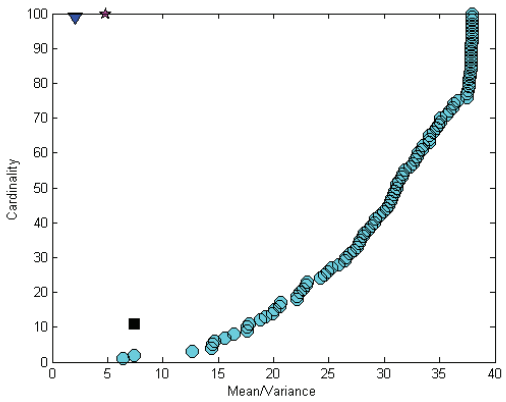

FIGURE 19. Efficient frontier of the biobjective cardinality/mean-variance problem for FF100. See the caption of Figure 4 for an explanation of the various symbols.

Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 12–05

EFFICIENT CARDINALITY/MEAN-VARIANCE PORTFOLIOS

R. P. BRITO AND L. N. VICENTE

ABSTRACT: A number of variants of the classical Markowitz mean-variance optimization model for portfolio selection have been investigated to render it more realistic. Recently, it has been studied the imposition of a cardinality constraint, setting an upper bound on the number of active positions taken in the portfolio, in an attempt to improve its performance and reduce transactions costs. However, one can regard cardinality as an objective function itself, thus adding another goal to those two traditionally considered (the variance and the mean of the return).

In this paper, we suggest a new approach to directly compute sparse portfolios by reformulating the cardinality constrained Markowitz mean-variance optimization model as a biobjective optimization problem, allowing the investor to analyze the efficient tradeoff between mean-variance and cardinality, in a general scenario where short-selling is allowed.

Since cardinality is a nonsmooth objective function, one has chosen a derivative-free algorithm (based on direct multisearch) for the solution of the biobjective optimization problem. For the several data sets obtained from the FTSE 100 index and the Fama/French benchmark collection, direct multisearch was capable of quickly determining (in-sample) the efficient frontier for the biobjective cardinality/mean-variance problem.

Our results showed that a number of efficient cardinality/mean-variance portfolios (with values of cardinality not high) overcome the naive strategy in terms of out-of-sample performance measured by the Sharpe ratio, which is known to be extremely difficult.

KEYWORDS: Portfolio selection, cardinality, sparse portfolios, multiobjective optimization, efficient frontier, derivative-free optimization.
AMS SUBJECT CLASSIFICATION (2010): 90B50, 90C56.

1. Introduction

Asset allocation is a core decision made by investors. Determining a strategy which effectively maximizes the expected utility for each investor is a challenging and arduous task. A first thought on asset allocation could lead us to consider as an ideal strategy one that maximizes the portfolio profit, measured for instance by the portfolio expected return. However, such a strategy is known to be incomplete, and therefore a poor one.

Date: February 19, 2012.
Support for this research was provided by FCT under the grant PTDC/MAT/098214/2008.

In fact, one knows since the pioneer work of Markowitz [20] that a rational investor has typically two goals in mind: to maximize the portfolio return (given, e.g., by the portfolio expected return) and to minimize the portfolio risk (described, e.g., by the portfolio variance). Traditionally, the Markowitz mean-variance optimization model is taken as a quadratic program (QP), intended to minimize the portfolio risk (variance) for a given level of expected return, over a set of feasible portfolios. By varying the level of expected return, the Markowitz model determines the so-called efficient frontier, as the set of nondominated portfolios regarding the two goals (variance and mean of the return). The rational investor can thus makes choices, by analyzing the tradeoff between expected return and variability of the investment, over a set of appropriate portfolios.

Several modifications to the classical Markowitz model or alternative methodologies have since then been proposed. One resulting from a simple observation was suggested in an article by DeMiguel, Garlappi, and Uppal [15]. These authors analyzed a number of methodologies inspired on the classic model of Markowitz and showed that none were able to significantly and consistently overcome the naive strategy, that is to say, the one in which the available investor's wealth is divided equally among the available securities. One possible explanation is related to the ill conditioning of the objective function of the Markowitz model (given by the variance of the return).

One of the important issues to consider in portfolio selection is how to handle transaction costs. There are well known modifications that can be made in the Markowitz model to incorporate transaction costs, such as to bound the turnover, which basically amount to further linear constraints in the QP. A recent technique to keep transaction costs low consists of selecting sparse portfolios, i.e., portfolios with few active positions, by imposing a cardinality constraint. Such a constraint, however, changes the classical QP into a MIQP (mixed-integer quadratic programming), which can no longer be solved in polynomial time.

In this paper, we suggest an alternative approach to the cardinality constrained Markowitz mean-variance optimization model, reformulating it directly as a biobjective problem, allowing the investor to analyze the tradeoff between cardinality and mean-variance, in a general scenario where short-selling is permitted. Such an approach allows us to find the set of nondominated points of biobjective problems in which an objective is smooth and

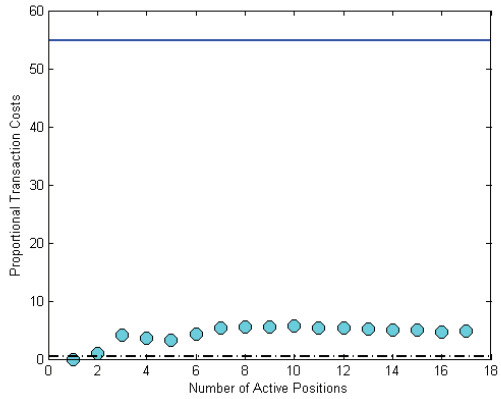

FIGURE 17. Transaction costs of the efficient cardinality/mean-variance portfolios for FF17. See the caption of Figure 16 for an explanation of the various symbols and lines.

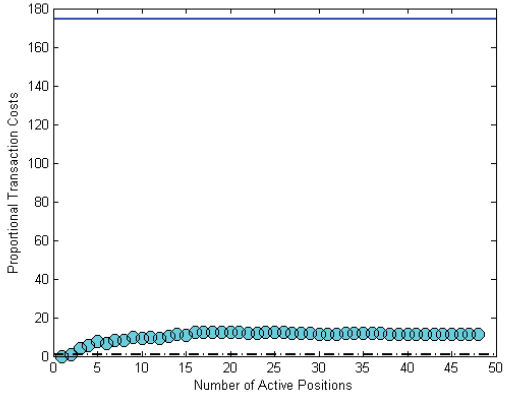

FIGURE 18. Transaction costs of the efficient cardinality/mean-variance portfolios for FF48. See the caption of Figure 16 for an explanation of the various symbols and lines.

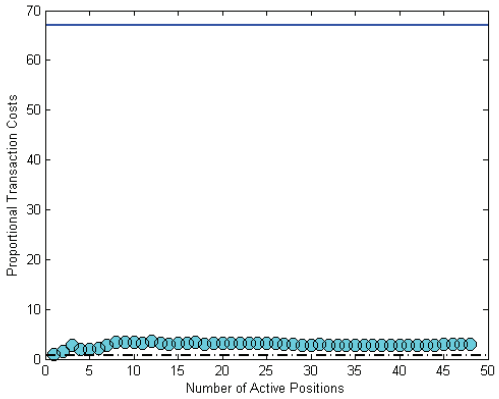

FIGURE 15. Transaction costs of the efficient cardinality/mean-variance portfolios for DTS3. See the caption of Figure 13 for an explanation of the various symbols and lines.

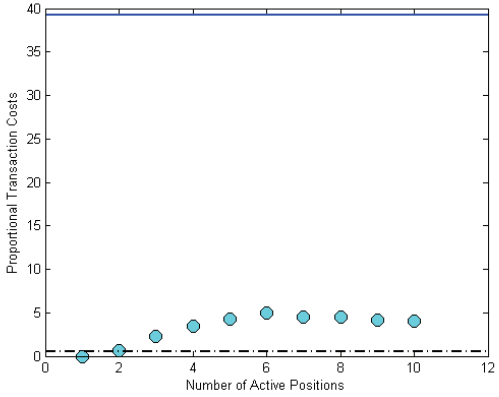

FIGURE 16. Transaction costs of the efficient cardinality/mean-variance portfolios for FF10.
— Markowitz mean-variance -.- Markowitz minimum variance

combines mean and variance and the other is nonsmooth (the cardinality or ℓ_0 norm of the vector of portfolio positions). The mean-variance objective function can take a number of forms. A parameter free possibility is given by profit per unity of risk (a nonlinear function obtained by dividing the expected return by its variance). Another possibility is a linear combination of the expected return and the variance of the return (a quadratic function).

Given the lack of derivatives of the cardinality function, we decided then to apply a directional derivative-free algorithm for the solution of the biobjective optimization problem. Such methods do not require derivatives, although their convergence results typically assume some weak form of smoothness such as Lipschitz continuity. Direct multisearch is a derivative-free multiobjective methodology for which one can show some type of convergence in the discontinuous case. More importantly, it exhibited excellent numerical performance on a comparison to a number of other multiobjective optimization solvers. We applied direct multisearch to determine (in-sample) the set of efficient or nondominated cardinality/mean-variance portfolios.

To illustrate our approach, we gathered several data sets from the FTSE 100 index (for returns of single securities) and from the Fama/French benchmark collection (for returns of portfolios), computed the efficient cardinality/mean-variance portfolios using (in-sample) optimization, and measured their out-of-sample performance using a rolling-sample approach. We found that a large number of sparse portfolios for the FTSE data sets, among the efficient cardinality/mean-variance ones, consistently overcome the naive strategy in terms of out-of-sample performance measured by the Sharpe ratio. This effect is also clearly visible for the FF data sets, where the performance of a large portion of the cardinality/mean-variance efficient frontier outperforms, in most of the instances, the naive strategy. The transactions costs are shown to be relatively low for all efficient cardinality/mean-variance portfolios, with a moderate increase with cardinality.

The organization of our paper is as follows. In the next section, we formulate the classical Markowitz model for portfolio selection, describe the naive strategy, and formulate the problem with cardinality constraint. In Section 3, we reformulate the cardinality constrained Markowitz mean-variance optimization model as a biobjective problem for application of multiobjective optimization. In Section 4, we present the empirical results. Some basic

material about derivative-free optimization and direct multisearch for multiobjective optimization is described in Appendix A. Finally, in Section 5 we summarize our findings and discuss future research.

In the remaining of this section we present some concepts and notation from multiobjective optimization used in our paper. In a multiobjective optimization problem one optimizes 'simultaneously' multiple objective functions (sometimes 'conflicting'). A constrained nonlinear multiobjective problem can be written in the form:

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & F(x) \equiv (f_1(x), \ldots, f_m(x))^\top \\ \text{subject to} & x \in \Omega \subset \mathbb{R}^n, \end{array} \tag{1}$$

involving m objective functions or objective function components $f_i : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}, i = 1, \ldots, m$, and a feasible region Ω.

In the presence of several objective functions, the minimizers of one function are not necessarily the minimizers of another. To define some sort of optimality, it is crucial to have a way of comparing different points, such as the concept of Pareto dominance. Given two points x, y in Ω, we say that $x \prec y$ (x dominates y) if and only if

$$F(x) \prec_F F(y) \iff F(y) - F(x) \in \mathbb{R}^m_+ \setminus \{0\},$$

where $\mathbb{R}^m_+$ is the nonnegative orthant $\mathbb{R}^m_+ = \{z \in \mathbb{R}^m : z \geq 0\}$. Note that the use of the nonnegative orthant induces a strict partial order.

A set of points in Ω is nondominated when no point in the set is dominated by another one in the set. The concept of Pareto dominance is thus used to characterize optimality in multiobjective optimization, by defining the set of Pareto optimizers or nondominated points. More formally, a point $x_* \in \Omega$ is said to be a (global) Pareto optimizer or a nondominated point of F in Ω if there is no $y \in \Omega$ such that $y \prec x_*$. The Pareto front or efficient frontier is the mapping by F of such set of Pareto optimizers.

2. Portfolio selection models

2.1. The classical Markowitz mean-variance model. Portfolios consist of securities (shares or bonds, for example, or classes or indices of the same). Suppose the investor has a certain wealth to invest in a set of N securities. The return of each security i is described by a random variable R_i, whose average can be computed (from estimation based on historical data).

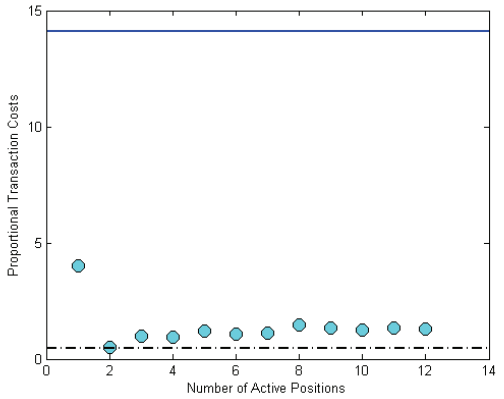

FIGURE 13. Transaction costs of the efficient cardinality/mean-variance portfolios for DTS1.
— Markowitz mean-variance -.- Markowitz minimum variance

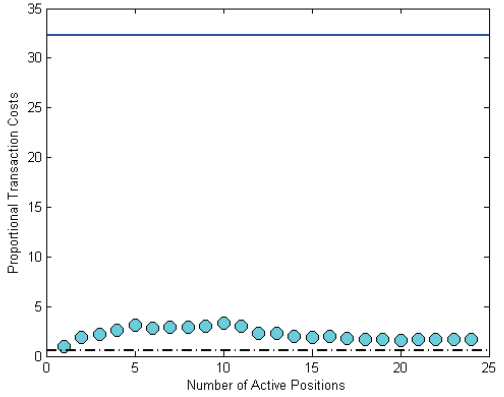

FIGURE 14. Transaction costs of the efficient cardinality/mean-variance portfolios for DTS2. See the caption of Figure 13 for an explanation of the various symbols and lines.

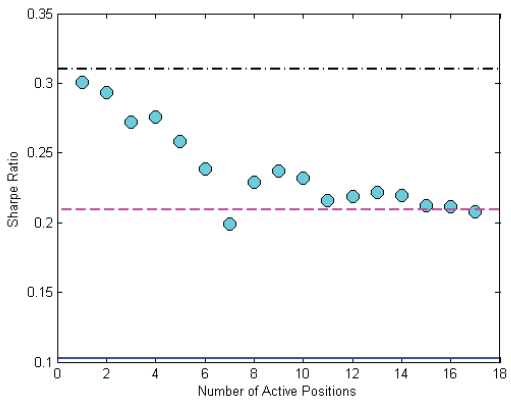

FIGURE 11. Out-of-sample performance for FF17. See the caption of Figure 10 for an explanation of the various symbols and lines.

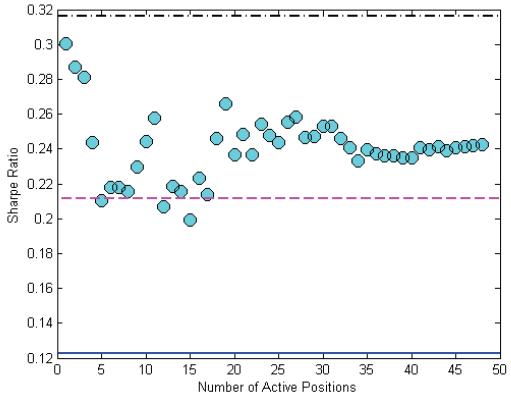

FIGURE 12. Out-of-sample performance for FF48. See the caption of Figure 10 for an explanation of the various symbols and lines.

Let $\mu_i = E(R_i)$, $i = 1, \ldots, N$, denote the expected returns of the securities. Let also w_i, $i = 1, \ldots, N$, represent the proportions of the total investment to allocate in the individual securities. The portfolio return is assumed linear in $w_1, \ldots, w_N$, and thus the portfolio expected return can be written as

$$E(R) \;=\; E(w_i R_1 + \cdots + w_N R_N) \;=\; w_1\mu_1 + \cdots + w_N\mu_N \;=\; \mu^\top w$$

with

$$\mu \;=\; (\mu_1, \ldots, \mu_N)^\top \quad \text{and} \quad w \;=\; (w_1, \ldots, w_N)^\top.$$

The portfolio variance, in turn, is calculated by

$$V(R) \;=\; E\Big(\Big[\sum_{i=1}^{N} w_i R_i - E\Big(\sum_{i=1}^{N} w_i R_i\Big)\Big]^2\Big).$$

So,

$$V(R) \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} E[(R_i - \mu_i)(R_j - \mu_j)]w_i w_j.$$

Representing each entry i, j of the covariance matrix Q by

$$\sigma_{ij} = E[(R_i - \mu_i)(R_j - \mu_j)],$$

one has

$$V(R) \;=\; w^\top Q w,$$

where Q is symmetric and positive semi-definite (and typically assumed positive definite).

As said before, a portfolio is defined by an $N \times 1$ vector w of weights representing the proportion of the total funds invested in the N securities. This vector of weights is thus required to satisfy the constraint

$$\sum_{i=1}^{N} w_i \;=\; e^\top w \;=\; 1,$$

where e is the $N \times 1$ vector of entries equal to 1. Lower bounds on the variables, of the form $w_i \geq 0$, $i = 1, \ldots, n$, can be also considered if short selling is undesirable. In general, we will say that $L_i \leq w_i \leq U_i$, $i = 1, \ldots, N$, for given lower L_i and upper U_i bounds on the variables.

Markowitz's model [20, 21] is based on the formulation of a mean-variance optimization problem. By solving this problem, we identify a portfolio of minimum variance among all which provide an expected return not below a

certain target value r. The aim is thus to minimize the risk from a given level of return. The formulation of this problem can be described as:

$$\begin{array}{ll} \min\limits_{w\in\mathbb{R}^N} & w^\top Q w \\ \text{subject to} & \mu^\top w \geq r, \\ & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1, \ldots, N. \end{array} \tag{2}$$

Problem (2) is a convex quadratic programming problem (QP), for which the first order necessary conditions are also sufficient for (global) optimality. See [11, 22] for a survey of portfolio optimization.

The classical Markowitz mean-variance model can be seen as way of solving the biobjective problem which consists of simultaneously minimizing the portfolio risk (variance) and maximizing the portfolio profit (expected return)

$$\begin{array}{ll} \min\limits_{w\in\mathbb{R}^N} & w^\top Q w \\ \max\limits_{w\in\mathbb{R}^N} & \mu^\top w \\ \text{subject to} & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1, \ldots, N. \end{array} \tag{3}$$

In fact, it is easy to prove that a solution of (2) is nondominated, efficient or Pareto optimal for (3). Efficient portfolios are thus the ones which have the minimum variance among all that provide at least a certain expected return, or, alternatively, those that have the maximal expected return among all up to a certain variance. The efficient frontier (or Pareto front) is typically represented as a 2-dimensional curve, where the axes correspond to the expected return and the standard deviation of the return of an efficient portfolio.

Another way to determine the efficient frontier is to scalarize the biobjective problem (3), introducing a weighting parameter $\lambda \in [0, 1]$ and considering

$$\begin{array}{ll} \min\limits_{w\in\mathbb{R}^N} & \lambda w^\top Q w - (1-\lambda)\mu^\top w \\ \text{subject to} & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1, \ldots, N. \end{array}$$

Here $\lambda = 0$ leads to the maximization the expected return ('independently' of the risk involved) and $\lambda = 1$ corresponds to the minimization of the risk ('independently' of the expected return involved). Values of $\lambda \in (0, 1)$

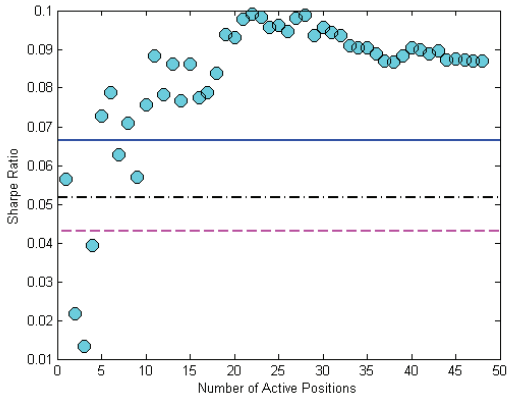

FIGURE 9. Out-of-sample performance for DTS3. See the caption of Figure 7 for an explanation of the various symbols and lines.

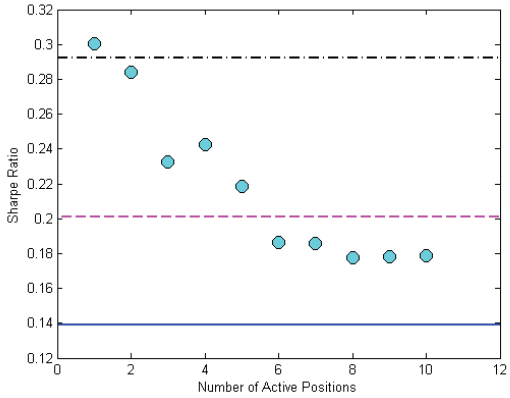

FIGURE 10. Out-of-sample performance for FF10 measured by the Sharpe ratio over all the out-of-sample periods.
- - Naive — Markowitz mean-variance -.- Markowitz minimum variance

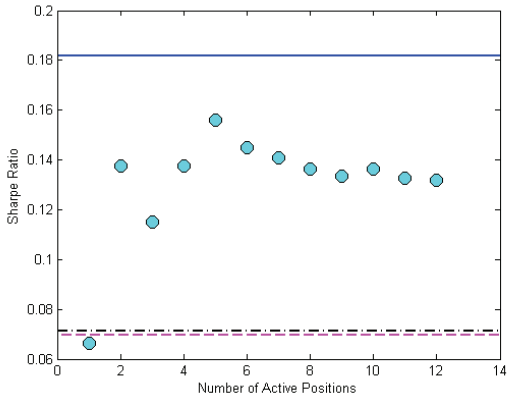

FIGURE 7. Out-of-sample performance for DTS1 measured by the Sharpe ratio over all the out-of-sample periods.
- - Naive — Markowitz mean-variance -.- Markowitz minimum variance

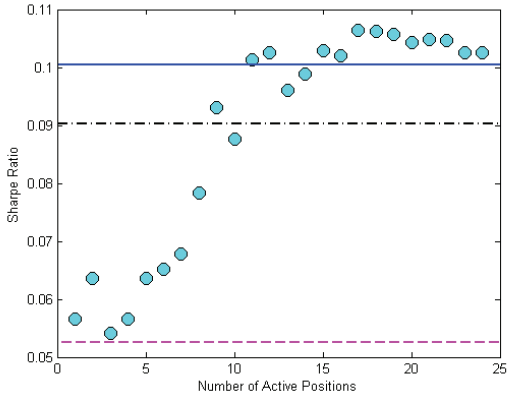

FIGURE 8. Out-of-sample performance for DTS2. See the caption of Figure 7 for an explanation of the various symbols and lines.

provide the tradeoff between risk and expected return, generating efficient portfolios between the two extremes of the efficient frontier.

2.2. The naive strategy $1/N$**.** The naive strategy is the one in which the available investor's wealth is divided equally among the securities available

$$w_i = \frac{1}{N}, \quad i = 1, \ldots, N.$$

This strategy has diversification as its main goal, it does not involve optimization, and it completely ignores the data.

Although a number of theoretical models have been developed in the last years, many investors pursuing diversification revert to the use of the naive strategy to allocate their wealth (see [2]). DeMiguel, Garlappi, and Uppal [15] evaluated fourteen models across seven empirical data sets and showed that none is consistently better than the naive strategy. A possible explanation for this phenomenon lies on the fact that the naive strategy does not involve estimation and promotes 'optimal' diversification. The naive strategy is therefore an excellent benchmarking strategy.

2.3. The cardinality constrained Markowitz mean-variance model. Since the appearance of the classical Markowitz mean-variance model, a number of methodologies have been proposed to render it more realistic. The classical Markowitz model assumes a perfect market without transaction costs or taxes, but such costs are an important issue to consider as far as the portfolio selection is concerned, especially for small investors.

Recently, it has been studied the addition of a constraint that sets an upper bound on the number of active positions taken in the portfolio, in an attempt to improve performance and reduce transactions costs. Such a cardinality constraint is defined by limiting $\text{card}(x) = |\{i \in \{1, ..., N\} : x_i \neq 0\}|$ and leads to cardinality constrained portfolio selection problems. In particular, the cardinality constrained Markowitz mean-variance optimization problem has the form:

$$\begin{array}{ll} \min\limits_{w \in \mathbb{R}^N} & w^\top Q w \\ \text{subject to} & \mu^\top w \geq r, \\ & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1, \ldots, N, \end{array} \tag{4}$$

where $K \in \{1, \ldots, N\}$. Although $\text{card}(x)$ is not a norm, it is frequently called the ℓ_0 norm in the literature, $\|x\|_0 = \text{card}(x)$.

By introducing binary variables, one can rewrite the problem as a mixed-integer quadratic programming (MIQP) problem:

$$\begin{array}{rl} \min\limits_{w,y \in \mathbb{R}^N} & w^\top Q w \\ \text{subject to} & \mu^\top w \geq r, \\ & e^\top y \leq K, \\ & e^\top w = 1, \\ & L_i y_i \leq w_i \leq U_i y_i, \quad i = 1, \ldots, N, \\ & y_i \in \{0, 1\}, \quad i = 1, \ldots, N. \end{array} \tag{5}$$

However such MIQPs are known to be hard combinatorial problems. The number of sparsity patterns in w (i.e., number of different possibilities of having K nonzeros entries) is $\binom{N}{K} = N!/[(N-K)!K!]$. Although there are exact algorithms for the solution of MIQPs (see [3, 4, 5, 25]), many researchers and portfolio managers prefer to use heuristics approaches (see [1, 7, 8, 16, 19, 26]). Some of these heuristics vary among evolutionary algorithms, tabu search, and simulated annealing (see [16, 26]).

Promotion of sparsity is also used in the field of signal and imaging processing, where a new technique called compressed sensing has been intensively studied in the recent years. Essentially one aims at recovering a desired signal or image with the least possible amount of basis components. The major developments in compressed sensing have been achieved by replacing the ℓ_0 norm by the ℓ_1 one, the latter being a convex relation of the former and known to also promote sparsity. The use of the ℓ_1 norm leads to recovering optimization problems solvable in polynomial time (in most of the cases equivalent to linear programs), and a number of sparse optimization techniques have been developed for the numerical solution of such problems. These ideas have already been used in portfolio selection primarily to promote regularization of ill conditioning (of the estimation of data or of the variance of the return itself). DeMiguel et al. [14] constrained the Markowitz classical model by imposing a bound on the ℓ_1 norm of the vector of portfolio positions, among other possibilities. Brodie et al. [6] focus on a modification to the Markowitz mean-variance classical model by the incorporation of a term involving a multiple of the ℓ_1 norm of the vector of portfolio positions.

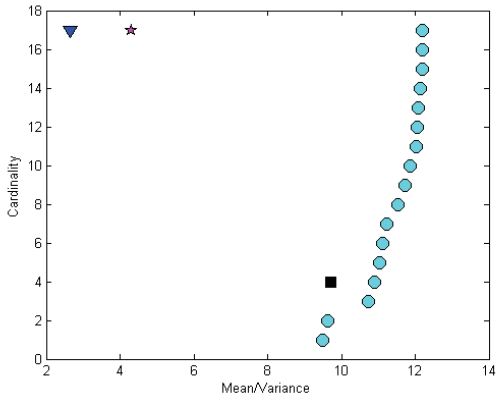

FIGURE 5. Efficient frontier of the biobjective cardinality/mean-variance problem for FF17. See the caption of Figure 4 for an explanation of the various symbols.

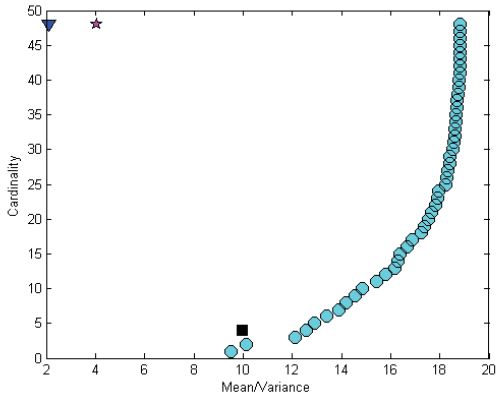

FIGURE 6. Efficient frontier of the biobjective cardinality/mean-variance problem for FF48. See the caption of Figure 4 for an explanation of the various symbols.

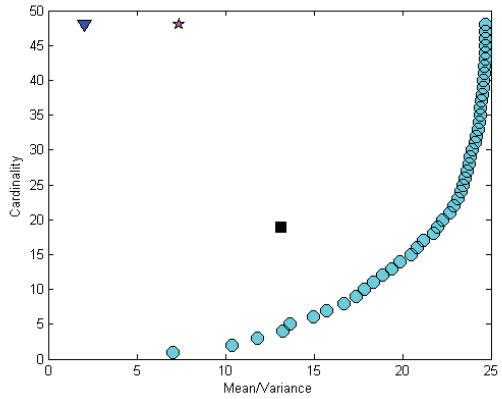

FIGURE 3. Efficient frontier of the biobjective cardinality/mean-variance problem for DTS3. See the caption of Figure 1 for an explanation of the various symbols.

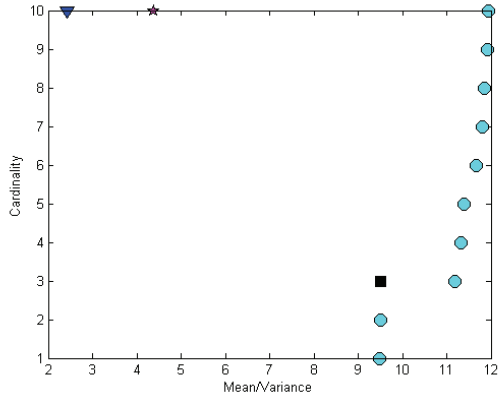

FIGURE 4. Efficient frontier of the biobjective cardinality/mean-variance problem for FF10.
★ Naive ▼ Markowitz mean-variance ■ Markowitz minimum variance

Inspired by sparse reconstruction (see, for instance, [5]), they also proposed an heuristic for the solution of the problem.

3. The cardinality/mean-variance biobjective model

Although the cardinality constrained Markowitz mean-variance model described in (4) provides an alternative to the classical Markowitz model in the sense of realistically limiting the number of active positions in a portfolio, it is dependent on the parameter K, the maximum number of such positions. Thus, one has to vary K to obtain various levels of cardinality or sparsity, and for each value of K solve an MIQP of the form (5).

The alternative suggested in this paper is to consider the cardinality function as an objective function itself. At a first glance, one could see the problem has a triobjective optimization problem by

- minimizing the variance of the return $w^\top Qw$,
- maximizing the expected return $\mu^\top w$,
- minimizing the cardinality $\text{card}(w)$,

over the set of feasible portfolios. However, one would then get a tridimensional efficient frontier, difficult to visualize. One could consider some form of projection of this frontier onto a plane but it is unclear how to proceed.

On the other hand, investors may find it useful to directly analyze the tradeoff between cardinality and mean-variance. One possibility is to linearly combine variance and mean of the return, assigning equal weights to both, reducing the problem to a biobjective, by

- minimizing the mean-variance $\frac{1}{2}w^\top Qw - \frac{1}{2}\mu^\top w$,
- minimizing the cardinality $\text{card}(w)$,

over the set of feasible portfolios. More rigorously, the cardinality/mean-variance biobjective optimization problem would then posed as

$$\begin{array}{rl} \min\limits_{w\in\mathbb{R}^N} & \frac{1}{2}w^\top Qw - \frac{1}{2}\mu^\top w \\ \min\limits_{w\in\mathbb{R}^N} & \text{card}(w) \\ \text{subject to} & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1,\ldots,N. \end{array} \tag{6}$$

A drawback of this approach is that it depends on a choice of a parameter, set above to 0.5. A parameter-free alternative is to consider a Sharpe ratio type objective function, by

- maximizing expected return per variance $\frac{\mu^\top w}{w^\top Q w}$,
- minimizing the cardinality $\text{card}(w)$,

over the set of feasible portfolios. In this case, the cardinality/mean-variance biobjective optimization problem is posed as

$$\begin{array}{ll} \min\limits_{w\in\mathbb{R}^N} & -\frac{\mu^\top w}{w^\top Q w} \\ \min\limits_{w\in\mathbb{R}^N} & \text{card}(w) \\ \text{subject to} & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1,\ldots,N. \end{array} \tag{7}$$

By solving (7) (or (6)), we identify a cardinality/mean-variance efficient frontier. A portfolio in this frontier is such that there exists no other feasible one which simultaneously presents a lower cardinality and a lower mean-variance measure. Given such an efficient frontier and a mean-variance target, an investor may directly find the answers to the questions of what is the optimal (lowest) cardinality level that can be chosen and what are the portfolios leading to such a cardinality level. The construction of sparse or lower cardinality portfolios is crucial in portfolio selection, since it makes portfolio management easier and it reduces transaction costs (the fewer securities entering in the construction of a portfolio, the lower are these costs).

Problem (7) has two objective functions and linear constraints. The first objective $f_1(w) = -\mu^\top w / w^\top Q w$ is nonlinear but smooth. However, the second objective function $f_2(w) = \text{card}(w) = |\{i \in \{1,...,N\} : w_i \neq 0\}|$ is piecewise linear discontinuous, consequently nonlinear and nonsmooth. We have thus decided to solve the biobjective optimization problem (7) using a derivative-free solver, based on direct multisearch. A brief description of direct multisearch and the solver `dms` is given in the Appendix A.

4. Empirical performance of efficient cardinality/mean-variance portfolios

Now we report a number of experiments made to numerically determine and assess the efficient cardinality/mean-variance frontier. We applied direct multisearch (see Appendix A.1) to determine the Pareto front or efficient frontier of the biobjective optimization problem (7) (according to Appendix A.2). We tested three data sets collected from the FTSE 100 index and three others from the Fama/French benchmark collection (see

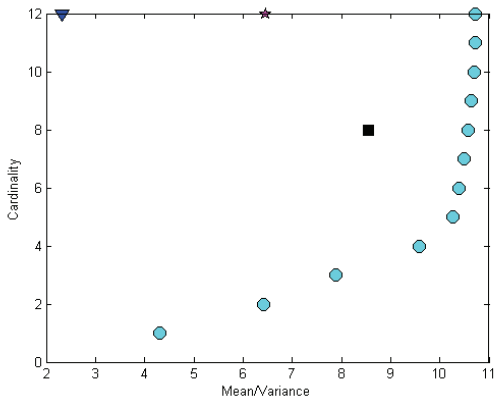

FIGURE 1. Efficient frontier of the biobjective cardinality/mean-variance problem for DTS1.
★ Naive ▼ Markowitz mean-variance ■ Markowitz minimum variance

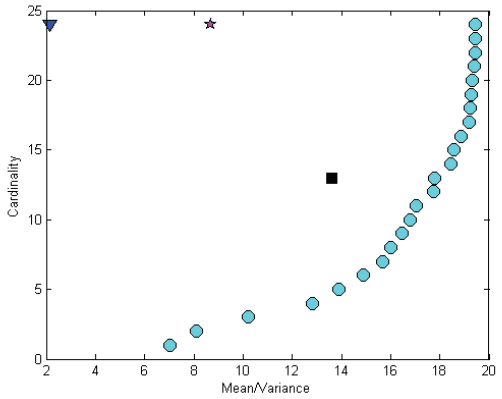

FIGURE 2. Efficient frontier of the biobjective cardinality/mean-variance problem for DTS2. See the caption of Figure 1 for an explanation of the various symbols.

[10] A. R. Conn, K. Scheinberg, and L.N. Vicente. *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization, Philadelphia, 2009.
[11] G. Cornnuejols and R. Tütüncü. *Optimizations Methods in Finance*. Cambridge University Press, Cambridge, 2007.
[12] A. L. Custdio, J. F. A. Madeira, A. I. F. Vaz, and L. N. Vicente. Direct multisearch for multiobjective optimization. *SIAM J. Optim.*, 21:1109–1140, 2011.
[13] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6:182–197, 2002.
[14] V. DeMiguel, L. Garlappi, F. J. Nogales, and R. Uppal. A generalized approach to portfolio optimization: Improving performance by constrained portfolio norms. *Management Science*, 55:798–812, 2009.
[15] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *The Review of Financial Studies*, 22:1915–1953, 2009.
[16] J. E. Fieldsend, J. Matatko, and M. Peng. Cardinality constrained portfolio optimisation. volume 3177 of *Lecture Notes in Comput. Sci.*, pages 788–793. Springer, Berlin, 2004.
[17] R. Jagannathan and T. Ma. Risk reduction in large portfolios: Why imposing the wrong constraints hfelps. *The Journal of Finance*, 58:1651–1684, 2003.
[18] T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev.*, 45:385–482, 2003.
[19] D. Lin, S. Wang, and H. Yan. A multiobjective genetic algorithm for portfolio selection. Working Paper. Istitute of Systems Science, Academy of Mathematics and Systems Science Chinese Academy of Sciences, Beijing, China, 2001.
[20] H. M. Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91, 1952.
[21] H. M. Markowitz. Portfolio selection: Efficient diversification of investments. *Cowles Foundation Monograph*, 16, 1959.
[22] M. C. Steinbach. Markowitz revisited: Mean-variance models in financial portfolio analysis. *SIAM Rev.*, 43:31–85, 2001.
[23] The MathWorks^{TM}. *MATLAB®*.
[24] L. N. Vicente and A. L. Custódio. Analysis of direct searches for discontinuous functions. *Math. Program.*, 2012, to appear.
[25] J. P. Vielma, S. Ahmed, and G. L. Nemhauser. A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs. *INFORMS J. Comput.*, 20:438–450, 2008.
[26] M. Woodside-Oriakhi, C. Lucas, and J. E. Beasley. Heuristic algorithms for the cardinality constrained efficient frontier. *European J. Oper. Res.*, 213:538–550, 2011.

R. P. Brito
Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
(rpedro.brito@gmail.com).

L. N. Vicente
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
(lnv@mat.uc.pt).

Subsection 4.1). The efficient frontiers obtained by the initial in-sample optimization are given in Subsection 4.2. The out-of-sample performance of the cardinality/mean-variance efficient portfolios, measured by a rolling-sample approach, is described in Subsection 4.3. The proportional transaction costs of each cardinality/mean-variance efficient portfolio are reported in Subsection 4.4. The section is ended with a discussion of the obtained results.

4.1. Data sets. For the first three data sets we collected daily data for securities from the FTSE 100 index, from 01/2003 to 12/2007 (five years). Such data is public and available from the site `http://www.bolsapt.com`. The three data sets are referred to as DTS1, DTS2, and DTS3, and are formed by 12, 24, and 48 securities, respectively. The composition of these data sets is given in Table 1. Based on the daily closing prices p_{it} of each session, we calculated the daily returns, using the continuous rates

$$r_{it} \;=\; \ln\left(\frac{p_{it}}{p_{it-1}}\right), \quad i = 1, \ldots, N, \quad t = 1, \ldots, T,$$

and the discrete rates

$$r_{it} \;=\; \frac{p_{it} - p_{it-1}}{p_{it-1}}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T,$$

where N ($N = 12, 24, 48$) is the number of securities and T ($T = 1303$) is the number of observations. We used the daily continuous returns for the in-sample optimization (estimation of Q and μ) and the daily discrete returns for the out-of-sample analysis.

We also included in our experiments three data sets from the Fama/French benchmark collection (FF10, FF17, and FF48, with cardinalities 10, 17, and 48), using the monthly returns from 07/1971 to 06/2011 (forty years) given there for a number of industry security sectors. More information on these security sectors (or portfolios of securities) can be found in `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

For all data sets, and given the observed returns, the estimate used for the vector μ of expected returns (see Section 2.1) was given by the arithmetic mean of the observations

$$\mu_i \;\simeq\; \frac{1}{T}\sum_{t=1}^{T} r_{it}, \quad i = 1, \ldots, N.$$

SECURITIES	
3 I GROUP (1,2,3)	JOHNSON MATTHEY P (3)
AMEC (1,2,3)	LEGAL & GENERAL (3)
ANGLO AMERICAN (1,2,3)	LLOIDS BANKING GR (3)
ANTOFAGASTA (1,2,3)	LONMIN (3)
ASSOCIAT BRIT FOO (1,2,3)	MARKS & SPENCER (3)
ASTRAZENECA (1,2,3)	MORRINSON SUPERMKT (3)
AVIVA (1,2,3)	NEXT (3)
B SKY B GROUP (1,2,3)	OLD MUTUAL (3)
BAE SYSTEMS (1,2,3)	PEARSON (3)
BARCLAYS (1,2,3)	PRUDENTIAL (3)
BG GROUP (1,2,3)	REED ELSEVIER PLC (3)
BHP BILLITON (1,2,3)	RENTOKIL INITIAL (3)
BP (2,3)	REXAM (3)
BRIT AMER TOBACCO (2,3)	RIO TINTO (3)
BRIT LAND CO REIT (2,3)	ROYAL BK SCOTL GR (3)
BRITISH AIRWAYS (2,3)	RSA INSUR GRP (3)
CAB & WIRE WRLD (2,3)	SABMILLER (3)
CAPITA GRP (2,3)	SAGE GRP (3)
COBHAM (2,3)	SAINSBURY (3)
DIAGEO (2,3)	SCHRODERS (3)
HAMMERSON REIT (2,3)	SEVERN TRENT (3)
IMPERIAL TOBACCO (2,3)	SHIRE (3)
INTERNATIONAL POW (2,3)	UNITED UTILITIES (3)
INVENSYS (2,3)	VODAFONE GRP (3)

TABLE 1. Composition of the three data sets from the FTSE 100 index. In brackets we indicate the data set to which each security belongs to.

The estimate for the entries of the covariance matrix Q (see Section 2.1) was calculated by

$$\sigma_{ij} \simeq \frac{1}{T}\sum_{t=1}^{T}(r_{it}-\mu_i)(r_{jt}-\mu_j), \quad i=1,\ldots,N, \quad j=1,\ldots,N.$$

Both μ and Q were computed using MATLAB [23] functions.

to the `dms` solver was then

$$\begin{array}{rl} \min\limits_{w(1:N-1)\in\mathbb{R}^{N-1}} & -\frac{\mu^\top w}{w^\top Q w} \\ \min\limits_{w(1:N-1)\in\mathbb{R}^{N-1}} & \sum_{i=1}^{N-1}\mathbf{1}_{\{|w_i|>\epsilon\}} \\ \text{subject to} & -b \le w_i \le b, \quad i=1,\ldots,N-1, \\ & -b \le 1-\sum_{i=1}^{N-1} w_i \le b, \end{array} \tag{10}$$

where w_N in $-\mu^\top w/w^\top Q w$ was replaced by $1-\sum_{i=1}^{N-1} w_i$.

We used all the default parameters of `dms` (version 0.2) with the following four exceptions. First, we needed to increase the maximum number of function evaluations allowed (from 20000 to 2000000 for $N(=n)$ up to 50) given the dimension of our portfolios, as well as to require more accuracy by reducing the step size tolerance from 10^{-3} to 10^{-7}. Then we turned off the use of the cache of previously evaluated points to make the runs faster (the default version of `dms` keeps such a list to avoid evaluating points too close to those already evaluated). Lastly, we realized that initializing the list of feasible nondominated points with a singleton led to better results than initializing it with a set of roughly N points as it happens by default. Thus, we set the option `list` of `dms` to zero, which, given the bounds on the variables, assigns the origin to the initial list.

References

[1] F. Bach, S. D. Ahipasaoglu, and A. d'Aspremont. Convex relaxations for subset selection. ArXiv: 1006.3601, 2010.

[2] S. Benartzi and R. H. Thaler. Naive diversification strategies in defined contribution saving plans. *The American Economic Review*, 91:79–98, 2001.

[3] D. Bertsimas and R. Shioda. Algorithm for cardinality-constrained quadratic optimization. *Comput. Optim. Appl.*, 43:1–22, 2009.

[4] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Math. Program.*, 74:121–140, 1996.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[6] J. Brodie, I. Daubechies, C. De Mol, D. Giannone, and I. Loris. Sparse and stable Markowitz portfolios. *Proc. Natl. Acad. Sci. USA*, 106:12267–12272, 2009.

[7] F. Cesarone, A. Scozzari, and F. Tardella. Efficient algorithms for mean-variance portfolio optimization with hard real-world constraints. *Giornale dell'Istituto Italiano degli Attuari*, 72:37–56, 2009.

[8] T. J. Chang, N. Meade, J. E. Beasley, and Y. M. Sharaiha. Heuristics for cardinality constrained portfolio optimisation. *Comput. Oper. Res.*, 27:1271–1302, 2000.

[9] F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, (reprinted by SIAM, Philadelphia), 1990.

(contained in the positive spanning sets D_k) must then be asymptotically dense in the unit sphere for x_* to be Pareto-Clarke stationary (see [12] for all the details). Such a result can be further generalized for discontinuous objective functions following the steps in [24].

The comprehensive numerical experience reported in [12] showed that direct multisearch performed the best among all derivative-free solvers for multiobjective optimization, even using a relatively simple implementation with an empty search step. In particular, direct multisearch clearly outperformed the very popular NSGA-II [13]. This benchmarking was done in a test set of more than 100 problems, among them some with discontinuous and nonconvex Pareto fronts. A number of tools and metrics were used to summarize the numerical findings, including data and performance profiles for the presentation of the results and purity and spread metrics to measure the quality of the obtained Pareto fronts.

A.2. Using `dms` to determine efficient cardinality/mean-variance portfolios. A few modifications to problem (7) were required to make it solvable by a multiobjective derivative-free solver, in particular by a direct multisearch one. In practice the first modification to (7) consisted of approximating the true cardinality, by introducing a tolerance ϵ,

$$\begin{array}{rl} \min\limits_{w\in\mathbb{R}^N} & -\frac{\mu^\top w}{w^\top Q w} \\ \min\limits_{w\in\mathbb{R}^N} & \sum_{i=1}^N \mathbf{1}_{\{|w_i|>\epsilon\}} \\ \text{subject to} & e^\top w = 1, \\ & L_i \leq w_i \leq U_i, \quad i = 1,\ldots,N. \end{array}$$

chosen as $\epsilon = 10^{-8}$ ($\mathbf{1}$ represents the indicator function). Secondly, we selected symmetric bounds on the variables $L_i = -b$ and $U_i = b$,

$$\begin{array}{rl} \min\limits_{w\in\mathbb{R}^N} & -\frac{\mu^\top w}{w^\top Q w} \\ \min\limits_{w\in\mathbb{R}^N} & \sum_{i=1}^N \mathbf{1}_{\{|w_i|>\epsilon\}} \\ \text{subject to} & e^\top w \;=\; 1, \\ & -b \;\leq\; w_i \;\leq\; b, \quad i = 1,\ldots,N, \end{array}$$

setting $b = 10$. Finally, we eliminated the constraint $e^\top w = 1$ since direct search methods do not cope well with equality constraints. The version fed

4.2. In-sample optimization. We then applied the solver `dms` (version 0.2) to compute the efficient frontier (or Pareto front) of the cardinality/mean-variance biobjective optimization problem (7). A few modifications to (7) were made before applying the solver as well as a few changes to the solver default parameters (the details are described in Appendix A.2).

We present results for the initial in-sample optimization. For the FTSE data sets this sample is from 03/2003 to 12/2006 and for the FF data sets is from 07/1971 to 06/1996. Figures 1–3 and Figures 4–6 contain the plots of the efficient frontiers calculated for, respectively, the FTSE and FF data sets.

In all these plots we also marked three other portfolios. The first one is the $1/N$ portfolio corresponding to the naive strategy. The other two are classical Markowitz related portfolios. One is obtained by optimizing a linear combination of mean and variance (allowing short-selling)

$$\begin{array}{rl} \min\limits_{w\in\mathbb{R}^N} & \frac{1}{2} w^\top Q w - \frac{1}{2}\mu^\top w \\ \text{subject to} & e^\top w = 1. \end{array} \tag{8}$$

The other is obtained by minimizing variance under no short-selling

$$\begin{array}{rl} \min\limits_{w\in\mathbb{R}^N} & w^\top Q w \\ \text{subject to} & e^\top w \;=\; 1, \\ & w \geq 0. \end{array} \tag{9}$$

Both instances were solved using the `quadprog` function from the MATLAB Optimization Toolbox. Regarding problem (9), it is known that not allowing short-sale has a regularizing effect on minimum-variance Markowitz portfolio selection (see [17]) and leads to portfolios of low cardinality.

4.3. Out-of-sample performance. The analysis of out-of-sample performance relies on a rolling-sample approach. For the FTSE data sets we considerer 12 periods (months) of evaluation. We begin by computing the efficient frontier (or Pareto front) of the cardinality/mean-variance biobjective optimization problem (7) for the in-sample time window from 01/2003 to 12/2006 (see Subsection 4.2). We then held fixed each portfolio and observed its returns over the next period (January 2007). Then we discarded January 2003 and brought January 2007 into the sample. We repeated this process until exhausting the 12 months of 2007. We applied the same rolling-sample approach to the FF data sets, considering an initial in-sample time window

from 07/1971 to 06/1996 (see Subsection 4.2) and 15 periods of evaluation (the 15 next years).

In each period of evaluation, the out-of-sample performance was then measured by the Sharpe ratio

$$S = \frac{m - r_f}{\sigma},$$

where m is the mean return, r_f is the return of the risk-free asset[1], and σ is the standard deviation. The results (over all the periods of evaluation) are given in Figures 7–9 for the FTSE portfolios and in Figures 10–12 for the FF ones.

4.4. Transaction costs. Since one is rebalancing portfolios for each out-of-sample period, one can compute the transaction costs of such a trade. We set the proportional transaction cost equal to 50 basis points per transaction (as usually assumed in the literature). Thus the cost of a trade over all assets is given by

$$0.5 \sum_{i=1}^{N} \mid w_{it+1} - w_{it} \mid$$

with $t = 1, \ldots, 11$ for the FTSE data sets and $t = 1, \ldots, 14$ for the FF data sets. The results are given in Figures 13–15 for the FTSE portfolios and in Figures 16–18 for the FF ones.

4.5. Discussion of the results. Contrary to one could think, given the intractability of $f_2(w) = \text{card}(w)$ and the fact that no derivatives are being used for $f_1(w) = -\mu^\top w / w^\top Q w$, direct multisearch (the solver `dms`) was capable of quickly determining (in-sample) the efficient frontier for the biobjective optimization problem (7). For instance, for the data sets of roughly 50 assets, a regular laptop takes a few dozens of seconds to produce the efficient frontiers.

For the portfolios constructed using the FTSE 100 index data (portfolios of individual securities), a large number of our sparse portfolios, among the efficient cardinality/mean-variance ones, consistently overcame the naive strategy and at least one of the two classical Markowitz models, in terms

[1]For the FTSE data sets we used the 3 month Treasury-Bills UK. Such data is public and made available by the Bank of England, at the site `http://www.bankofengland.co.uk`. For the FF data sets we used the 90-day Treasury-Bills US. Such data is public and made available by the Federal Reserve, at the site `http://www.federalreserve.gov`.

(1) **Selection of an iterate point:** Order the list L_k in some way and select the first item $(x; \alpha) \in L_k$ as the current iterate and step size parameter (thus setting $(x_k; \alpha_k) = (x; \alpha)$).
(2) **Search step:** Compute a finite set of points $\{z_s\}_{s \in S}$ and evaluate F_Ω at each element. Set $L_{add} = \{(z_s; \alpha_k), s \in S\}$.
Form L_{trial} by eliminating dominated points from $L_k \cup L_{add}$. If $L_{trial} \neq L_k$ declare the iteration (and the search step) successful, set $L_{k+1} = L_{trial}$, and skip the poll step.
(3) **Poll step:** Choose a positive spanning set D_k. Evaluate F_Ω at the set of poll points $P_k = \{x_k + \alpha_k d : d \in D_k\}$. Set $L_{add} = \{(x_k + \alpha_k d; \alpha_k), d \in D_k\}$.
Form L_{trial} by eliminating dominated points from $L_k \cup L_{add}$. If $L_{trial} \neq L_k$ declare the iteration (and the poll step) as successful and set $L_{k+1} = L_{trial}$. Otherwise, declare the iteration (and the poll step) unsuccessful and set $L_{k+1} = L_k$.
(4) **Step size parameter update:** If the iteration was successful then maintain or increase the corresponding step size parameters: $\alpha_{k,new} \in [\alpha_k, \gamma \alpha_k]$ and replace all the new points $(x_k + \alpha_k d; \alpha_k)$ in L_{k+1} by $(x_k + \alpha_k d; \alpha_{k,new})$, when success is coming from the poll step, or $(z_s; \alpha_k)$ in L_{k+1} by $(z_s; \alpha_{k,new})$, when success is coming from the search; replace also $(x_k; \alpha_k)$, if in L_{k+1}, by $(x_k; \alpha_{k,new})$. Otherwise decrease the step size parameter: $\alpha_{k,new} \in [\beta_1 \alpha_k, \beta_2 \alpha_k]$ and replace the poll pair $(x_k; \alpha_k)$ in L_{k+1} by $(x_k; \alpha_{k,new})$.

The goal of direct multisearch is to approximate the true Pareto front, although theoretically one is only able to prove that there is a limit point in a stationary form of this front, as no aggregation or scalarization technique is incorporated. In fact, it is possible to prove under standard assumptions that direct multisearch (globalized by integer lattices or sufficient decrease) generates a sequence of (unsuccessful) iterates driving the step size to zero and converging to a candidate for Pareto minimizer of the original problem, meaning to a point that satisfies some form of Pareto stationary. Essentially, one is able to prove at such a limit point x_* that, given a direction d (in the cone tangent to Ω at x_*), there exists at least one objective function component $j \in \{1, \ldots, m\}$ such that $f_j^\circ(x_*; d) \geq 0$. Here the directional derivative $f_j^\circ(x_*; d)$ is defined in the Clarke [9] way if f_j is only assumed Lipschitz continuous near x_*. The set of directions used by the algorithm

These methods maintain a list of feasible nondominated points. In both search and poll steps, the objective function components are evaluated at a finite set of points. A new trial list is then formed after adding the new points to the current list and removing all nondominated ones that might have appeared. Successful iterations correspond then to list changes, i.e., to iterations where the trial list differs from the current one.

Poll centers are chosen from the list. Note that each point in the list is associated with a step size parameter. Polling consists of evaluating the function at points obtained by adding to the poll center a multiple (defined by the step size) of a set of directions. These directions typically form a positive spanning set (in other words, a set of directions that spans $\mathbb{R}^n$ with nonnegative coefficients).

In direct multisearch, constraints are handled using an extreme barrier function

$$F_\Omega(x) = \begin{cases} F(x) & \text{if } x \in \Omega, \\ (+\infty, \ldots, +\infty)^\top & \text{otherwise,} \end{cases}$$

which prevents adding infeasible points to the current list of nondominated points.

Direct multisearch is described below following the algorithmic framework in [12]. A number of details are omitted and the reader is referred to [12] for a complete description. In particular, we omitted the two known globalization strategies (generation of points in integer lattices and imposition of sufficient decrease), under which the algorithm is known to be globally convergent, in the sense of yielding some form of convergence independently of the starting point or starting list. In practice, these globalization strategies amount to very minor modifications to the algorithm.

Algorithm A.1 (Direct Multisearch for MOO).

Initialization

Choose $x_0 \in \Omega$ with $f_i(x_0) < +\infty, \forall i \in \{1, \ldots, m\}$, $\alpha_0 > 0$, $0 < \beta_1 \leq \beta_2 < 1$, and $\gamma \geq 1$. Initialize the list of nondominated points and corresponding step size parameters ($L_0 = \{(x_0; \alpha_0)\}$ in case of a singleton).

For $k = 0, 1, 2, \ldots$

of out-of-sample performance measured by the Sharpe ratio. This effect has even happened for the largest data set (DTS3 with 48 securities), where the demand for sparsity is more relevant.

For the portfolios constructed using the Fama/French benchmark collection (where securities are portfolios rather than individual securities), the scenario is different since the behavior of the naive strategy is even more difficult to outperform. Still, a large number of sparse efficient cardinality/mean-variance portfolios consistently overcame the naive strategy. For the FF data, we also made the same set of experiments using instead 10 periods (years) for the out-of-sample evaluation. The Sharpe ratio of the naive strategy increased proportionally more than the Sharpe ratio of the efficient cardinality/mean-variance portfolios. This is due to the nature of the data and indicates a tendency of robustness of these portfolios in comparison to the naive strategy.

In both cases, FTSE and FF data, the transaction costs of the efficient cardinality/mean-variance portfolios are much lower than the classical mean-variance portfolio (solution of problem (8)) and slightly higher than the minimum-variance portfolio (solution of problem (9)). Note that the minimum-variance portfolio does not allow short-selling, and so the weights at the outset are much more limited, thus leading to better results. Finally we point out that the increase with cardinality is moderate.

We also computed the cardinality/mean-variance efficient frontier for the data set FF100, where portfolios are formed on size and book-to-market (see Figure 19). (This time we needed a budget of the order of 10^7 function evaluations, see Appendix A.2.) We point out that the data in FF100 was incomplete and filled in by interpolation.

We remark that FF48 and FF100 are the data sets also used in [6]. In this paper, as we said before, the authors focus on a modification to the Markowitz classical model by the incorporation of a term involving a multiple of the ℓ_1 norm of the vector of portfolio positions. Despite the different sparse-oriented techniques and different strategies for evaluating out-of-sample performance, in both approaches (theirs and ours), sparse portfolios are found overcoming the naive strategy. In our approach one computes sparse portfolios satisfying an efficient or nondominant property and one does it directly and in single run, whereas in [6], there is a need to vary a tunable parameter and select the portfolios according to some criterion to be

met (for example, sparsity). It is unclear what sort of efficient or nondominant property their portfolios satisfy. Moreover, we provide results for all cardinality values (from 1 to 48 in FF48 and from 1 to 100 in FF100), while in [6] the authors report results for cardinality values from 4 and 48 (FF48) and from 3 to 60 (FF100). We therefore claim to have a more direct way of dealing with sparsity, which offers a complete determination of an efficient frontier for all cardinalities.

Finally, we point out that we also made the same set of experiments using instead the cardinality/mean-variance biobjective optimization problem (6). Complete Pareto fronts were also determined. The results, in terms of out-of-sample performance, were not as good. A number of other alternatives were also tried, leading to the same conclusion (success in determining the efficient cardinality/mean-variance frontiers, but not so good out-of-sample performance).

5. Conclusions and perspectives for future work

In this paper we have developed a new methodology to deal with the computation of mean-variance Markowitz portfolios with pre-specified cardinalities. Instead of imposing a bound on the maximum cardinality or including a penalization or regularization term into the objective function (in classical Markowitz mean-variance models), we took the more direct approach of explicitly considering the cardinality as a separate goal. This led us to a cardinality/mean-variance biobjective optimization problem (7) whose solution is given in the form of an efficient frontier or Pareto front, thus allowing the investor to tradeoff among these two goals when having transaction costs and portfolio management in mind. In addition, and surprisingly, a significant portion of the efficient cardinality/mean-variance portfolios (with cardinality values considerably lower than the number N of securities) have exhibited superior out-of-sample performance (under reasonably low transaction costs that only increase moderately with cardinality).

We solved the biobjective optimization problem (7) using a derivative-free solver running direct multisearch. Direct-search methods based on polling are known in general to be slow but extremely robust due their directional properties. Such a feature is crucial given the difficulty of the problem (one discontinuous objective function, the cardinality, and discontinuous Pareto fronts). We have observed the robustness of direct multisearch, in other words, its capability of successfully solving a vast majority of the instances

(all in our case) even if at the expense of a large budget of function evaluations.

Direct multisearch was applied off-the-shelf to determine the cardinality/mean-variance efficient frontier. The structure of problem (7), or of its practical counterpart (10), was essentially ignored. One can use the fact that the first objective function is smooth and of known derivatives (or even quadratic in the case of the alternative problem (6)) to speed up the optimization and reduce even further the budget of function evaluations. Moreover, we also point out that it is trivial to run the poll step of direct multisearch in a parallel mode.

The use of derivative-free single or multiobjective optimization opens the research range of future work in sparse or dense portfolio selection. In fact, since derivative-free algorithms only rely on zero order information, they are applicable to any objective function of black-box type. One can thus use any measure to quantify the profit and risk of a portfolio. The classical Markowitz model assumes that the return of a portfolio is a linear combination of the returns of the individual securities. Also, it implicitly assumes a Gaussian distribution for the return, letting its variance be a natural measure of risk. However, it is known from the analysis of stylized facts that the distribution for the return of securities exhibits tails which are fatter than the Gaussian ones. Practitioners consider other measures of risk and profit better tailored to reality. Our approach to compute the cardinality/mean-variance efficient frontier is ready for application in such general scenarios.

Appendix A. Direct multisearch for multiobjective optimization

A.1. A description of direct multisearch. Direct multisearch is a class of methods for the solution of multiobjective optimization problems of the form (1) without the use of derivatives, which does not aggregate or scalarize any of the objective function components. It essentially generalizes all direct-search methods of directional type from single to multiobjective optimization. As in direct search [10, 18], each iteration of direct multisearch is organized around a search step and a poll step, and it is the latter the one responsible for the convergence properties. The search step is optional and when included it aims at improving numerical performance. Direct multisearch tries, however, to capture the whole Pareto front (or efficient frontier) from the polling procedure itself.