncratic risk branching rule (Section 3.2),
- Bonmin’s branch-and-bound algorithm with the portfolio risk branching rule (Section 3.3.2),
- MINLP_BB’s branch-and-bound algorithm.

Figure 6 shows that the *IP* solution approach using the dynamic portfolio risk branching rule is by far the most robust method for problems with round lot constraints. The *IP* method is the only one solving to optimality *all* 100-problem instances, and finds the optimal solution for 83% of the 200-problem instances, while *none* of the three other methods can solve to optimality *any* of those twelve problem instances. A few additional comments are in order. First, the *MF* approach does not find any feasible integer solution for any of the problem instances that it cannot solve to optimality (i.e., 43% and 100% of the 100-stock and 200-stock instances, respectively). The *IR* does not find any integer feasible solution for any of the 200-problem instances. On the other hand, *MBB* always finds a feasible integer solution, and has an average optimality gap of 0.204% and 1.039% for the 100-stock and 200-stock instances, respectively.

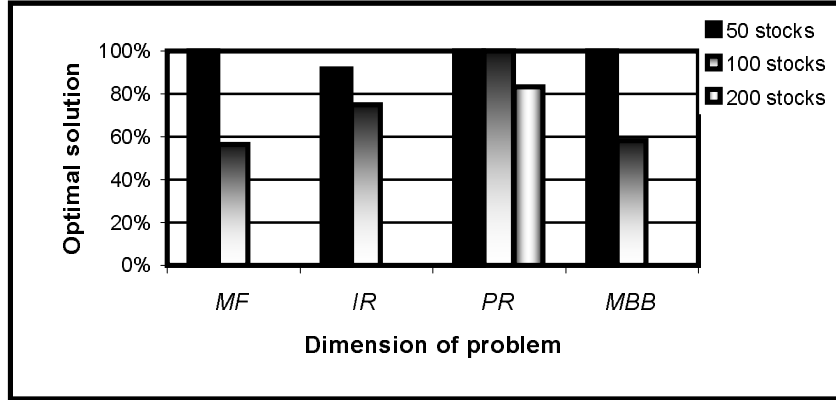


Figure 6: Quality of solution for problems with round lot constraints

Figure 7 shows that *PR* is not only the most robust but also the fastest regardless of the dimensionality of the problem. The average computing times (i.e., irrespective of whether one considers all instances [left-side in Figure 7], or only those solved to optimality by all approaches [right-side in Figure 7]) of *PR* are very significantly lower than those of the other methods. It appears that the difference in speed between *PR* and any of the other three methods increases with the size of the problem; indeed, *PR* is

- 1.22 (instances solved to optimality) and 1.44 (all instances) times faster than *MBB* (the second-fastest method) for the 50-stock instances;
- 3.65 (instances solved to optimality) and 10.62 (all instances) and times faster than *MBB* for the 100-stock instances.

No speed comparison can be drawn for the 200-stock instances since *PR* is the only method solving some (83%) of them to optimality.

Finally, we note that the relative difference between the optimal values found by Bonmin and MINLP_BB are always smaller than 10^{-4} . In all but 6 cases the value found by Bonmin is better than the one found by MINLP_BB. In those 6 cases where MINLP_BB finds a better solution the largest relative difference is $2 * 10^{-5}$.

4.2.3 Model with diversification constraints

The results displayed in Table 5 are related to the 36 problem instances with cardinality-type diversification constraints.

The results have been obtained by setting L_{min} (the minimum number of sectors in which the investor must allocate his capital) to 10, 15 and 20 for the problem instances comprising 50, 100 and 200 assets, respectively, and by setting s_{min} (minimal position in any of the K_{min} sectors) to 1% for all problem instances. The results obtained with the following three integer solution methods

- Bonmin's branch-and-bound algorithm with branching on the most fractional integer variable,

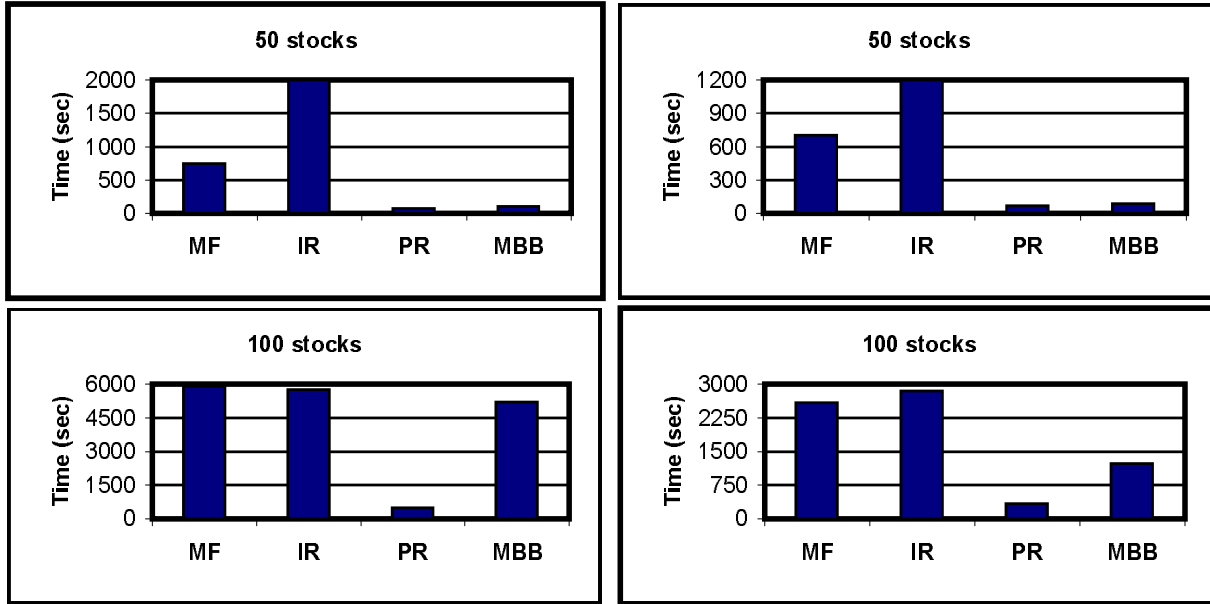


Figure 7: Average computing time for instances with round lot constraints

- Bonmin’s branch-and-bound algorithm with the idiosyncratic risk branching rule (Section 3.2),
- MINLP_BB’s branch-and-bound algorithm

are given in Table 5.

The results in Table 5 indicate that the three methods above solve to optimality all 36 instances in very limited computing time. The average computing times for the slowest and fastest methods (respectively *MF* and *MBB*) are equal to 126 sec and 69 sec. Clearly, the problems with diversification constraints appear the easiest to solve.

The relative difference between the optimal values found by Bonmin and MINLP_BB is again in the order of 10^{-4} and the solution found with Bonmin is always smaller than that obtained with MINLP_BB.

4.3 Impact of integer trading constraints

We discuss below the impact on the various types of integer trading constraints. In particular, we analyze

- the difficulty of solving the problem associated with each type of constraints. The difficulty is evaluated with respect to the average computing time per type of models and for each problem size (50, 100, 200 stocks). Figure 8 shows that the computational time is an increasing function in the number of stocks, and highlights the following hierarchy in terms of problem complexity:
 1. problem with cardinality constraints,
 2. problem with buy-in threshold constraints,

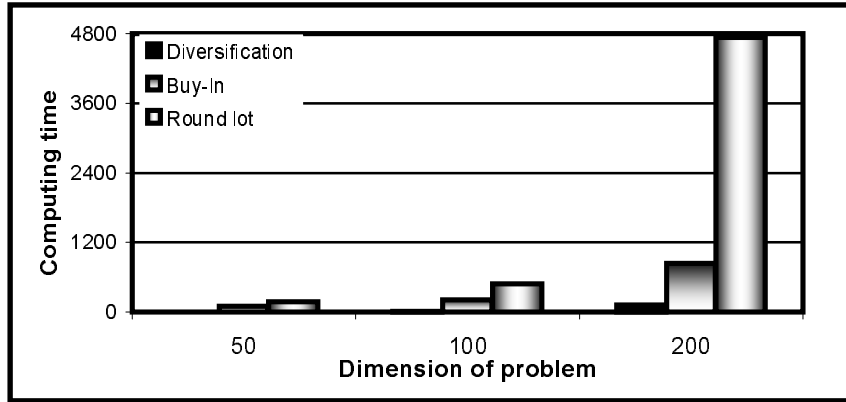


Figure 8: Average computing time per model type and problem dimension

3. problem with round lot constraints.

The largest problems (i.e., 200-stock instances) with diversification constraints require less computing time on average than the least complex (i.e., 50-stock instances) problems with round lot constraints. The accrued complexity of those latter is due to the presence of general integer variables which implicitly require the detention of an integer number of shares of any asset included in the optimal portfolio.

- the impact of the buy-in threshold constraints. Table 1 presents detailed results about the composition of the optimal portfolio for each combination of model type (without integer constraints, with diversification, round lot and buy-in threshold constraints) and problem size. The notation N^P and N^{SP} respectively denote the average number of positions in the optimal portfolio and the average number of positions which are greater than the threshold imposed by the buy-in constraints. The threshold w_{\min} is equal to 2%, 3% and 5% for the 50-, 100- and 200-stock instances, respectively. Table 1 shows that the buy-in constraints drastically change the structure of the optimal portfolio. The optimal portfolio with buy-in constraints is less diversified than the optimal portfolio obtained with any of the other three approaches. The optimal portfolio with buy-in constraints has positions in 16, 24 and 10 assets for 50-, 100- 100-, and 200-stock instances, respectively. These number must be contrasted to those of the optimal portfolios without any integer constraints (24, 30, 34), with diversification constraints (26, 37, 41), and with round lot constraints (24, 28, 30).
- the impact of the diversification constraints. In addition to constraining the holding of positions in a pre-defined number of industrial sectors, the diversification constraints, as shown by Table 1, have also for effect that the investor detains positions in a larger number of assets (at least, on average, 20.5% of the available assets) and detains a larger number of small positions (at least, on average, 56.76%).

		No integer constraint (7)	Diversification constraint (19)	Round lot constraint (18)	Buy-in threshold constraint (16)
50-stock instances ($w_{min} = 0.02\%$)	N^P	24	26	24	16
	N^{SP}	14	16	13	0
	N^{SP}/N^P	58.33%	61.54%	54.17%	0%
100-stock instances ($w_{min} = 0.05\%$)	N^P	30	37	28	24
	N^{SP}	13	21	11	0
	N^{SP}/N^P	43.33%	56.76%	39.29%	0%
200-stock instances ($w_{min} = 0.03\%$)	N^P	34	41	30	10
	N^{SP}	32	39	28	0
	N^{SP}/N^P	94.12%	95.12%	93.33%	0%

Table 1: Concentration effect of buy-in threshold constraint

- the impact of the round lot constraints. The requirement to buy shares by large lots has for effect to limit the number of active positions which is smaller than that for the model without integer constraints and with diversification constraints.

5 Conclusion

In this paper, we study the probabilistic Markowitz' mean-variance portfolio optimization model with integer-based trading constraints. We consider real-world trading constraints, such as the need to diversify the investments in a number of industrial sectors, the non-profitability of holding small positions, or the constraint of buying stocks by lots, which are modeled with integer variables. We account for the uncertainty in the estimation of the expected asset return through the introduction of a stochastic constraint ensuring that the expected return of the portfolio exceeds the prescribed return with a high confidence level. We derive stochastic integer formulations for each type of trading constraints, show under which conditions they are convex, reformulate them as quadratic integer programming problems, and develop exact solution techniques.

A key additional contribution of this paper is that it develops an *exact* solution approach for portfolio optimization problems in which uncertainty in the estimate of the expected return and real-life market restriction modeled with integer constraints are *simultaneously* considered. The joint presence of integrality restrictions and of a non-linear, probabilistic constraint explains the complexity of solving such problems, for which very few solvers can be efficiently used.

The proposed solution approaches consist of non-linear branch-and-bound algorithms. In particular, we propose two new branching techniques that we implement in a nonlinear branch-and-bound algorithm. The first one is a static branching rule, called idiosyncratic risk branching, while the second one is an integrated, dynamic branching rule, called portfolio risk branching. The latter updates, at each node in the branch-and-bound tree, the branching priorities given to the integer variables depending on their impact on the variance of the portfolio.

We evaluate the efficacy of four exact integer solution approaches on 36 problem instances containing up to 200 assets and constructed using the stocks included in the S&P 500 Index. We have not found any other computational study considering so many assets for a stochastic portfolio optimization model subject to integer constraints. Computational results show that the solution approach using the portfolio risk branching rule is the most performing one, both in terms of speed and robustness (i.e., percentage of problems solved to optimality), and that it scales well. We also derive a hierarchy of the integer trading constraints and give insights about the impact (concentration effect) of the buy-in threshold and diversification constraints.

The algorithmic results presented in this paper pave the way for multiple extensions. Our ongoing research relates to the algorithmic developments described below. The scalability of the proposed solution approaches could lead to their application to problems of larger dimension. The running time could be further reduced by relying on a second-order cone programming solver to optimize the continuous relaxations of the second-order cone problems (i.e., polynomial running time) at each node in the branch-and-bound tree. Other trading constraints (i.e., "transaction cost", "tax lot", "maximum number of transaction" constraints, etc.) leading to the formulation of other types of second-order cone problems with integer variables deserve attention. Branch-and-cut solution approaches could also be considered.

References

- [1] Avérous J., Meste M. 1997. Skewness for Multivariate Distributions: Two Approaches. *The Annals of Statistics* 25 (5), 1984-1997.
- [2] Basic Open Source Solver for Mixed Integer Nonlinear Programming: <https://projects.coin-or.org/Bonmin>
- [3] Bawa V.S., Brown S.J., Klein R.W. 1979. *Estimation Risk and Optimal Portfolio Choice*. North-Holland. Amsterdam, The Netherlands.
- [4] Ben-Tal, A., A. Nemirovski. 2002. Robust Optimization - Methodology and Applications. *Mathematical Programming* 92, 453-480.
- [5] Bienstock D. 1996. Computational Study of a Family of Mixed-Integer Quadratic Programming Problems. *Mathematical Programming* 75, 121-140. 48, 28-43.
- [6] Bonami P., Biegler L.T., Conn A.R., Cornuéjols G., Grossmann I.E., Laird C.D., Lee J., Lodi A., Margot F., Sawaya N., Wachter A. 2005. An Algorithmic Framework for Convex Mixed-integer Nonlinear Programs. *IBM Research Report*.
- [7] Ceria S., Stubbs R.A. 2006. Incorporating Estimation Errors into Portfolio Selection: Portfolio Construction. *Journal of Asset Management* 7 (2), 109-127.

- [8] Cheon M.-S., Ahmed S., Al-Khayyal F. 2006. A Branch-Reduce-Cut Algorithm for the Global Optimization of Probabilistically Constrained Linear Programs. *Mathematical Programming* 108 (2-3), 617-634.
- [9] Chopra V., Ziemba W.T. 1993. The Effects of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice. *Journal of Portfolio Management* 19, 6-11.
- [10] Cornuejols G., Tütüncü R. 2006. *Optimization Methods in Finance*. Cambridge University Press.
- [11] Costa O.L.V., Paiva A.C. 2002. Robust Portfolio Selection Using Linear-Matrix Inequality. *Journal of Economic Dynamics & Control* 26, 889-909.
- [12] El Ghaoui L., Oks M., Oustry F. 2003. Worst-Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach. *Operations Research* 51 (4), 543-556.
- [13] Fletcher R., Leyffer, S. 1998. User Manual for FilterSQP. *Numerical Analysis Report NA/181*, Dundee University.
- [14] Frangioni A., Gentile C. 2006. Perspective Cuts for a Class of Convex 0-1 Mixed Integer Programs. *Mathematical Programming* 106,225-236.
- [15] Goldfarb D., Iyengar G. 2003. Robust Portfolio Selection Problems. *Mathematics of Operations Research* 28 1-38.
- [16] Henrion R., Strugarek C. 2007. Convexity of Chance Constraints with Independent Random Variables. *Computational Optimization and Applications*. Forthcoming.
- [17] Jobst N.J., Horniman M.D., Lucas C.A., Mitra G. 2001. Computational Aspects of Alternative Portfolio Selection Models in the Presence of Discrete Asset Choice Constraints. *Quantitative Finance* 1, 1-13.
- [18] Kataoka S. 1963. A Stochastic Programming Model. 1963. *Econometrica* 31 (1-2), 181-196.
- [19] Konno H., Yamamoto R. 2005. Integer Programming Approaches in Mean-Risk Models. *Computational Management Science* 2 (4), 339-351.
- [20] Leyffer S. 1999. *User Manual for MINLP_BB*, University of Dundee.
- [21] Linderoth J.T. and Savelsbergh M.V.P. 1999. A Computational Study of Branch and Bound Search Strategies for Mixed Integer Programming. *INFORMS Journal on Computing* 11, 173-187.
- [22] Markowitz H.M. 1952. Portfolio Selection. *The Journal of Finance* 7, 77-91.
- [23] Marshall A., Olkin I. 1979. *Inequalities: Theory of Majorization and its Applications*. Academic Press, New York.

- [24] Mulvey J., Erkan H.G. 2003. Decentralized Risk Management for Global P/C Insurance Companies. *Working Paper*. In: *Applications of Stochastic Programming*. Eds: Wallace S.W., Ziemba W.T. MPS-SIAM Series on Optimization, Philadelphia, PA, 503-530.
- [25] van de Panne C., Popp W. 1963. Minimum Cost Cattle Feed under Probabilistic Constraints. *Management Science* 9, 405-430.
- [26] Prékopa A. 1974. Programming under Probabilistic Constraints with a Random Technology Matrix. *Mathematische Operationsforschung und Statistik* 5, 109-116.
- [27] Prékopa A. 1995. *Stochastic Programming*. Kluwer. Boston, MA. *Computers and Chemical Engineering* 16, 937-947.
- [28] Ruszczyński A., Shapiro A. 2003. *Stochastic Programming: Handbook in Operations Research and Management Science 10*. Elsevier.
- [29] Scherer B., Martin D. 2005. *Introduction to Modern Portfolio Optimization*. Springer, New York.
- [30] Standard&Poor's. 2005. S&P and MSCI Barra Announce Revisions to the Global Industry Classification Standard (GICS). *S&P Technical Report*.
- [31] Tütüncü R. H., Koenig M. 2004. Robust Asset Allocation. *Annals of Operations Research* 132, 157-187.
- [32] Wächter A., Biegler L.T. 2006. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming* 106 (1), 25-57.
- [33] Weintraub A., Vera J. 1991. A Cutting Plane Approach for Chance Constrained Linear Programs. *Operations Research* 39 (5), 776-785.

Name	Most Fractional			Idiosyncratic Risk			Portfolio Risk			MINLP_BB		
	Final Gap	Time	#nodes	Final Gap	Time	#nodes	Final Gap	Time	#nodes	Final Gap	Time	#nodes
050_01	*	97.4	1250	*	10.2	148	*	11.4	142	*	90.9	478
050_02	*	23.9	161	*	2.7	39	*	2.3	30	*	30.9	506
050_03	*	62.9	860	*	7.1	114	*	6.9	96	*	30.3	2140
050_04	*	31.4	433	*	5.2	87	*	6.2	86	*	30.2	1080
050_05	*	2194.7	167	*	5.0	81	*	4.3	58	*	30.8	312
050_06	*	40.1	563	*	23.7	393	*	22.1	316	*	30.5	1746
050_07	*	36.3	384	*	8.5	129	*	7.2	98	*	30.3	682
050_08	*	107.5	1438	*	6.5	109	*	6.1	86	*	90.4	3398
050_09	*	49.7	224	*	3.6	55	*	4.5	52	*	30.2	492
050_10	*	37.5	384	*	6.7	98	*	8.4	96	*	30.4	576
050_11	*	43.9	527	*	7.4	123	*	8.8	104	*	30.7	2294
050_12	*	6.9	62	*	7.2	58	*	1.7	18	*	30.6	214
100_01	*	99.0	257	*	14.3	47	*	7.1	20	*	30.2	440
100_02	*	60.1	212	*	17.9	75	*	11.0	34	*	90.7	892
100_03	*	201.5	696	*	12.7	49	*	11.7	38	*	210.5	1788
100_04	*	40.9	147	*	16.0	64	*	11.6	36	*	90.3	1116
100_05	*	2993.1	9865	*	301.3	1201	*	360.2	1234	*	7890.9	56994
100_06	*	435.1	1560	*	89.7	368	*	60.2	206	*	1050.3	8396
100_07	*	122.5	454	*	29.4	125	*	23.0	82	*	150.2	1266
100_08	*	59.8	218	*	22.4	95	*	15.1	54	*	90.1	804
100_09	*	1186.7	4321	*	147.7	575	*	101.5	362	*	1170.7	8872
100_10	*	150.6	539	*	24.5	85	*	17.4	62	*	150.4	1262
100_11	*	116.5	472	*	16.5	59	*	7.6	24	*	90.1	698
100_12	*	716.9	2487	*	108.8	412	*	97.6	348	*	870.6	6778
200_01	NS	> 10800	9342	*	488.5	388	*	413.3	358	13.56 %	> 10800	13100
200_02	NS	> 10800	8816	*	2054.1	1434	*	1768.2	1462	5.69%	> 10800	17400
200_03	NS	> 10800	8935	*	1092.8	785	*	636.6	536	1.32%	> 10800	23400
200_04	*	655.8	509	*	152.2	103	*	90.1	72	*	750.6	2230
200_05	NS	> 10800	9249	*	1960.1	1454	*	1766.9	1540	12.26%	> 10800	18100
200_06	NS	> 10800	9051	*	483.9	347	*	357.0	306	12.32%	> 10800	17700
200_07	NS	> 10800	8994	*	2080.3	1431	*	1252.5	1066	7.19%	> 10800	17600
200_08	NS	> 10800	9078	*	3240.2	2203	*	2149.5	1812	2.42%	> 10800	17900
200_09	NS	> 10800	9196	*	1493.9	1105	*	1801.8	1544	4.58%	> 10800	17800
200_10	*	1891.7	1603	*	342.1	253	*	269.9	232	*	2610.9	7314
200_11	*	1539.3	1321	*	336.2	253	*	307.8	260	*	1470.5	4668
200_12	NS	> 10800	8867	*	1681.6	1186	*	1026.7	832	4.14%	> 10800	18000

Table 2: Computational results for problems with buy-in thresholds constraints

Name	Most Fractional			Idiosyncratic Risk			Portfolio Risk			MINLP_BB		
	Final Gap	Time	#nodes	Final Gap	Time	#nodes	Final Gap	Time	#nodes	Final Gap	Time	#nodes
0050_01	*	331.67	6666	*	560.29	9080	*	71.94	1376	*	84.31	6992
050_02	*	1202.61	19108	0.003%	> 10800	35444	*	50.34	864	*	239.91	20016
050_03	*	1080.59	19540	*	624.19	9788	*	178.89	3160	*	254.70	20514
050_04	*	49.22	1122	*	78.09	1483	*	32.75	644	*	18.65	2064
050_05	*	199.33	3600	*	98.87	1596	*	37.14	672	*	43.10	3080
050_06	*	294.23	5984	*	4775.53	57750	*	96.34	1642	*	87.82	7244
050_07	*	265.08	5198	*	102.56	1796	*	50.67	950	*	70.30	6054
050_08	*	963.17	18510	*	58.84	1094	*	58.21	1110	*	247.86	15938
050_09	*	34.44	662	*	296.58	4508	*	20.14	384	*	10.13	900
050_10	*	123.82	2390	*	123.64	1992	*	38.33	756	*	35.71	3324
050_11	*	185.57	3804	*	118.72	1984	*	38.07	750	*	76.33	8054
050_12	*	4181.16	1076	*	6350.26	914	*	150.37	338	*	15.29	1208
100_01	*	645.55	2600	*	3972.24	12103	*	177.68	642	*	376.64	6300
100_02	NS	> 10800	46008	*	3542.65	11434	*	257.42	1014	1.33%	> 10800	104500
100_03	*	1740.26	7675	*	1156.38	3551	*	166.67	629	*	861.52	9862
100_04	*	7267.37	28870	*	1715.61	5910	*	485.79	1786	*	4169.63	61678
100_05	NS	> 10800	47385	NS	> 10800	34067	*	1492.14	5552	0.21%	> 10800	100000
100_06	NS	> 10800	46974	*	10252.58	32820	*	311.95	1198	0.27%	> 10800	105100
100_07	*	1517.37	6908	*	1056.07	3578	*	305.70	1168	*	508.37	5964
100_08	*	2348.11	10248	*	4903.24	21438	*	328.52	1320	*	575.11	5980
100_09	NS	> 10800	49687	*	5868.05	22300	*	994.20	4132	0.0042	> 10800	120100
100_10	*	1298.51	5644	0.02%	> 10800	45007	*	286.53	1068	*	847.22	16360
100_11	*	1977.42	8422	*	4309.18	18370	*	562.46	2134	*	906.56	10832
100_12	NS	> 10800	49169	0.01%	> 10800	46515	*	489.76	1806	0.21%	> 10800	104900
200_01	NS	> 10800	9804	NS	> 10800	10338	*	2322.59	1758	0.98%	> 10800	29200
200_02	NS	> 10800	8562	NS	> 10800	10296	NS	> 10800	7419	0.95%	> 10800	19000
200_03	NS	> 10800	9272	NS	> 10800	9149	*	2608.64	1862	0.42%	> 10800	19100
200_04	NS	> 10800	8705	NS	> 10800	8704	*	5595.12	3900	0.27%	> 10800	24200
200_05	NS	> 10800	9617	NS	> 10800	7773	*	6869.09	5658	2.79%	> 10800	24300
200_06	NS	> 10800	8229	NS	> 10800	8173	*	5437.88	4430	0.85%	> 10800	24300
200_07	NS	> 10800	9340	NS	> 10800	10625	*	5076.81	4436	0.92%	> 10800	23900
200_08	NS	> 10800	10757	NS	> 10800	10143	*	6662.85	6174	1.10%	> 10800	19100
200_09	NS	> 10800	10168	NS	> 10800	9771	NS	> 10800	9122	2.37%	> 10800	24400
200_10	NS	> 10800	8457	NS	> 10800	8698	*	2853.25	2250	0.30%	> 10800	19200
200_11	NS	> 10800	8600	NS	> 10800	9460	*	1138.78	914	1.24%	> 10800	24200
200_12	NS	> 10800	8593	NS	> 10800	8303	*	8741.88	7136	0.27%	> 10800	24400

Table 3: Computational results for problems with round lots constraints

Name	Most Fractional			Idiosyncratic Risk			MINLP		
	Final Gap	Time	#nodes	Final Gap	Time	#nodes	Final Gap	Time	#nodes
50_01	*	18.37	322	*	2.30	33	*	4.184	176
50_02	*	0.76	11	*	2.34	23	*	0.244	12
50_03	*	1.85	34	*	2.34	38	*	1.096	82
50_04	*	1.28	27	*	1.62	27	*	0.632	70
50_05	*	1.58	26	*	1.88	27	*	0.272	18
50_06	*	2.25	44	*	2.34	37	*	1.436	110
50_07	*	3.84	66	*	3.02	37	*	1.516	114
50_08	*	10.55	216	*	4.24	67	*	3.552	246
50_09	*	1.90	34	*	15.57	33	*	1.816	144
50_10	*	8.72	90	*	2.36	31	*	1.428	106
50_11	*	1.97	36	*	3.02	47	*	1.008	64
50_12	*	1.79	27	*	2.66	33	*	0.176	10
100_01	*	4.10	12	*	10.76	24	*	2.536	22
100_02	*	9.45	38	*	12.90	37	*	3.42	40
100_03	*	11.27	39	*	13.34	39	*	4.056	54
100_04	*	1.76	5	*	1.63	3	*	0.54	4
100_05	*	21.98	79	*	24.30	67	*	16.749	162
100_06	*	17.21	64	*	23.73	67	*	16.489	182
100_07	*	2.86	11	*	6.13	17	*	0.808	6
100_08	*	19.78	77	*	13.96	41	*	4.164	40
100_09	*	7.61	27	*	24.53	65	*	3.532	28
100_10	*	11.70	43	*	13.74	37	*	4.576	42
100_11	*	7.94	37	*	12.44	39	*	5.78	82
100_12	*	20.15	81	*	11.84	33	*	14.657	130
200_01	*	92.41	89	*	77.63	49	*	43.095	74
200_02	*	178.43	163	*	167.92	101	*	123.1	226
200_03	*	83.82	75	*	128.27	81	*	29.23	58
200_04	*	93.50	70	*	76.43	39	*	51.823	84
200_05	*	170.03	149	*	134.40	83	*	84.557	156
200_06	*	64.90	61	*	69.39	43	*	36.658	62
200_07	*	46.18	41	*	95.21	59	*	23.905	52
200_08	*	248.74	236	*	140.52	89	*	137.233	328
200_09	*	293.05	251	*	265.64	159	*	163.938	274
200_10	*	19.39	15	*	39.97	23	*	7.388	12
200_11	*	190.53	175	*	171.14	107	*	113.751	232
200_12	*	24.06	21	*	53.02	29	*	8.329	16

Table 4: Computational results for problems with diversification constraints