

Equity portfolio management with cardinality constraints and risk parity control using multi-objective particle swarm optimization

Massimiliano Kaucic

Department of Economics, Business, Mathematics and Statistics, University of Trieste, Piazzale Europa 1, Trieste 34127, Italy

ARTICLE INFO

Keywords:

Portfolio optimization
Cardinality constraint
Risk parity
Hybrid constraint-handling
MOPSO

ABSTRACT

The financial crisis and the market uncertainty of the last years have pointed out the shortcomings of traditional portfolio theory to adequately manage the different sources of risk of the investment process. This paper addresses the issue by developing an alternative portfolio design, that integrates risk parity into the cardinality constrained portfolio optimization model. The resulting mixed integer programming problem is handled by an improved multi-objective particle swarm optimization algorithm. Three hybrid approaches, based on a repair mechanism and different versions of the constrained-domination principle, are proposed to handle constraints. The efficiency of the algorithm and the effectiveness of the solution approaches are assessed through a set of numerical examples. Moreover, the benefits of adopting the proposed strategy instead of the cardinality constrained mean-variance approach are validated in an out-of-sample experiment.

1. Introduction

The approach developed by Markowitz in [Markowitz \(1952\)](#) models the portfolio construction process as an optimization problem involving the selection of promising assets and the allocation of capital among them. In this framework, the expected return of the portfolio is maximized while its variance, perceived as a risk for the investment, is minimized. A portfolio is efficient if it provides the maximum return for a given level of risk or, equivalently, if it has the minimum risk for a given level of desired return. Hence, the set of optimal portfolios requires the formulation and solution of a quadratic program. In the risk-return space, the set of optimal mean-variance tradeoffs forms the so-called efficient frontier, or Pareto front.

Several variants of the basic model have been investigated to render it more realistic ([Guerard Jr, 2009](#); [Kolm et al., 2014](#)). Recently, it has been suggested the inclusion of a cardinality constraint that limits the number of assets to hold in the portfolio in an attempt to control costs and improve its performance ([Liagkouras and Metaxiotis, 2015](#)). The associated mixed-integer quadratic program represents the so-called cardinality constrained portfolio optimization problem (CCPOP). It has been proven to be NP-hard ([Moral-Escudero et al., 2006](#)) and, consequently, the computation of the entire efficient frontier becomes computationally challenging. The various algorithms studied to solve this

issue can be roughly grouped in the following two categories. From one hand, there are exact methods ([Bertsimas and Shioda, 2009](#); [Li et al., 2006](#); [Shaw et al., 2008](#)), that provide optimal solutions but, as the problem size grows, they present unmanageable computational times. From the other hand, heuristic approaches ([Chang et al., 2000](#); [Crama and Schyns, 2003](#); [Cura, 2009](#); [Jobst et al., 2001](#); [Mutunge and Haugland, 2018](#); [Woodside-Oriakhi et al., 2011](#)), that guarantee the identification of approximate good, and sometimes optimal, solutions even for large size instances of CCPPOP within reasonable time.

In the last decade, researchers have tried to tackle the CCPPOP in its multi-objective form, by optimizing both objectives at the same time with the so-called multi-objective evolutionary algorithms ([Anagnostopoulos and Mamanis, 2010](#); [Anagnostopoulos and Mamanis, 2011](#); [Liagkouras and Metaxiotis, 2017](#)). The main advantage of these procedures, with respect to single objective algorithms, is that they are able to find the efficient frontier in a single run.

From an economic point of view, the traditional Markowitz's model treats both the above and the below target returns equally, while investors are more concerned about the probability of investment returns falling below the target return. Therefore, risks are under-estimated and portfolios that are downside efficient are ruled out. An alternative approach that properly models the loss aversion of investors is the replacement of variance with a downside risk measure. Nowadays, due to the regulatory importance of quantifying large losses in banking and insurance sectors, the so-called value-at-risk (VaR) occupies a leading position as risk management tool. It is defined as the maximum loss occurring over a

E-mail address: massimiliano.kaucic@deams.units.it

given period at a given confidence level. Although it is apparently simple and intuitive, VaR presents several disadvantages. Primarily, it ignores losses exceeding VaR and is not sub-additive, i.e., diversification of the portfolio may increase it (Artzner et al., 1999). In order to deal with these shortcomings, Rockafellar and Uryasev introduce the conditional value-at-risk (CVaR), which is the conditional expectation of losses above the VaR (Rockafellar et al., 2000).

The financial crisis of 2008 pointed out two other drawbacks of the Markowitz portfolio selection model: from one hand, relying on the estimates of expected returns, the optimal portfolio may be highly unstable; from the other hand, this approach tends to provide solutions excessively concentrated on few assets, thus failing in diversifying risks. A period characterized by uncertainty has followed. Consequently, investors have moved their attention toward novel portfolio strategies able to capture risk premia at lower risk levels (Roncalli, 2013). The management of different sources of risk involved in the investment process has thus become of primary relevance. Contrary to the Markowitz approach, risk-based strategies only encompass the risk dimension and neglect the performance one. In this manner, they do not require any explicit stock return forecasts and take into account solely drift and changes in risk views. Risk parity is certainly one of the most promising notions in this context (Chaves et al., 2011). It is defined in terms of risk contributions, where weights are adjusted so that all assets equally contribute to portfolio risk (Qian, 2006). Its theoretical properties have been studied in Maillard et al. (2010), while the empirical comparisons conducted in Braga (2015) indicate that risk parity portfolios have a lower risk than naive portfolios and have higher Sharpe ratio than the minimum variance portfolios. However, the existence and uniqueness of a solution to the risk parity portfolio selection problem can be guaranteed only in some special cases, for which ad hoc methods have been proposed in literature (Bai et al., 2016; Griveau-Billion et al., 2013; Maillard et al., 2010; Spinu, 2013). The only effective tool for the general case is represented by the use of a heuristic algorithm that is able to find approximate solutions (Hochreiter, 2015; Vijayalakshmi Pai and Michel, 2012).

The mutated economic conditions of the recent years have encouraged scholars and practitioners to develop investment strategies that balance performance and risk concentration by combining aggressive active portfolios, based on the Markowitz optimization model, with more conservative risk control approaches. A first attempt in this direction is represented by Roncalli (2014), where a generalized risk measure is introduced that takes into account both the portfolio return and volatility. The author suggests the use of the cyclical coordinate descent algorithm (Tseng, 2001) to solve numerically the resulting risk parity portfolio allocation problem. Feng and Palomar (2015) develop a framework gathering the most used risk parity formulations present in the literature with different performance objectives, and propose the use of a successive convex optimization method as solver. An extension of the procedure, that includes asset selection, is given in Feng and Palomar (2016). The novel formulation represents portfolio cardinality as a regularization term added to the objective function.

Unlike the aforementioned studies, that tackle the portfolio selection problem by optimizing a weighted sum of the involved criteria, we develop a multi-objective optimization framework in order to generate an even spread of points that accurately represents the complete Pareto optimal set. Our approach introduces risk parity in the definition of the admissible set of the mean-variance portfolio optimization model through mixed integer constraints. Furthermore, we use a tolerance threshold to identify solutions that are approximately at parity for the cases in which optimal portfolios do not exist.

Risk parity has been usually considered in asset-class strategies, while its application to equities is limited due to the large size of

the investable sets and to the corresponding elaborate risk models (see, for instance, de Carvalho et al., 2012; de Carvalho et al., 2014 and Siu, 2014).

From a modelling point of view, we contribute to the literature by proposing a novel equity portfolio design that aims at overcoming these difficulties with the introduction into the classical mean-variance framework of both a cardinality constraint and a risk parity condition. The first is used to limit the number of assets in the portfolio, while the second properly controls the distribution of risks among its constituents. The resulting risk parity based cardinality constrained portfolio optimization problem will be denoted by RP-CCPOP. From a computational point of view, we develop a variant of the multi-objective particle swarm optimization (MOPSO) algorithm proposed by Coello et al. (2004). In our algorithm, candidate solutions are represented by a pair of vectors. The first of them includes portfolio weights, while the second conveys possible rebalancing information, that will be used to modify the first vector according to the paradigms of particle swarm optimization. Further, a swap mutation operator has been added to increase exploration capabilities. The implementation of a novel scheme to handle at the same time cardinality, buy-in thresholds, budget and risk parity constraints is also considered. This procedure combines a repair mechanism for the first three constraints with a domination principle to compare solutions in terms of deviations from risk parity. Three alternatives are taken into account to this end, namely the constrained domination principle by Deb (2000), a self-adaptive tolerance method similar to the one developed by Zhang and Rangaiah (2012) and the self-adaptive penalty technique presented in Woldesenbet et al. (2009). Once a constraint-handling method is selected from these alternatives, the resulting hybrid constraint-handling multi-objective particle swarm optimization algorithm will be generally denoted by HMOPSO. Table 1 points out the contributions of the proposed algorithm with respect to some popular PSO variants adopted in the portfolio optimization field.

The rest of the paper is organized as follows. In Section 2, we formally describe the equity portfolio optimization problem, and discuss different implementations of the risk parity principle. Section 3 reports the description of the developed multi-objective evolutionary optimization algorithm with its three variants of the constraint-handling technique. Section 4 is devoted to the discussion and assessment of the proposed algorithms on a set of real-world experiments. The best alternative is then used to validate the novel portfolio investment strategy with respect to the standard cardinality constrained portfolio optimization model in an out-of-sample setting. Finally, the concluding remarks are presented in Section 5.

2. Problem formulation

2.1. Cardinality constrained portfolio optimization

Consider a financial market consisting of n risky assets modeled through a probability space. A portfolio is represented by a weight vector $\mathbf{x} = (x_1, \dots, x_n)^\top$, where $x_i \in \mathbb{R}$ denotes the proportion of capital invested in asset i . Assume that agents act their decisions over a one-period horizon and the random rate of return of asset i at the end of the investment period, denoted as R_i , has a continuous probability density function (pdf), then the rate of return of a portfolio \mathbf{x} is the random variable

$$R(\mathbf{x}) = \sum_{i=1}^n x_i R_i \quad (1)$$

having the pdf $f_{R(\mathbf{x})}$ induced by that of $(R_1, \dots, R_n)^\top$. $R(\mathbf{x})$ is measurable in (R_1, \dots, R_n) with expected value

Table 1

Particle swarm optimization algorithms adopted for solving portfolio optimization problems with cardinality constraints and risk parity. In the third column, SO and MO indicate single-objective and multi-objective optimization, respectively.

Algorithm	Portfolio model	Methodology	Contribution
PSO Cura (2009)	CCPOP	SO	the hybrid solution representation by Streichert et al. (2004) is used for the particles; PSO and discrete PSO are integrated; a repair mechanism handles constraints
VC-HPSO Yaakob and Watada (2010)	CCPOP	SO	particle velocity is clamped; a mutation operator is added and parameters are adaptively updated
NS-MOPSO Mishra et al. (2014)	CCPOP	MO	the hybrid solution representation by Streichert et al. (2004) is used for the particles; the non-dominated sorting is adapted to satisfy both the objectives and the constraints; velocity is clamped and its direction changes with a fixed probability
DMS-MOPSO Liang and Qu (2013)	mean-variance with no-short selling	MO	solutions are represented by real vectors; the non-dominated sorting is adapted to satisfy both the objectives and the constraints; a set of small sized swarms and a self-learning strategy are used
HMOPSO	CCPOP and RP-CCPOP	MO	solutions are represented by real vectors; the non-dominated sorting is adapted to satisfy the objectives; velocity is clamped and its direction changes with a fixed probability; a swap mutation operator is added for obtaining more efficient solutions; three hybrid mechanisms are added to handle constraints

$$\mu_p(\mathbf{x}) = \mathbb{E}(R(\mathbf{x})) = \mathbf{x}^\top \boldsymbol{\mu} \quad (2)$$

and variance given by

$$\sigma_p^2(\mathbf{x}) = \mathbf{x}^\top C \mathbf{x} \quad (3)$$

where $\mathbb{E}(\cdot)$ denotes the expectation, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ and C are the vector of expected values and the covariance matrix of the rates of return of the risky assets, respectively.

According to the modern portfolio theory (Markowitz, 1952), investors' decisions are based on the first two moments of the rate of return distribution of portfolios. The preference relation is thus the following

Definition 1. Let $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathbf{x}, \mathbf{y} \in \mathcal{X}$. Then \mathbf{x} (Pareto) dominates \mathbf{y} with respect to \mathcal{X} , in symbols $\mathbf{x} \prec \mathbf{y}$, if and only if $\mu_p(\mathbf{x}) \geq \mu_p(\mathbf{y})$ and $\sigma_p^2(\mathbf{x}) \leq \sigma_p^2(\mathbf{y})$ with at least one strict inequality.

The asset allocation problem reduces to a bi-objective optimization problem where portfolio expected return (2) is maximized and portfolio variance (3) is minimized. Thus, a portfolio $\mathbf{x}^* \in \mathcal{X}$ is (Pareto) optimal if and only if it is non-dominated with respect to \mathcal{X} , i.e., there does not exist another portfolio $\mathbf{x} \in \mathcal{X}$ that dominates \mathbf{x}^* . The subset of non-dominated portfolios in \mathcal{X} represents the so-called Pareto set. Its image in the mean-variance space is called efficient frontier, or Pareto front.

The starting point of our portfolio design is the model developed by Chang et al. (2000), which extends the standard mean-variance formulation by adding a cardinality constraint to control the number of active positions, i.e., assets with positive weight, taken in the portfolio and the corresponding costs.

Assuming each asset weight lies between a lower bound l_i and an upper bound u_i , with $0 < l_i < u_i \leq 1$, $i = 1, \dots, n$, the resulting mixed-integer bi-objective optimization problem can be written as

$$\text{minimize} \quad (-\mu_p(\mathbf{x}), \sigma_p^2(\mathbf{x})) \quad (4)$$

$$\text{subject to} \quad \sum_{i=1}^n x_i = 1 \quad (5)$$

$$\sum_{i=1}^n \delta_i = K \quad (6)$$

$$\delta_i \in \{0, 1\} \quad i = 1, \dots, n \quad (7)$$

$$\delta_i l_i \leq x_i \leq \delta_i u_i \quad i = 1, \dots, n. \quad (8)$$

Eq. (5) represents the budget constraint and ensures that all the capital is invested in the portfolio. The cardinality constraint given

in (6) imposes that exactly K active positions from the n assets present in the investible universe are selected. Note that K is a parameter decided by the decision-maker. The binary variables δ_i in (7), $i = 1, \dots, n$, are employed to model the inclusion or exclusion of asset i in the portfolio and have the following meaning

$$\delta_i = \begin{cases} 0, & \text{if asset } i \text{ is not included in the portfolio} \\ 1, & \text{if asset } i \text{ is included in the portfolio.} \end{cases} \quad (9)$$

Constraint (8) ensures that, if asset i is in the portfolio, i.e. $\delta_i = 1$, its proportion x_i lies between l_i and u_i . In the case the asset is not included in the portfolio, i.e., $\delta_i = 0$, constraint (8) imposes the corresponding weight x_i to be 0. Floor constraints, related to l_i , are used to control costs while ceiling constraints, related to u_i , prevent portfolios from excessive concentration on a singular asset. Note that the assumption $l_i > 0$ imposes no-shortselling.

From now on, the cardinality constrained portfolio optimization problem (4)–(8) will be denoted by CCPOP.

2.2. Risk-parity portfolios

CCPOP simultaneously manages profit and risk (Beasley, 2013). However, this approach may be unsuitable for investors having reduction of large losses and diversification across sources of risk as primary goals. An allocation strategy solely relying on risk views to manage risk may be advisable in this case (de Carvalho et al., 2012).

The level of diversification within a portfolio can be measured by the distribution of risk contributions of its constituents (Kolm et al., 2014). By representing the risk of a portfolio as the volatility of its rates of return, i.e. $\sigma_p(\mathbf{x}) = \sqrt{\mathbf{x}^\top C \mathbf{x}}$, we have the following risk decomposition

$$\sigma_p(\mathbf{x}) = \sum_{i=1}^n x_i \frac{\partial \sigma_p(\mathbf{x})}{\partial x_i} = \sum_{i=1}^n x_i \left(\frac{C \mathbf{x}}{\sqrt{\mathbf{x}^\top C \mathbf{x}}} \right)_i = \sum_{i=1}^n x_i \frac{(C \mathbf{x})_i}{\sqrt{\mathbf{x}^\top C \mathbf{x}}} \quad (10)$$

where $(C \mathbf{x})_i$ is the i th component of the vector $C \mathbf{x}$. Denote the risk contribution from asset i by $RC_i(\mathbf{x}) = x_i \frac{\partial \sigma_p(\mathbf{x})}{\partial x_i}$, and define the corresponding relative risk contribution as

$$RRC_i(\mathbf{x}) = \frac{RC_i(\mathbf{x})}{\sigma_p(\mathbf{x})} = \frac{x_i (C \mathbf{x})_i}{\sigma_p^2(\mathbf{x})}. \quad (11)$$

The paper will focus on the following risk control strategy known as risk parity (Roncalli, 2013).

Definition 2. Let C be a positive definite covariance matrix of order n , a portfolio $\mathbf{x} \in \mathcal{X}$ with K active positions is called a risk

parity (RP) portfolio with respect to C if it satisfies the condition

$$RC_i(\mathbf{x}) = \frac{\sigma_p(\mathbf{x})}{K} \quad (12)$$

for all i such that $\delta_i = 1$, where δ_i is given in (9).

An RP portfolio is thus characterized by the requirement of having equal risk contributions from each active asset. Based on (11), a portfolio $\mathbf{x} \in \mathcal{X}$ attains risk parity if and only if $RRC_i(\mathbf{x}) = 1/K$, for all i such that $\delta_i = 1$.

This type of portfolio presents appealing properties for diversification purposes, as reported in Maillard et al. (2010). However, existence and uniqueness of a solution satisfying (12) are only guaranteed in the full invested and long-only setting (Kaya and Lee, 2012). Thus, using numerical methods becomes fundamental to solve RP portfolio selection problems with more complex constraints.

The problem of identifying an RP portfolio in a feasible set \mathcal{X} can be rewritten as an optimization problem with a cost function quantifying the deviation from risk parity. A first example of such a function considers all pairwise differences in squared among the active assets as follows

$$f_1^{RP}(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \delta_i \delta_j (x_i(C\mathbf{x})_i - x_j(C\mathbf{x})_j)^2. \quad (13)$$

One can alternatively measure deviations of RRCs from the average value in squared or absolute sense, obtaining the following cost functions

$$f_2^{RP}(\mathbf{x}) = \sum_{i=1}^n \delta_i \left(\frac{x_i(C\mathbf{x})_i}{\mathbf{x}^\top C\mathbf{x}} - \frac{1}{K} \right)^2 \quad (14)$$

and

$$f_3^{RP}(\mathbf{x}) = \sum_{i=1}^n \delta_i \left| \frac{x_i(C\mathbf{x})_i}{\mathbf{x}^\top C\mathbf{x}} - \frac{1}{K} \right| \quad (15)$$

respectively. In all the proposed cases, portfolios with cost function value equal to zero achieve risk parity.

Due to the non-smoothness and non-convexity of f_1^{RP} , f_2^{RP} and f_3^{RP} , approximation and heuristic approaches are used to solve the corresponding minimization problems (see, for instance, Hochreiter, 2015 and Feng and Palomar, 2015).

2.3. Reconciling profit with risk control

In this paper, we propose to manage performance and risk dimensions at the same time by integrating the RP condition (12) into CCPOP. The resulting portfolio optimization problem, which will be denoted by RP-CCPOP, can be formulated as follows

$$\text{minimize} \quad (-\mathbf{x}^\top \boldsymbol{\mu}, \mathbf{x}^\top C\mathbf{x}) \quad (16)$$

$$\text{subject to} \quad \sum_{i=1}^n x_i = 1 \quad (17)$$

$$\sum_{i=1}^n \delta_i = K \quad (18)$$

$$\delta_i \in \{0, 1\} \quad i = 1, \dots, n \quad (19)$$

$$\delta_i l_i \leq x_i \leq \delta_i u_i \quad i = 1, \dots, n. \quad (20)$$

$$x_i(C\mathbf{x})_i = \delta_i \frac{\mathbf{x}^\top C\mathbf{x}}{K} \quad i = 1, \dots, n \quad (21)$$

where $0 < l_i < u_i \leq 1$, for $i = 1, \dots, n$, and the equality constraint (21) represents the risk parity condition.

3. Constrained multi-objective particle swarm optimization

Particle swarm optimization (PSO) is a distributed behavioral algorithm that performs multidimensional search by mimicking the movements of a bird flock or a fish schooling that searches for food (Eberhart and Kennedy, 1995). The PSO mechanism is based on the communication of information about good solutions through the swarm. In this manner, the particles will tend to move toward good areas in the search space.

The adaptation of this paradigm to deal with portfolio management problems has proved to be very effective (see, for instance, Ertenlice and Kalayci, 2018 and the references therein for a survey on the subject). In particular, the promising results in the solution of CCPOP obtained by Mishra et al. (2014) justify our development of a new version of the multi-objective particle swarm optimization (MOPSO) algorithm to handle cardinality, quantity and risk parity constraints at the same time.

3.1. Multi-objective particle swarm portfolio optimization

The algorithm that we propose operates on a set of S candidate solutions, denoted as \mathcal{P} . At each iteration t , the s th candidate solution is formed by a pair of n -dimensional real vectors $(\mathbf{x}_s(t), \mathbf{v}_s(t))$, for $s = 1, \dots, S$, where $\mathbf{x}_s(t) = (x_{s1}(t), \dots, x_{sn}(t))^\top$ gathers the weights of the s -th portfolio to be optimized and $\mathbf{v}_s(t) = (v_{s1}(t), \dots, v_{sn}(t))^\top$ is an update vector conveying the information about possible rebalancing of \mathbf{x}_s . The set of decision variables δ_i , $i = 1, \dots, n$, in Problem (16)–(21) are implicitly handled as follows: for each weight vector $\mathbf{x}_s(t)$, the K assets with the highest weights enter the corresponding portfolio while zero weight is assigned to the remaining $n - K$ assets. In this manner, the cardinality constraint is implicitly satisfied while the assumption $l_i > 0$ for all $i = 1, \dots, n$ avoids the indecisiveness case represented by $\delta_i = 1$ and $x_i = 0$.

Using a Pareto ranking scheme based on Definition 1, there are two types of promising solutions that play a part in determining $\mathbf{v}_s(t)$: from one hand, the most recent non-dominated portfolio \mathbf{x}_s^{best} related to candidate solution s and, from the other hand, all the non-dominated portfolios found along the optimization process. We denote this second set as \mathcal{G}^{best} and fix its maximal size to S . Following the procedure stated in Coello et al. (2004), if the number of elements of \mathcal{G}^{best} exceeds S , some portfolios are pruned as follows: first, an adaptive grid divides the mean-variance space into a number div of squares, then the most densely populated squares are truncated. At this point, a fitness value is assigned to each square that contains members of \mathcal{G}^{best} equal to the result of dividing any number $\gamma > 1$ by the number of resident portfolios. Thus, a more densely populated square is given a lower score. A roulette wheel selection of a square in the objective space according to its score is performed and a member \mathbf{g}_s^{best} of that square is randomly chosen. The update rule for \mathbf{v}_s becomes

$$\mathbf{v}_s(t+1) = w\mathbf{v}_s(t) + c_1 \text{rand}_1(\mathbf{x}_s^{best} - \mathbf{x}_s(t)) + c_2 \text{rand}_2(\mathbf{g}_s^{best} - \mathbf{x}_s(t)) \quad (22)$$

for $s = 1, \dots, S$. The term $w\mathbf{v}_s(t)$, called in literature momentum, regulates the importance of $\mathbf{v}_s(t)$ in $\mathbf{v}_s(t+1)$ through the inertia weight w . The second term, called cognitive component, quantifies how much displacement, starting at $\mathbf{x}_s(t)$, is needed to reach the corresponding non-dominated portfolio \mathbf{x}_s^{best} . The social component is represented by the third addend and quantifies the performance of portfolio s relative to the best portfolio selected in \mathcal{G}^{best} .

Higher values for w result in a good global search characteristic, while lower values produce a refined localized search. Parameters c_1 and c_2 are positive acceleration coefficients used to weight the contribution of cognitive and social components respectively, while

$rand_1$ and $rand_2$ represent two random numbers generated by a uniform distribution on the interval $[0,1]$.

The portfolio composition of the s th candidate solution can now be modified according to the following formula

$$\mathbf{x}_s(t+1) = \mathbf{x}_s(t) + \mathbf{v}_s(t+1) \quad (23)$$

for $s = 1, \dots, S$.

All the weight vectors \mathbf{x}_s^{best} and the external archive \mathcal{G}_{best} are updated after every iteration.

3.2. Solution clamping and mutation

Since large weight updates may cause portfolios to leave the domain boundaries, we consider the following clamping procedure proposed by [Xu and Chen \(2006\)](#) to control the portfolio rebalancing components

$$v_{si}(t) = \begin{cases} v_i^{\min}, & \text{if } v_{si}(t) < v_i^{\min} \\ v_i^{\max}, & \text{if } v_{si}(t) > v_i^{\max} \\ v_{si}(t), & \text{otherwise} \end{cases} \quad (24)$$

for $s = 1, \dots, S$ and $i = 1, \dots, n$, where v_i^{\min} and v_i^{\max} are minimum and maximum percentage of capital that are allowed to be moved for asset i respectively.

In order to promote diversity and higher explorative capability, we assign to $v_{si}(t)$ a probability p_{m_1} of being specified in a different direction as in [Mishra et al. \(2014\)](#). Moreover, a swap mutation operator is applied to each portfolio $\mathbf{x}_s(t)$ in order to explore un-reached areas of the search space. The operator swaps the weights of two constituents in every candidate solution with a probability p_{m_2} as follows. Let $\mathbf{x} \in \mathcal{X}$ be a portfolio subject to swap mutation and a and b be two randomly chosen positions such that $x_a = 0$ and $l_b \leq x_b \leq u_b$ then the mutated \mathbf{x} is defined componentwise as

$$\tilde{x}_i = \begin{cases} x_i, & \text{if } i \neq a \text{ and } i \neq b \\ l_a + \frac{x_b - l_b}{u_b - l_b} (u_a - l_a), & \text{if } i = a \\ 0, & \text{if } i = b. \end{cases} \quad (25)$$

The new candidate solution $\tilde{\mathbf{x}}$ maintains K active positions and satisfies quantity constraints for a and removes asset b from the portfolio.

3.3. Candidate solutions initialization and constraint-handling techniques

The set of candidate solutions is initialized randomly with assets positions uniformly distributed between their respective bounds l_i and u_i . By applying the proposed selection method, we obtain a set of particles that satisfy cardinality, floor and ceiling constraints. The repair transformations reported below deal with the full investment condition.

Proposition 1 ([Meghwani and Thakur, 2017](#)). Let $\mathbf{x} \in \mathbb{R}^n$ and $I_+ = \{i = 1, \dots, n \mid x_i > 0\}$, $|I_+| = K$, and assume

1. $l_i \leq x_i \leq u_i$ for all $i \in I_+$,
2. $\sum_{i \in I_+} l_i < 1$,
3. $\sum_{i \in I_+} u_i > 1$.

Then, the portfolio $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)^\top$ defined as $\tilde{x}_i = 0$ for $i \notin I_+$, and, for all $i \in I_+$,

$$\tilde{x}_i = \begin{cases} u_i - \frac{(u_i - x_i)}{\sum_{j \in I_+} (u_j - x_j)} (\sum_{j \in I_+} u_j - 1), & \text{if } \sum_{j \in I_+} x_j < 1 \\ x_i, & \text{if } \sum_{j \in I_+} x_j = 1 \\ l_i + \frac{(x_i - l_i)}{\sum_{j \in I_+} (x_j - l_j)} (1 - \sum_{j \in I_+} l_j), & \text{if } \sum_{j \in I_+} x_j > 1 \end{cases} \quad (26)$$

preserves the invested positions of \mathbf{x} , satisfies the constraints $l_i \leq \tilde{x}_i \leq u_i$, for all $i \in I_+$, and $\sum_{j=1}^n \tilde{x}_j = 1$.

This update rule does not modify portfolio constituents and involves proportionate distribution of allocation over non-zero assets in order to follow the required budget, floor and ceiling constraints.

The risk parity equality constraints in (21) are dealt with by applying a suitable constraint-handling technique based on the Pareto dominance principle. In particular, we will analyze the hybridization capabilities of the following procedures with the repair mechanism (26):

1. The constrained-domination principle proposed in [Deb \(2000\)](#);
2. The constrained method developed in [Zhang and Rangaiah \(2012\)](#);
3. The self-adaptive penalty technique presented in [Woldesenbet et al. \(2009\)](#).

Before introducing the characteristics of these constraint-handling methods, we transform (21) into inequality constraints as follows

$$\left| x_i(\mathbf{C}\mathbf{x})_i - \delta_i \frac{\mathbf{x}^\top \mathbf{C}\mathbf{x}}{K} \right| \leq \tau, \quad i = 1, \dots, n \quad (27)$$

where $\tau \geq 0$ is a tolerance parameter according to which infeasible solutions close to the feasible region are retained in order to improve the search effectiveness.

The violation of constraint i in (21) for an individual \mathbf{x} can be measured as

$$c_i(\mathbf{x}, \tau) = \max \left\{ \left| x_i(\mathbf{C}\mathbf{x})_i - \delta_i \frac{\mathbf{x}^\top \mathbf{C}\mathbf{x}}{K} \right| - \tau, 0 \right\}, \quad i = 1, \dots, n. \quad (28)$$

Let c_i^{\max} denote the maximum of $c_i(\mathbf{x}, \tau)$ over all portfolios in the current iteration and define

$$\phi_i(\mathbf{x}, \tau) = \begin{cases} \frac{c_i(\mathbf{x}, \tau)}{c_i^{\max}}, & \text{if } c_i^{\max} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (29)$$

for $i = 1, \dots, n$, the overall constraint violation can be expressed as

$$\phi(\mathbf{x}, \tau) = \frac{1}{n} \sum_{i=1}^n \phi_i(\mathbf{x}, \tau). \quad (30)$$

Note that c_i^{\max} varies during the evolution in order to balance the contribution of each constraint in the problem irrespective of their differing ranges.

3.3.1. Constrained domination principle

The constrained-domination principle, shortly CDP, compares pairwise solutions according to the following three feasibility criteria:

- i) When two feasible portfolios are compared, the one Pareto dominating the other is chosen;
- ii) When a feasible portfolio is compared with an infeasible portfolio, the feasible one is chosen;
- iii) When two infeasible portfolios are compared, the one with the lowest overall constraint violation is chosen.

Therefore, the CDP always considers feasible solutions better than infeasible ones. When two feasible solutions are considered, they are only compared in terms of their objective function values since the corresponding overall constraint violations are zero. In the case of two infeasible solutions, the comparison is based on the sum of constraints violations. This rule aims at pushing infeasible individuals to the feasible region.

3.3.2. Self-adaptive tolerance constrained domination method

Our variant of the self-adaptive tolerance constrained method proposed in Zhang and Rangaiah (2012) first exploits the projection mechanism of Proposition 1 for the budgeting and quantity constraints and then adaptively relaxes the risk parity constraint (21) according to Eqs. (28)–(30), with τ changing over time as follows

$$\tau(t) = \begin{cases} \text{median} \{ \phi(\mathbf{x}_s(t), 0) \mid s = 1, \dots, S \}, & \text{if } t = 0 \\ \tau(t-1)(1-r_f(t)), & \text{otherwise} \end{cases} \quad (31)$$

where $r_f(t)$ denotes the feasibility ratio of the portfolios at iteration t , i.e. the ratio of the number of feasible portfolios at time t to the size S of the set \mathcal{P} .

Since the extent of relaxation depends on the feasibility of the elements in \mathcal{P} , when fewer portfolios are feasible, the rate of reducing τ will be slower. Hence, more generations will be required to find feasible solutions. For example, when all individuals in the current set are infeasible, τ will not be reduced until some feasible portfolios are found.

By incorporating this adaptive relaxation of constraints into the CDP approach, the method provides a dynamic sorting that depends on the number of risk parity deviations in \mathcal{P} . More specifically, we have the following

Definition 3. Let \mathbf{x}, \mathbf{y} be two portfolios satisfying (17)–(20) and τ be defined as in (31), then we say that \mathbf{x} τ -dominates \mathbf{y} and write $\mathbf{x} \prec_{\tau} \mathbf{y}$ if and only if

$$\begin{cases} \mathbf{x} \prec \mathbf{y}, & \text{if } \phi(\mathbf{x}, 0) \leq \tau(t) \text{ and } \phi(\mathbf{y}, 0) \leq \tau(t) \\ \mathbf{x} \prec \mathbf{y}, & \text{if } \phi(\mathbf{x}, 0) = \phi(\mathbf{y}, 0) \\ \phi(\mathbf{x}, 0) < \phi(\mathbf{y}, 0), & \text{otherwise} \end{cases} \quad (32)$$

where $\mathbf{x} \prec \mathbf{y}$ denotes Pareto dominance as stated in Definition 1 and $\tau(t)$ is the tolerance level at time t .

This ordering compares two solutions using only their objective function values if both solutions are feasible, slightly feasible (as determined by the value of τ) or even if they have the same sum of constraint violations. Conversely, if both solutions are infeasible, they are compared based on only their overall constraint violations. In the extreme case in which $r_f(t) = 1$, (32) reduces to the CDP feasibility rules in that generation.

3.3.3. Self-adaptive penalty constrained domination method

Similar to the method just described, the self-adaptive penalty technique uses the information gathered from the search process to control the amount of penalty added to infeasible individuals. More specifically, two types of penalty are added to each infeasible individual to identify the best infeasible individuals in the current population. The amount of the added penalties depends on the number of feasible individuals present in the current population. On the one hand, if there is a small proportion of feasible individuals, then infeasible individuals with a higher amount of constraint violation will have a higher amount of penalty. On the other hand, if there is a large number of feasible individuals, then small penalties will be added to infeasible individuals with high fitness values. These two penalties will allow the algorithm to switch between finding more solutions that are feasible and searching for optimal solutions at any time during the search process.

The procedure involves the following computations. First, the values of the two objective functions, namely μ_p and σ_p^2 , of each candidate portfolio \mathbf{x}_s in \mathcal{P} are normalized using the formulas

$$f_1(\mathbf{x}_s) = \frac{\max_{s'} \mu_p(\mathbf{x}_{s'}) - \mu_p(\mathbf{x}_s)}{\max_{s'} \mu_p(\mathbf{x}_{s'}) - \min_{s'} \mu_p(\mathbf{x}_{s'}) + \epsilon} \quad (33)$$

and

$$f_2(\mathbf{x}_s) = \frac{\sigma_p^2(\mathbf{x}_s) - \min_{s'} \sigma_p^2(\mathbf{x}_{s'})}{\max_{s'} \sigma_p^2(\mathbf{x}_{s'}) - \min_{s'} \sigma_p^2(\mathbf{x}_{s'}) + \epsilon} \quad (34)$$

where ϵ is a positive value for preventing the denominators from becoming zero in the case $\min_{s'} \mu_p(\mathbf{x}_{s'}) = \max_{s'} \mu_p(\mathbf{x}_{s'})$ or $\min_{s'} \sigma_p^2(\mathbf{x}_{s'}) = \max_{s'} \sigma_p^2(\mathbf{x}_{s'})$. We have set $\epsilon = 10^{-6}$ in the numerical analysis as Ishibuchi et al. (2017). These scaling functions map the original objective space \mathbb{R}^2 to $[0, 1]^2$.

The overall constraint violation of the risk parity is then calculated according to Eqs. (28)–(30).

Depending on the feasibility ratio r_f , the final value of the j th modified objective function for \mathbf{x}_s is defined as

$$F_j(\mathbf{x}_s) = \begin{cases} \phi(\mathbf{x}_s, \tau), & \text{if } r_f = 0 \\ \sqrt{f_j(\mathbf{x}_s)^2 + \phi(\mathbf{x}_s, \tau)^2} + (1-r_f)\phi(\mathbf{x}_s, \tau) + r_f f_j(\mathbf{x}_s), & \text{if } r_f \neq 0 \text{ and } \mathbf{x}_s \notin \mathcal{X} \\ f_j(\mathbf{x}_s), & \text{if } r_f \neq 0 \text{ and } \mathbf{x}_s \in \mathcal{X} \end{cases} \quad (35)$$

with $j = 1, 2$. The population can now be sorted by CDP based on F_1 and F_2 .

The pseudocode of HMOPSO is outlined in Algorithm 1, where

Algorithm 1: HMOPSO for RP-CCPOP.

Input : $\mathbf{l}, \mathbf{u}, t_{\max}, S, c_1, c_2, w, p_{m_1}, p_{m_2}, div, \gamma, \mathbf{v}^{\min}, \mathbf{v}^{\max}, \tau, CHT$

Output: $\mathcal{G}_{\text{best}}$

- 1 Randomly generate a set of portfolios $\mathbf{x}_s(0), s = 1, \dots, S$
 - 2 Set the corresponding rebalancing vectors $\mathbf{v}_s(0) = 0$
 - 3 Set $\mathcal{P} = \{(\mathbf{x}_s(0), \mathbf{v}_s(0)) : s = 1, \dots, S\}$
 - 4 Modify each member of \mathcal{P} using (26) and evaluate its overall constraint violation for the risk parity constraint using (28)–(30)
 - 5 Initialize $\mathbf{x}_s^{\text{best}} = \mathbf{x}_s(0), s = 1, \dots, S$
 - 6 Initialize $\mathcal{G}^{\text{best}}$ based on CHT
 - 7 Let $t = 0$
 - 8 **while** $t \leq t_{\max}$ **do**
 - 9 **for** $s = 1$ **to** S **do**
 - 10 Select $\mathbf{g}_s^{\text{best}} \in \mathcal{G}^{\text{best}}$ on the basis of CHT
 - 11 Update $\mathbf{x}_s^{\text{best}}$ on the basis of CHT and $\mathbf{x}_s(t)$
 - 12 **end**
 - 13 $t = t + 1$
 - 14 Update $\mathbf{v}_s(t)$ using (22) and adjust it through (24)
 - 15 **for** $i = 1$ **to** n **do**
 - 16 **if** $\text{rand} < p_{m_1}$ **then**
 - 17 $\mathbf{v}_{si}(t) = -\mathbf{v}_{si}(t)$
 - 18 **end**
 - 19 **end**
 - 20 Update \mathcal{P} using (23)
 - 21 Modify \mathcal{P} using (26)
 - 22 Mutate \mathcal{P} using (25)
 - 23 Evaluate \mathcal{P} according to CHT
 - 24 Update $\mathcal{G}^{\text{best}}$ according to CHT
 - 25 **end**
-

CHT indicates the variant of the hybrid operator that is implemented.

4. Performance metrics

Several criteria are considered to assess the quality of the proposed HMOPSO and the profitability of the designed portfolio optimization model.

4.1. In-sample performance metrics

According to the prescriptions in [Fonseca et al. \(2005\)](#), multiple runs are employed in order to obtain a sample of competing fronts for each algorithm and for each instance of the test problems.

Since the true Pareto fronts are unknown, the effectiveness of each procedure has been measured on the basis of three criteria:

- i) The diversification level attained by the optimal portfolios;
- ii) The capacity to generate an appropriate number of non-dominated solutions that satisfy the risk parity principle at least approximately;
- iii) The capacity of the approximated Pareto front to converge toward the true Pareto front by spanning a suitable portion of non-dominated risk-return profiles.

Accordingly, optimal solution sets with large number of well-diversified feasible non-dominated portfolios, enclosing large areas of the risk-return space and scattering evenly are generally desirable.

Two performance metrics are used to summarize the in-sample experimental results. The first indicator is the Herfindahl index of risk contributions ([Roncalli, 2013](#)), defined as

$$h(\mathbf{x}) = \sum_{i=1}^n \left(\frac{x_i(\mathbf{C}\mathbf{x})_i}{\mathbf{x}^T \mathbf{C}\mathbf{x}} \right)^2. \quad (36)$$

It simultaneously quantifies the magnitude of deviations from risk parity and the diversification level of the portfolio $\mathbf{x} \in \mathcal{X}$. If there are K active positions, then it is easy to verify that $1/K \leq h(\mathbf{x}) \leq 1$. When $h(\mathbf{x}) = 1$, the risk contribution is concentrated in only one asset. For $h(\mathbf{x}) = 1/K$ the portfolio is fully diversified among the active assets, thus satisfying condition (12).

Preliminary to the introduction of the second performance metric, which will be used to compare the overall quality of the approximation sets, is the notion of hypervolume indicator ([Zitzler and Thiele, 1998](#)). The hypervolume of a given (approximated) Pareto front PF measures the size of the region of the objective space dominated by at least one of the members of PF and bounded by a selected reference point, that is dominated by all the points of PF . In our context, with two objectives to minimize, the reference point is $\mathbf{r} = (r_1, r_2)$ and the hypervolume is computed by

$$HV(PF, \mathbf{r}) = \text{Leb} \left(\bigcup_{(y_1, y_2) \in PF} [y_1, r_1] \times [y_2, r_2] \right) \quad (37)$$

where $\text{Leb}(A)$ is the Lebesgue measure of a set A . In this manner, HV evaluates the closeness of the solutions to the optimal set and captures the spread of the solutions over the objective space. Thus, larger hypervolume indicator values imply a larger distance from the reference point and indicate solutions are scattered more evenly in the objective space.

Despite its effectiveness in evaluating an approximation, the hypervolume indicator presents some drawbacks. From one hand, it is sensitive to the relative scaling of the objectives and to the presence or absence of extremal points in a front ([Zitzler et al., 2007](#); [Li et al., 2015](#)). This results in favoring an approximation set which has better converged for the least-scaled objective function or has more points that are extremal. On the other hand, hypervolume depends strongly on the choice of the reference point ([Auger et al., 2009](#)). Different reference points may result in different evaluations. Moreover, all possible vectors in the comparison must dominate the reference point; otherwise, the evaluations of the indica-

tor may not make any sense, for example, reporting negative values ([Lizárraga, 2009](#)).

In this study, assuming the approximation sets have cardinality $S \geq 2$, the following procedure is proposed to avoid these pitfalls. First, the objective function values are projected onto the interval $[0, 1]$ by Eqs. (33) and (34). Second, noting that in the $[0, 1]^2$ space the worst solution for our optimization problem is represented by the point $(1, 1)$, we choose as reference point the vector

$$\mathbf{r} = (1 + r, 1 + r) \quad (38)$$

where $r > 0$. In the experiments we will set $r = \sqrt{2}/(S - 1)$, adapting ([Cao et al., 2015](#)). With this choice of \mathbf{r} , fronts with uniformly distributed points are preferred. The evaluation of the hypervolume is conducted as follows.

Proposition 2. Let $\widehat{PF} = \{(y_1^s, y_2^s) \in [0, 1]^2 \mid s = 1, \dots, S\}$, $S \geq 2$, be a normalized discrete Pareto front for the RP-CCPOP, then the corresponding hypervolume indicator with respect to the reference point (38) can be formulated as

$$HV(\widehat{PF}, \mathbf{r}) = \sum_{s=1}^{S-1} (y_1^{s+1} - y_1^s)(1 - y_2^s) + (1 - y_1^s)(1 - y_2^s) + r(1 - y_1^s) + r(1 - y_2^s) + r^2. \quad (39)$$

Proof. Without loss of generality, let us assume that the solutions in \widehat{PF} are sorted by the first criterion in ascending order, i.e., $y_1^s \leq y_1^{s+1}$, for $s = 1, \dots, S - 1$. The hypervolume indicator with respect to $\mathbf{r} = (1 + r, 1 + r)$ is then

$$\begin{aligned} HV(\widehat{PF}, \mathbf{r}) &= (y_1^2 - y_1^1)(1 - y_2^1 + r) + \dots + (y_1^S - y_1^{S-1})(1 - y_2^{S-1} + r) \\ &\quad + (1 - y_1^S + r)(1 - y_2^S + r) \\ &= \sum_{s=1}^{S-1} (y_1^{s+1} - y_1^s)(1 - y_2^s) + (1 - y_1^S)(1 - y_2^S) + r(1 - y_1^1) \\ &\quad + r(1 - y_2^S) + r^2. \end{aligned}$$

□

We can now describe the second performance metric used in the experiments, the so-called contribution rate indicator ([Cao et al., 2015](#)). To this end, let PF_1 , PF_2 and PF_3 be the approximation fronts generated by the three versions of HMOPSO and denote by PF their combination based on Pareto dominance definition. Since the true Pareto front is unknown and PF is the best available approximation, PF is regarded as a surrogate for the true Pareto front. The procedure to compute the contribution rate indicator is as follows.

1. From each approximated Pareto front PF_i , $i = 1, 2, 3$, obtain the subset $PF'_i = PF_i \cap PF$, i.e., the set of criterion vectors in PF_i which are not dominated by PF .
2. Normalize PF and PF'_i , $i = 1, 2, 3$, by employing (33) and (34) with $\epsilon = 0$ and using as extreme values those in PF . Denote these sets as \widehat{PF} and \widehat{PF}'_i , $i = 1, 2, 3$, respectively.
3. Calculate the hypervolume indicators $HV(\widehat{PF}, \mathbf{r})$ and $HV(\widehat{PF}'_i, \mathbf{r})$, $i = 1, 2, 3$, with respect to the reference point (38) according to Eq. (39).
4. Define the contribution rate associated to each approximation set as

$$CR(\widehat{PF}'_i, \widehat{PF}, \mathbf{r}) = \frac{HV(\widehat{PF}'_i, \mathbf{r})}{HV(\widehat{PF}, \mathbf{r})}, \quad i = 1, \dots, n. \quad (40)$$

Values of $CR(\widehat{PF}'_i, \widehat{PF}, \mathbf{r})$ close to 1 suggest that the i th approximation front is very close to the surrogate Pareto front. Moreover, if $CR(\widehat{PF}'_i, \widehat{PF}, \mathbf{r}) > CR(\widehat{PF}'_j, \widehat{PF}, \mathbf{r})$, with $i \neq j$, then PF_i is a better approximation than PF_j , and the comparison of the contribution rate gives a sense of the magnitude of the difference.

Table 2

Data sets from [Bruni et al. \(2016\)](#) with corresponding time window, number of weeks and number of market constituents (n) used in the estimation of parameters, and portfolio cardinalities K imposed in each application of HMOPSO.

Data set name	Time window	Weeks	Assets (n)	Cardinality (K)
DowJones	02/1990 – 04/2016	1363	28	5 10 15 20
FF49Industries	07/1969 – 07/2015	2325	49	5 10 20 40
NASDAQ100	11/2004 – 04/2016	596	82	5 10 20 40

Table 3

Configuration of HMOPSO parameters for the experimental analysis.

Parameter	Value
w	0.4
c_1	1.494
c_2	1.494
v^{\min}	0.002
v^{\max}	0.2
p_{m_1}	0.5
p_{m_2}	0.1
div	10
γ	10
S	200
t_{\max}	1000

Table 4

Average number of non-dominated solutions included into the external archive $\mathcal{G}_{\text{best}}$ of the proposed algorithm adopting the constrained domination principle (HMOPSO-CDP), with self-adaptive tolerance (HMOPSO-ST) and using self-adaptive penalty (HMOPSO-SP) classified by data set and cardinality instance (K).

Data set	K	HMOPSO-CDP	HMOPSO-ST	HMOPSO-SP
DowJones	5	199.00	198.90	58.95
	10	186.20	165.75	56.45
	15	126.75	119.95	42.55
	20	100.00	101.50	25.25
FF49Industries	5	199.95	197.90	36.00
	10	171.30	150.45	46.95
	20	86.25	77.50	21.00
	40	79.65	80.85	9.70
NASDAQ100	5	199.10	199.55	33.50
	10	200.00	199.90	55.60
	20	162.95	145.30	34.50
	40	57.80	62.50	16.15

Table 5

Optimal values of Herfindahl index of risk contributions, h_{\min} , with the corresponding mean and standard deviation values attained by the proposed algorithm adopting the constrained domination principle (HMOPSO-CDP), with self-adaptive tolerance (HMOPSO-ST) and using self-adaptive penalty (HMOPSO-SP) classified by data set and cardinality instance (K).

Data set	K	h_{\min}	HMOPSO-CDP		HMOPSO-ST		HMOPSO-SP	
			Mean	std	Mean	std	Mean	std
DowJones	5	0.2	0.2177	0.0427	0.2341	0.0734	0.2081	0.0037
	10	0.1	0.1203	0.0328	0.1245	0.0429	0.1092	0.0043
	15	0.0667	0.0956	0.0332	0.0986	0.0441	0.0758	0.0039
	20	0.05	0.0889	0.0350	0.0896	0.0391	0.0572	0.0034
FF49Industries	5	0.2	0.2163	0.0307	0.2150	0.0306	0.2092	0.0045
	10	0.1	0.1136	0.0187	0.1102	0.0186	0.1078	0.0035
	20	0.05	0.0682	0.0189	0.0608	0.0168	0.0564	0.0030
	40	0.025	0.0366	0.0075	0.0358	0.0077	0.0268	0.0015
NASDAQ100	5	0.2	0.2072	0.0107	0.2140	0.0262	0.2052	0.0027
	10	0.1	0.1073	0.0067	0.1082	0.0125	0.1068	0.0027
	20	0.05	0.0580	0.0054	0.0564	0.0049	0.0567	0.0024
	40	0.025	0.0302	0.0041	0.0287	0.0032	0.0283	0.0013

Table 6

Wilcoxon rank-sum test for the Herfindahl index of risk contributions (36): HMOPSO using self-adaptive penalty (HMOPSO-SP) vs HMOPSO adopting the constrained domination principle (HMOPSO-CDP) and HMOPSO with self-adaptive tolerance (HMOPSO-ST) classified by data set and cardinality (K). The symbol “ \uparrow ” means that HMOPSO-SP has yielded better results than the algorithm in the column at a 5% confidence level, “ \downarrow ” indicates that the algorithm in the column is statistically better than HMOPSO-SP and “ \leftrightarrow ” is used when there is no statistical difference between algorithms.

Data set	K	HMOPSO-CDP	HMOPSO-ST
DowJones	5	\downarrow	\uparrow
	10	\uparrow	\uparrow
	15	\uparrow	\uparrow
	20	\uparrow	\uparrow
FF49Industries	5	\leftrightarrow	\downarrow
	10	\uparrow	\downarrow
	20	\uparrow	\downarrow
	40	\uparrow	\uparrow
NASDAQ100	5	\uparrow	\uparrow
	10	\leftrightarrow	\leftrightarrow
	20	\uparrow	\downarrow
	40	\uparrow	\downarrow

A rigorous assessment of the robustness of the results can then be performed on the basis of nonparametric inference testing. In particular, we use the Wilcoxon’s rank-sum test for independent samples at the 5% significance level to judge whether the results obtained with the best performing procedure differ from the results of the rest of competitors in a statistically significant way.

4.2. Out-of-sample performance metrics

We evaluate the out-of-sample performance of each portfolio optimization model with a set of statistics. Let us denote by t_0 and T the last in-sample and the last out-of-sample time respectively and let us indicate by r_t^{out} , $t = t_0 + 1, \dots, T$, the portfolio rates of return for each strategy in the out-of-sample period, respectively. The first performance measure we introduce is the Sharpe ratio ([Sharpe, 1966](#)), which evaluates the compensation earned per unit of portfolio total risk. It is defined as the ratio between the annualized average μ^{out} of r_t^{out} , and the annualized sample standard deviation σ^{out} as follows

$$SR = \frac{\mu^{\text{out}}}{\sigma^{\text{out}}}. \quad (41)$$

Table 7

Mean and standard deviation values of the CR performance metric (40) for the proposed algorithm adopting the constrained domination principle (HMOPSO-CDP), with self-adaptive tolerance (HMOPSO-ST) and using self-adaptive penalty (HMOPSO-SP) divided by data set and cardinality (K).

Data set	K	HMOPSO-CDP		HMOPSO-ST		HMOPSO-SP	
		Mean	std	Mean	std	Mean	std
DowJones	5	0.8635	0.0506	0.9916	0.0049	0.4616	0.2104
	10	0.9634	0.0144	0.9506	0.0220	0.5480	0.0561
	15	0.9644	0.0165	0.9408	0.0165	0.4226	0.1212
	20	0.9744	0.0089	0.9445	0.0222	0.1258	0.1614
FF49Industries	5	0.9750	0.0119	0.9636	0.0175	0.5824	0.0589
	10	0.9823	0.0104	0.9205	0.0334	0.5763	0.0651
	20	0.9788	0.0180	0.6527	0.2332	0.3835	0.0824
	40	0.9712	0.0175	0.9488	0.0329	0.0748	0.1270
NASDAQ100	5	0.8779	0.0867	0.9877	0.0093	0.5048	0.2633
	10	0.9559	0.0282	0.9733	0.0218	0.3772	0.3204
	20	0.9250	0.1065	0.8526	0.0852	0.3671	0.3115
	40	0.9279	0.0756	0.7940	0.1415	0.5391	0.1931

Table 8

Wilcoxon rank-sum test for the CR performance metric (40): HMOPSO adopting the constrained domination principle (HMOPSO-CDP) vs. HMOPSO using self-adaptive tolerance (HMOPSO-ST) and HMOPSO with self-adaptive penalty (HMOPSO-SP) classified by data set and cardinality (K). The symbol “ \uparrow ” means that HMOPSO-CDP has yielded better results than the algorithm in the column at a 5% confidence level, “ \downarrow ” indicates that the algorithm in the column is statistically better than HMOPSO-CDP and “ \leftrightarrow ” is used when there is no statistical difference between algorithms.

Data set	K	HMOPSO-ST	HMOPSO-SP
DowJones	5	\uparrow	\uparrow
	10	\leftrightarrow	\uparrow
	15	\uparrow	\uparrow
	20	\uparrow	\uparrow
FF49Industries	5	\uparrow	\uparrow
	10	\uparrow	\uparrow
	20	\uparrow	\uparrow
	40	\uparrow	\uparrow
NASDAQ100	5	\downarrow	\uparrow
	10	\downarrow	\uparrow
	20	\uparrow	\uparrow
	40	\uparrow	\uparrow

Table 9

Mean total CPU times (in seconds) for the proposed algorithm adopting the constrained domination principle (HMOPSO-CDP), with self-adaptive tolerance (HMOPSO-ST) and using self-adaptive penalty (HMOPSO-SP) classified by data set and cardinality (K).

Data set	K	HMOPSO-CDP	HMOPSO-ST	HMOPSO-SP
DowJones	5	33.43	32.36	19.15
	10	43.57	44.20	32.15
	15	52.89	53.41	45.26
	20	62.91	63.39	58.61
FF49Industries	5	43.40	45.35	32.14
	10	67.69	70.10	57.91
	20	112.95	115.75	111.15
	40	214.67	218.95	215.45
NASDAQ100	5	67.85	68.62	55.29
	10	114.49	118.31	104.98
	20	209.73	216.05	201.56
	40	404.56	406.80	396.87

The attractiveness of the two investment opportunities is also evaluated by the so-called Omega ratio (Keating and Shadwick, 2002), which gauges favorable and unfavourable excess returns with respect to a user-defined threshold. In this paper, the threshold is

set to 0 and the resulting measure can be written in the form

$$\Omega = \frac{\sum_{t=t_0+1}^T r_t^{out} \mathbf{1}_{\{r_t^{out} > 0\}}}{-\sum_{t=t_0+1}^T r_t^{out} \mathbf{1}_{\{r_t^{out} < 0\}}} \quad (42)$$

where, for a logical proposition q , $\mathbf{1}_q = 1$ if q is true, and $\mathbf{1}_q = 0$ otherwise. Higher values of SR and Ω are preferable.

The downside risk is measured by the maximum drawdown, which represents the maximum loss from a peak to a trough, before a new peak is attained. More precisely, let us consider the cumulative out-of-sample portfolio returns, which correspond to the values of wealth after t periods

$$W_t = W_{t-1}(1 + r_t^{out}) \quad (43)$$

with $t = t_0 + 1, \dots, T$ and $W_{t_0+1} = 1$, then the drawdowns are defined as

$$DD_t = -\frac{W_t - \max_{t_0+1 \leq s \leq t} W_s}{\max_{t_0+1 \leq s \leq t} W_s}.$$

The maximum drawdown, which corresponds to the largest loss achieved over the out-of-sample, is

$$MDD = \max_{t_0+1 \leq s \leq T} DD_t. \quad (44)$$

Two other statistics we also consider to provide further information about the left tail of the portfolio return distribution, namely the value-at-risk and the conditional value-at-risk. After sorting out-of-sample portfolio returns in ascending order, i.e. $r_{(t_0+1)}^{out} \leq r_{(t_0+2)}^{out} \leq \dots \leq r_{(T)}^{out}$, the value-at-risk at the confidence level α can be defined as

$$VaR_\alpha = -r_{(\lfloor \alpha(T-t_0-1) \rfloor)}^{out} \quad (45)$$

where $\alpha \in (0, 1)$. The conditional value-at-risk at the confidence level α can be estimated as

$$CVaR_\alpha = -\frac{\sum_{t=t_0+1}^T r_{(t)}^{out} \mathbf{1}_{\{r_{(t)}^{out} \leq -VaR_\alpha\}}}{\sum_{t=t_0+1}^T \mathbf{1}_{\{r_{(t)}^{out} \leq -VaR_\alpha\}}}. \quad (46)$$

In the experimental comparison we will set $\alpha = 10\%$.

The average of the diversification index introduced by Woerheide and Persson (1993) is used to measure the diversification level of optimal portfolios. It is given by

$$DI = \frac{1}{T-t_0} \sum_{t=t_0+1}^T \left(1 - \sum_{i=1}^n x_{i,t}^2 \right) \quad (47)$$

in which $x_{i,t}$ is the portfolio weight of asset i at time t . A greater value of DI implies a greater diversification level.

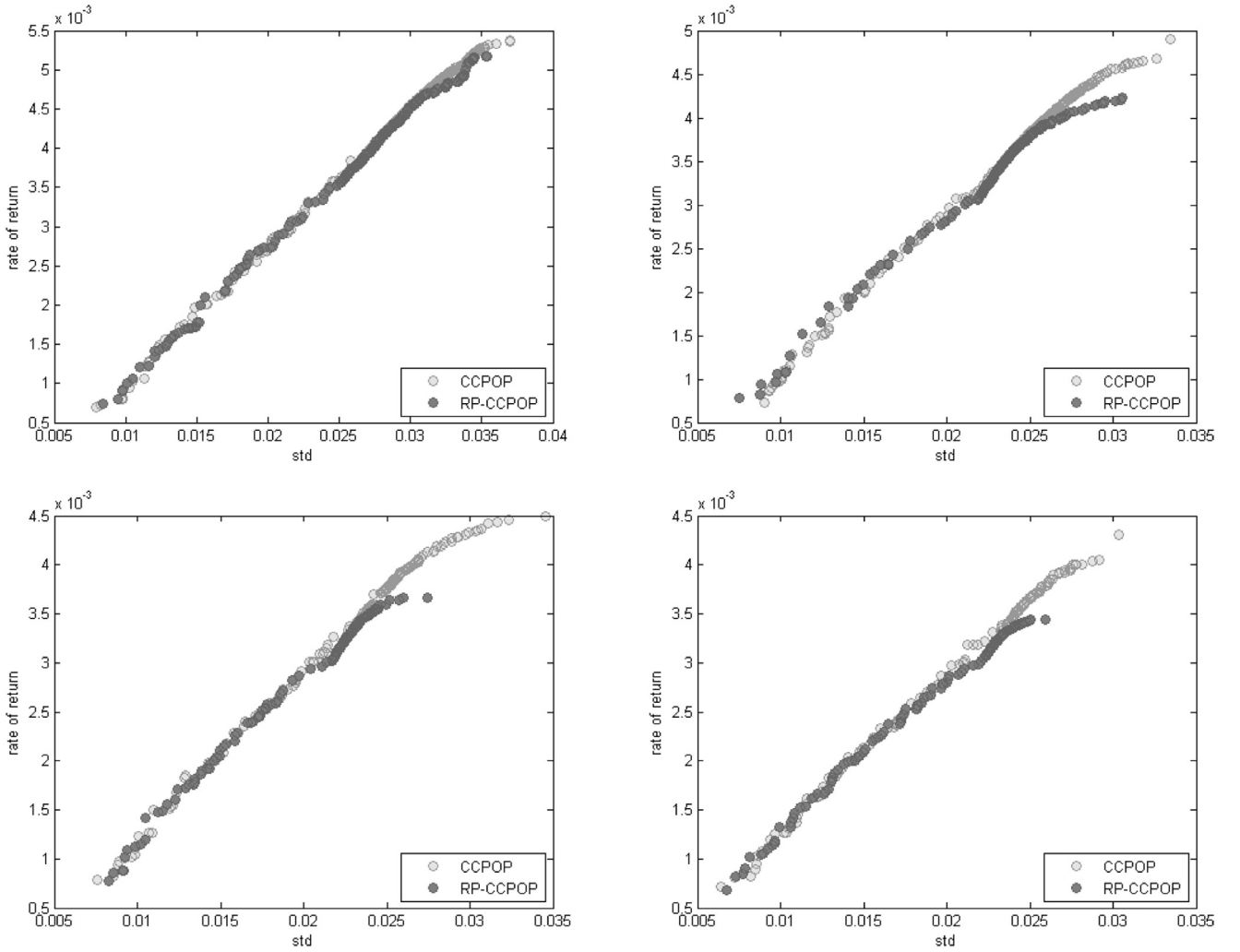


Fig. 1. Efficient frontiers for CCPOP and RP-CCPOP models related to the DowJones data set with K in $\{5, 10, 15, 20\}$. The top left chart refers to portfolios with cardinality $K = 5$, the top right displays portfolios with $K = 10$, the remaining plots regard solutions with $K = 15$ (on the left) and with $K = 20$ (on the right), respectively.

Table 10

Mean and standard deviation values of the CR performance metric (40) for the NS-MOPSO and HMOPSO* algorithms classified by data set and cardinality (K). The “rank-sum” column reports the corresponding results for the Wilcoxon rank-sum test. The symbol “ \uparrow ” means that NS-MOPSO has yielded better results than HMOPSO* at a 5% confidence level, “ \downarrow ” indicates that the algorithm in the column is statistically better than NS-MOPSO.

Data set	K	NS-MOPSO		HMOPSO*		
		Mean	std	Mean	std	Rank-sum
DowJones	5	0.7663	0.0389	0.9810	0.0087	\downarrow
	10	0.6806	0.0378	0.9784	0.0137	\downarrow
	15	0.4493	0.1959	0.9999	0.0001	\downarrow
	20	0.0000	0.0000	1.0000	0.0000	\downarrow
FF49Industries	5	0.8636	0.0375	0.8085	0.0730	\uparrow
	10	0.7799	0.0332	0.8981	0.0490	\downarrow
	20	0.2977	0.0979	0.9933	0.0045	\downarrow
	40	0.0000	0.0000	1.0000	0.0000	\downarrow
NASDAQ100	5	0.9217	0.0733	0.9798	0.0118	\downarrow
	10	0.8257	0.2372	0.9869	0.0141	\downarrow
	20	0.1772	0.2805	0.9986	0.0046	\downarrow
	40	0.0342	0.1158	1.0000	0.0000	\downarrow

Finally, to get an impression of the transaction costs involved, we calculate the average turnover over the out-of-sample period

as defined in DeMiguel et al. (2007)

$$turnover = \frac{1}{T - t_0} \sum_{t=t_0+1}^T \sum_{i=1}^n |x_{i,t+1} - x_{i,t}| \quad (48)$$

in which $x_{i,t}$ is the portfolio weight of asset i at time t and $x_{i,t+1}$ is the portfolio weight after rebalancing. The value of this statistic equals the average weekly amount of buy and sell transactions as a percentage of the portfolio value. Thus, a greater value of *turnover* indicates a more expensive investment strategy.

5. Computational analysis

This section is divided into two parts. First, we assess the capabilities of the variants of HMOPSO to generate well-distributed approximations to the Pareto front for the RP-CCPOP. Then, we compare the performance of the portfolios optimized in terms of the CCPOP with and without the risk parity constraint in order to highlight strengths and weaknesses of the proposed investment strategy.

5.1. Comparison of HMOPSO variants

The behavior of the aforementioned variants of HMOPSO is analyzed on the public data sets available in Bruni et al. (2016). Specifically, the following three groups of stocks are considered: Dow

Table 11

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the Dowjones data set with weekly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, "MDD" denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, "DI" is the mean of the diversification index for the portfolio weights and "turnover" provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	15	20	5	10	15	20
SR	0.4006	0.3538	-0.0293	0.0728	0.4150	1.0302	0.6449	0.3836
Ω	0.7931	0.9084	0.8497	0.8201	0.8647	0.8389	0.7967	0.8496
MDD	0.1362	0.1219	0.1486	0.1073	0.1345	0.0832	0.0549	0.0480
$VaR_{10\%}$	0.0291	0.0282	0.0297	0.0226	0.0304	0.0208	0.0169	0.0101
$CVaR_{10\%}$	0.0573	0.0539	0.0516	0.0443	0.0562	0.0349	0.0265	0.0189
DI	0.0280	0.0312	0.0312	0.0322	0.0289	0.0328	0.0344	0.0349
turnover	0.2544	0.5183	0.7180	0.6966	0.2577	0.5170	0.4668	0.2878

Table 12

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the FF49Industries data set with weekly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, "MDD" denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, "DI" is the mean of the diversification index for the portfolio weights and "turnover" provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	20	40	5	10	20	40
SR	2.0705	1.6621	1.2413	0.6384	2.0896	1.6291	1.0705	0.4317
Ω	1.6368	1.4159	1.0378	0.9862	1.5285	1.3914	0.9761	1.3565
MDD	0.0320	0.0492	0.0551	0.0563	0.0279	0.0527	0.0592	0.0359
$VaR_{10\%}$	0.0107	0.0158	0.0183	0.0173	0.0102	0.0194	0.0203	0.0079
$CVaR_{10\%}$	0.0183	0.0207	0.0234	0.0247	0.0183	0.0225	0.0257	0.0117
DI	0.0159	0.0179	0.0189	0.0198	0.0160	0.0180	0.0192	0.0195
turnover	0.1779	0.3693	0.6267	0.6674	0.1849	0.3226	0.5494	0.5573

Table 13

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the NASDAQ100 data set with weekly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, "MDD" denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, "DI" is the mean of the diversification index for the portfolio weights and "turnover" provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	20	40	5	10	20	40
SR	-0.3287	-0.2462	-0.3536	0.0271	-0.0503	-0.0603	0.0141	0.2433
Ω	0.5972	0.5675	0.5437	0.6848	0.6644	0.6618	0.6285	0.6872
MDD	0.2423	0.2405	0.2453	0.1802	0.1943	0.1958	0.1871	0.1547
$VaR_{10\%}$	0.0449	0.0427	0.0384	0.0410	0.0396	0.0383	0.0342	0.0382
$CVaR_{10\%}$	0.0798	0.0760	0.0739	0.0697	0.0760	0.0715	0.0648	0.0663
DI	0.0096	0.0108	0.0113	0.0119	0.0097	0.0109	0.0115	0.0119
turnover	0.3299	0.4330	0.5479	0.5079	0.3465	0.4053	0.5487	0.5728

Table 14

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the Dowjones data set with monthly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, "MDD" denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, "DI" is the mean of the diversification index for the portfolio weights and "turnover" provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	20	40	5	10	20	40
SR	0.4073	0.3253	0.1876	0.2919	0.2246	0.0584	-0.0158	0.1387
Ω	0.7936	0.8925	0.7842	0.8160	0.7408	0.8126	0.8528	0.7168
MDD	0.1372	0.1272	0.1338	0.1084	0.1387	0.1345	0.0725	0.0564
$VaR_{10\%}$	0.0304	0.0273	0.0314	0.0276	0.0287	0.0270	0.0150	0.0119
$CVaR_{10\%}$	0.0615	0.0542	0.0552	0.0504	0.0601	0.0531	0.0275	0.0220
DI	0.0280	0.0312	0.0319	0.0317	0.0287	0.0325	0.0343	0.0350
turnover	0.0734	0.1408	0.1618	0.1602	0.0744	0.1362	0.1071	0.0740

Table 15

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the FF49Industries data set with monthly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, " MDD " denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, " DI " is the mean of the diversification index for the portfolio weights and " $turnover$ " provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	20	40	5	10	20	40
SR	1.8358	1.4360	1.2331	0.3629	1.9389	1.6371	0.9724	1.1034
Ω	1.3864	1.2955	1.2148	0.9001	1.3294	1.3932	1.0200	1.3740
MDD	0.0375	0.0505	0.0607	0.0689	0.0376	0.0411	0.0599	0.0209
$VaR_{10\%}$	0.0139	0.0185	0.0198	0.0234	0.0127	0.0150	0.0212	0.0087
$CVaR_{10\%}$	0.0194	0.0216	0.0247	0.0273	0.0188	0.0209	0.0253	0.0115
DI	0.0159	0.0180	0.0189	0.0199	0.0159	0.0180	0.0192	0.0196
$turnover$	0.0665	0.0831	0.1607	0.1496	0.0814	0.1020	0.1402	0.1281

Table 16

Out-of-sample performance measures for the investment strategies related to CCPOP and RP-CCPOP models classified by cardinality (K) relative to the last 53 weeks of the NASDAQ100 data set with monthly rebalancing. "SR" is the annualized Sharpe ratio, " Ω " is the weekly Omega ratio, " MDD " denotes the maximum drawdown, " $VaR_{10\%}$ " and " $CVaR_{10\%}$ " represent respectively the value-at-risk and the conditional value-at-risk at the 10% significance level, " DI " is the mean of the diversification index for the portfolio weights and " $turnover$ " provides the mean turnover.

K	CCPOP				RP-CCPOP			
	5	10	20	40	5	10	20	40
SR	-0.3944	-0.2395	-0.1957	0.1566	-0.2690	-0.0820	0.0827	0.1607
Ω	0.6301	0.6681	0.6272	0.6121	0.7145	0.6560	0.7574	0.6656
MDD	0.3043	0.2532	0.2458	0.1749	0.2733	0.2198	0.1853	0.1599
$VaR_{10\%}$	0.0475	0.0410	0.0429	0.0389	0.0402	0.0439	0.0355	0.0402
$CVaR_{10\%}$	0.0850	0.0763	0.0743	0.0701	0.0853	0.0753	0.0671	0.0669
DI	0.0096	0.0108	0.0114	0.0119	0.0097	0.0109	0.0115	0.0119
$turnover$	0.1511	0.1146	0.1593	0.1361	0.1472	0.1216	0.1486	0.1454

Jones Industrial Average, Fama and French 49 Industry and NASDAQ 100. Columns 2–4 of Table 2 report the periods involved in the experiments with the corresponding number of weeks and assets. In order to analyze robustness and scaling of HMOPSO, and to identify the best choice among the proposed hybrid constraint-handling techniques, we have considered 12 instances of the RP-CCPOP by varying the number K of assets to include in the optimal portfolios according to the dimension n of the investible universe. The last column of Table 2 presents the cardinalities chosen.

The expected rates of return and covariance matrices of the stocks are calculated as sample estimates based on all the weekly data available in the corresponding data sets.

The algorithms have been implemented in a MATLAB environment and are run on a PC with Intel Core i7 3.0GHz with 8 GB RAM.

5.1.1. Parameters setting

Before performing the experiments, HMOPSO parameters have been tuned over the Dow Jones market following the procedure in Eiben and Smit (2011). The best values are listed in Table 3.

Moreover, the tolerance parameter τ for equality constraints relaxation has been fixed to 5×10^{-5} .

For each RP-CCPOP instance, we carried out 20 independent runs of the three versions of HMOPSO to collect statistics. The lower and upper bounds for the stocks weights are $l_i = 0.001$ and $u_i = 1$, $i = 1, \dots, n$, respectively. These are usual constraint values utilized in the literature (see, Anagnostopoulos and Mamanis, 2011; Meghwani and Thakur, 2017 and Lwin et al., 2014). Note that with this choice of bounds, conditions 2 and 3 of Proposition 1 are automatically verified by any combination of assets.

5.1.2. Experimental results and discussion

The first analysis focuses on the capability of HMOPSO to generate an appropriate number of feasible non-dominated solutions.

Table 4 reports the mean number of optimal solutions included in the external archive \mathcal{G}_{best} for each variant of the proposed algorithm, and for all instances of RP-CCPOP. Table 5 shows mean and standard deviation values of the Herfindahl index for risk contributions computed as in (36). Table 6 displays the Wilcoxon's rank-sum test for the h values of the optimal portfolios. We can notice that the variants of HMOPSO adopting the constrained domination principle and the self-adaptive tolerance to manage risk parity produce, on average, comparable number of optimal portfolios. Whereas, HMOPSO using the third constraint-handling mechanism generates archives with a size less than a quarter of those of its competitors. It should be noticed that, if we compare the h values of the optimal portfolios identified by the solvers with the corresponding attainable minimum values reported in the second column of Table 5, HMOPSO with self-adaptive penalty provides, on average, the best performance with the lowest standard deviation. The next table confirms the superiority of this variant with respect to that based on the constrained domination principle. The comparison with HMOPSO using self-adaptive tolerance is more complex. The implementation of a self-adaptive penalty outperforms the application of a self-adaptive tolerance in all instances of the Dow Jones data set. However, as the dimension of the market increases, the second mechanism attains better results.

Convergence and diversity are assessed in a single scale through the CR indicator (40). Table 7 displays the corresponding mean and standard deviation values. The results for this metric indicate the use of the constrained domination principle as the best choice in HMOPSO. The version based on the self-adaptive tolerance presents, in general, slightly worse results. Moreover, the Wilcoxon rank-sum test given in Table 8 corroborates these findings from the statistical point of view.

The mean total CPU times (in seconds) for the examined test problems are listed in Table 9. You can see that the proposed

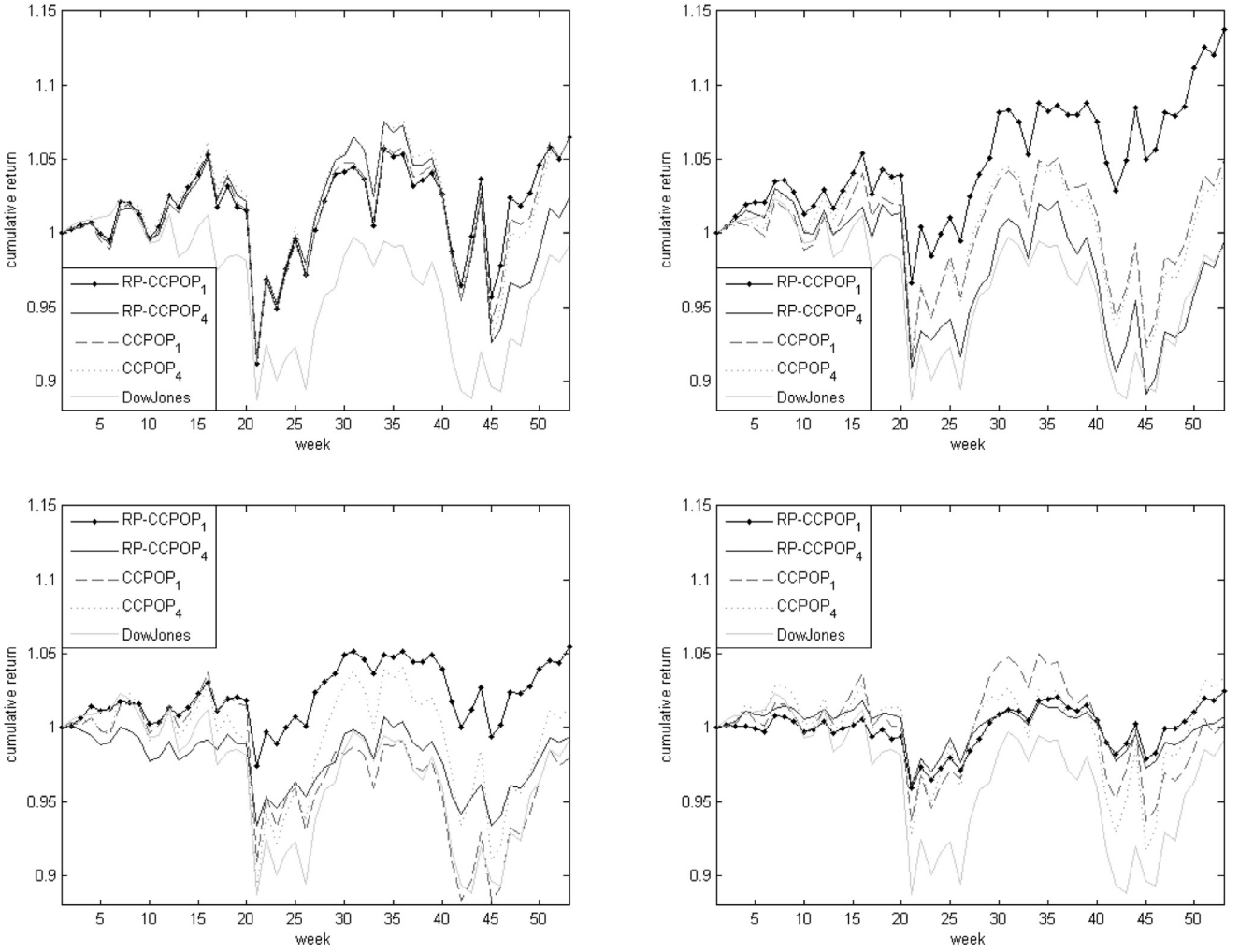


Fig. 2. Out-of-sample evolution of the wealth for the portfolios related to CCPOP and RP-CCPOP models over the last 53 weeks of the DowJones data set, taking into account weekly and monthly rebalancing strategies. The top left chart refers to portfolios with cardinality $K = 5$, the top right displays portfolios with $K = 10$, the remaining plots regard solutions with $K = 15$ (on the left) and with $K = 20$ (on the right), respectively.

algorithms present similar performance, with a slightly lower computational time for the HMOPSO with self-adaptive penalty.

Summing up, the developed algorithm exploiting the constrained domination principle is able to generate a large number of approximated RP portfolios with satisfactory risk-return tradeoff. The introduction into HMOPSO of a self-adaptive penalty provides the closest results to risk parity, but the sets of efficient portfolios it identifies are very small. In general, the integration of a self-adaptive tolerance to risk parity into HMOPSO guarantees the best balance between the number of optimal solutions produced and the quality of risk parity approximations.

5.2. Comparison between CCPOP and RP-CCPOP in terms of Pareto fronts

In this section, we compare the portfolios obtained by RP-CCPOP with those identified by the cardinality constrained mean-variance model. The comparison is based on the corresponding efficient frontiers.

We obtain a solver for mean-variance problems with cardinality, buy-in thresholds and budget constraints, denoted by HMOPSO*, by removing the mechanism related to the risk parity constraints from HMOPSO. We compare HMOPSO* with the non-dominated sorting multi-objective particle swarm optimization

(NS-MOPSO) algorithm proposed in Mishra et al. (2014), customized for CCPOP. Similar to HMOPSO*, it bases the exploration of the search space on the non-dominated sorting, uses a repair mechanism for the portfolio constraints, changes velocity direction with a fixed probability and adopts a clamping technique. The main differences are the following. First, NS-MOPSO implements the hybrid solution representation proposed by Streichert et al. (2004) for defining a portfolio. According to this approach, a binary vector specifies whether a particular asset participates in the portfolio, and a real-valued vector is used to compute the proportions of the budget invested in the assets. Second, it does not consider any mutation operator for the portfolio weights.

The structure of the experiments is the same as that described in Section 5.1, with the data sets and the cardinalities (K) listed in Table 2. The parameters used in HMOPSO* are those reported in Table 3. For the NS-MOPSO algorithm, we adopt the parameter setting recommended by Mishra et al. (2014), i.e., $w = 0.862$, $c_1 = c_2 = 2.05$, $v^{\min} = 0.06$ and $v^{\max} = 0.5$.

It is evident from the computational results given in Table 10, that the proposed simplified version of HMOPSO overperforms its competitor in almost all test problems in terms of contribution rate values. The Wilcoxon's rank-sum test values confirm these findings at a 5% significance level. In particular, we can note that the num-

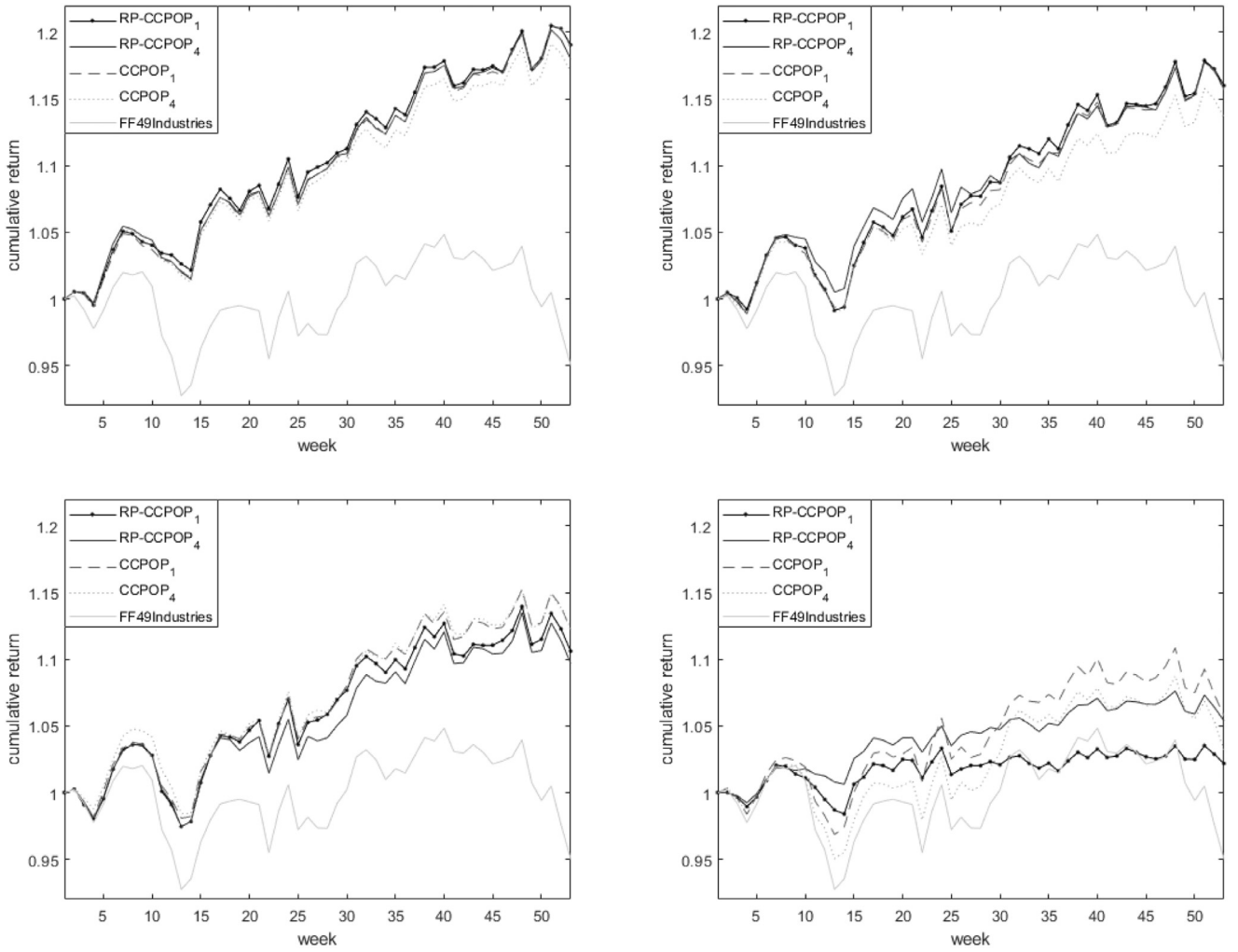


Fig. 3. Out-of-sample evolution of the wealth for the portfolios related to CCPOP and RP-CCPOP models over the last 53 weeks of the FF49Industries data set, taking into account weekly and monthly rebalancing strategies. The top left chart refers to portfolios with cardinality $K = 5$, the top right displays portfolios with $K = 10$, the remaining plots regard solutions with $K = 20$ (on the left) and with $K = 40$ (on the right), respectively.

ber of non-dominated solutions identified by HMOPSO* increases with respect to NS-MOSPO as the cardinality K increases.

Based on these results and the consideration made in the previous section, we will adopt the HMOPSO* for the CCPOP strategy while the variant of HMOPSO using a self-adaptive tolerance will be used for RP-CCPOP.

Fig. 1 displays the approximated Pareto fronts for the CCPOP solutions and those of the RP-CCPOP respectively, using the data from the Dow Jones market and varying the cardinality K in $\{5, 10, 15, 20\}$. The RP-CCPOP strategy is able to replicate the CCPOP risk-return profiles for low to medium risks. However, the risk parity condition limits the investment opportunities when allowable risks and cardinality increase. Therefore, for higher levels of risk, RP-CCPOP portfolios have lower profits or even are undefined. This finding is similar to every risk-based strategy.

5.3. Out-of-sample comparison of CCPOP and RP-CCPOP strategies

We now focus on the economic implications of the proposed portfolio optimization model. The magnitude of the potential gains an investor can realize, by using this strategy instead of the CCPOP model, is assessed out-of-sample.

5.3.1. Out-of-sample results and discussion

The three data sets listed in Table 2 are also used for the out-of-sample analysis. For each of them, we consider the last 53 observations as the test set and readjust optimal portfolio weights at regular calendar intervals, employing the so-called calendar based rebalancing procedure (Eakins and Stansell, 2007). In particular, two rolling window strategies are implemented: one combining 104 weeks to determine optimal portfolios and 1 week for the out-of-sample period, the other using 104 weeks in-sample and 4 weeks out-of-sample. For each in-sample window, the expected rates of return and covariance matrices of the stocks are first calculated as sample estimates, successively HMOPSO* is used to determine CCPOP solutions while the variant of HMOPSO using a self-adaptive tolerance is used for RP-CCPOP portfolios. From each set of non-dominated portfolios, we extract the one which presents the highest in-sample Sharpe ratio. The selected vectors of weights are then kept in the next out-of-sample period for evaluating the corresponding rates of return. The in-sample window is updated by removing the oldest data and including the most recent information, i.e., one week of rates of return in the case of weekly rebalancing and four weeks if the changes are needed monthly. This procedure is repeated until the last available week. Note that the same estimation window is used in both the procedures, since the

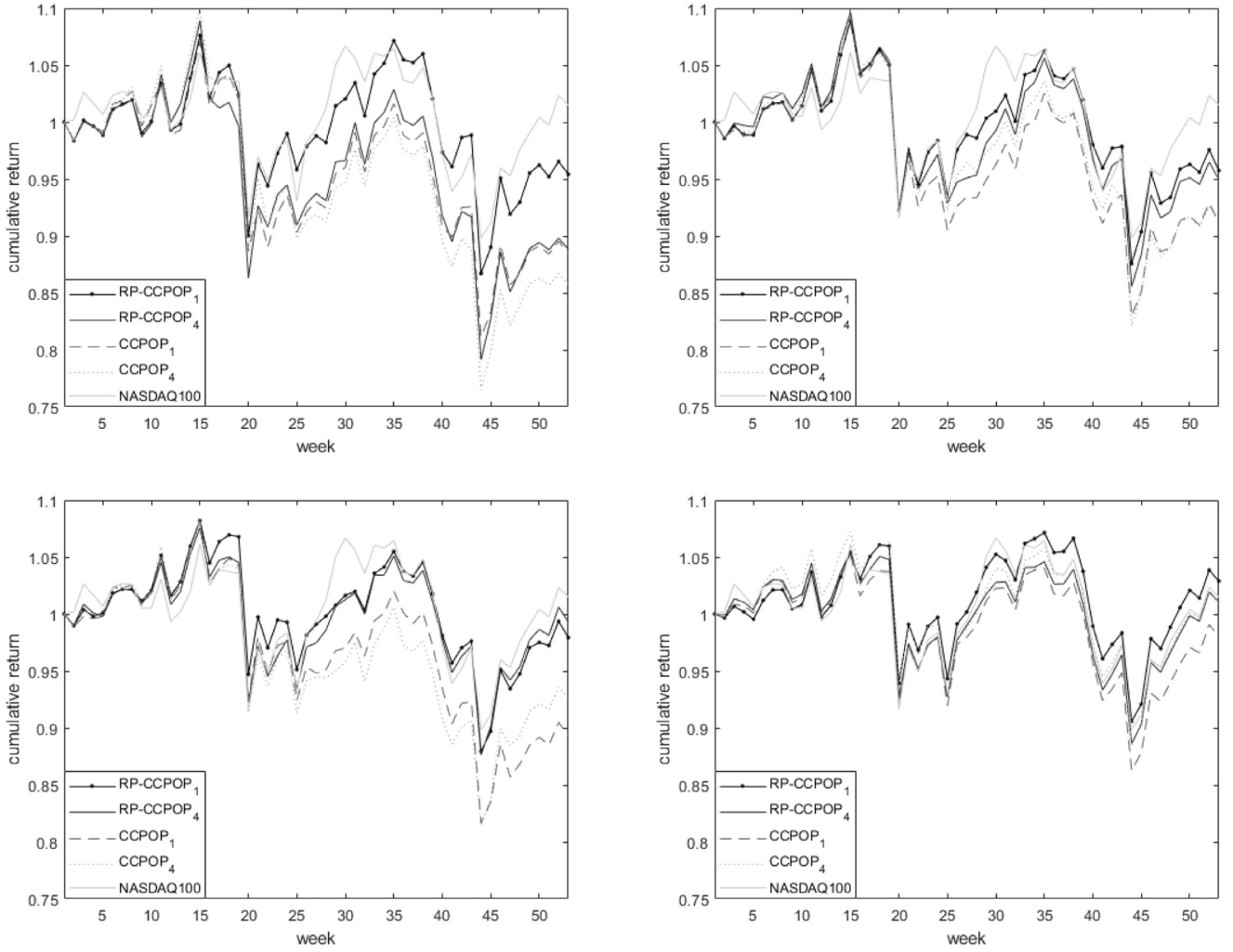


Fig. 4. Out-of-sample evolution of the wealth for the portfolios related to CCPOP and RP-CCPOP models over the last 53 weeks of the NASDAQ100 data set, taking into account weekly and monthly rebalancing strategies. The top left chart refers to portfolios with cardinality $K = 5$, the top right displays portfolios with $K = 10$, the remaining plots regard solutions with $K = 20$ (on the left) and with $K = 40$ (on the right), respectively.

impact of estimation error in determining optimal portfolios is not under investigation in the paper. In every case, the choice of a 2 years length for the estimation window seems reasonable, since there were no extraordinary economic changes during the period 2013–2016. On the other hand, we test the robustness of the solutions by varying the length of the rebalancing window.

Tables 11–13 report the performance statistics of the CCPOP and the RP-CCPOP models relative to the first rebalancing procedure for the DowJones, FF49Industries and NASDAQ100 data sets by varying the cardinality K . The most profitable strategies in terms of the annualized Sharpe ratio are those implementing RP-CCPOP. However, the quality of the results crucially depends on the choice of the portfolio cardinality. Indeed, for the first two data sets, portfolios with $K \leq 20$ attain the maximum values of SR (1.0302 and 2.0896). In the last experiment, a portfolio with 40 assets guarantees the best performance (0.2433). The CCPOP models involving portfolios with 15 to 40 assets attain the worst results. The two strategies present comparable values for the Omega ratio (42). With respect to the downside risk measures (44), (45) and (46), we can see that the proposed strategy produces the lowest extreme losses. Moreover, the level of diversification of the RP-CCPOP portfolios is slightly better than that attained by the CCPOP portfolios with a lower turnover.

Tables 14–16 list the results with monthly rebalancing. For the DowJones data set, the CCPOP portfolios are the most promising in terms of both the Sharpe ratio and the Omega ratio. In the second experiment, the SR criterion indicates the RP-CCPOP with $K = 5$ as the best choice, while Ω suggests the use of the CCPOP with $K = 5$. For the NASDAQ100, the RP-CCPOP models give the best results of the performance ratios. Overall, analyzing the behavior of the optimal solutions in terms of the tail risk, we note that RP-CCPOP portfolios improve their control on the losses with respect to the CCPOP portfolios as K increases. The proposed strategy shows a better level of diversification and lower turnover in all the experiments, analogously to the case of weekly rebalancing.

Figs. 2–4 report the evolution over the out-of-sample period of the wealth (43) for the RP-CCPOP portfolios with weekly and monthly rebalancing, denoted by RP-CCPOP₁ and RP-CCPOP₄ respectively, in comparison to the corresponding CCPOP counterparts, CCPOP₁ and CCPOP₄, and the related market index. We assume an initial wealth of 1\$ for all the portfolios. Overall, in the first two case studies, the analyzed strategies are more profitable than simply investing on the index. For the comparisons involving the data from NASDAQ 100, these strategies tend to replicate the benchmark. However, it can be noted that the number of constituents and the length of the rebalancing pe-

riod affect in a different measure their performance in the three experiments.

For the DowJones data set, the top left chart in Fig. 2 reveals that the strategies behave in a similar manner for small sized portfolios, irrespective to the frequency of the rebalancing. The next chart shows that CCPOP solutions do not change a lot their performance, as the cardinality increases. At the same time, RP-CCPOP₄ deteriorates, almost settling on the levels of the index, while the RP-CCPOP strategy with a weekly rebalancing, RP-CCPOP₁, guarantees the best performance, closing with a final rate of return of around 14%. RP-CCPOP₁ remains one of the most profitable choices, when continuing to increase the cardinality of optimal portfolios, even if it reduces the final wealth like all the other strategies. Moreover, the role of the risk parity constraint in managing extreme losses increases.

In Fig. 3 we can see that the four strategies behave in a very similar way and with high profits when they are applied to the FF49Industries data set for small to medium sized portfolios. The major differences appear in the bottom right chart, where the optimal portfolios with $K = 40$ are compared. In this case, we invest in almost all the index constituents. The RP-CCPOP₄ solutions show a slight improvement compared to the benchmark.

Fig. 4 points out the capability of the RP-CCPOP strategies to reduce the losses compared to the CCPOP strategies, especially with weekly rebalancing.

Taking into account all these findings, it could be argued that the best compromise between performance and risk control is attained by the RP-CCPOP strategy with a cardinality equal to 20%, 30% of the constituents of the market and with weekly rebalancing.

6. Conclusions

We have proposed an extension of the Markowitz mean-variance asset allocation framework, where cardinality, buy-in thresholds, budget constraints and risk parity conditions are handled at the same time. The resulting mixed-integer bi-objective portfolio optimization problem has been tackled by means of a novel version of the multi-objective particle swarm optimization algorithm, called HMOPSO. This uses a swap mutation operator to improve the exploration capabilities of the original algorithm, and handles the constraints by a hybrid technique that combines a repair mechanism with different instances of the domination principle.

The computational analysis on a set of real-world test problems confirms the effectiveness and the robustness of HMOPSO. In particular, the variant using a self-adaptive tolerance for relaxing the risk parity condition provides the best compromise between the quality of solutions produced and the number of optimal risk-return profiles identified on the approximated Pareto front. Additional investigations, concerning the standard cardinality constrained portfolio optimization, have highlighted the benefits of the novel optimization strategy for risk management purposes. Indeed, the combination of cardinality constraint and risk parity permits to simultaneously control costs and risk contributions. The resulting portfolios present losses which are consistently reduced. At the same time, profits are comparable, or even higher, with respect to the portfolios optimized in terms of the standard cardinality constrained mean-variance model over all the out-of-sample period.

In our future research, we plan to study the performance of other bio-inspired algorithms and constraint-handling techniques for the RP-CCPOP. At the same time, it will be interesting to study the influence of different risk measures on the structure of the investible portfolios, when risk parity is included as a constraint.

List of acronyms and abbreviations

The following acronyms and abbreviations are used in the text:

CCPOP	cardinality constrained portfolio optimization problem
RP-CCPOP	risk parity based cardinality constrained portfolio optimization problem
MOPSO	multi-objective particle swarm optimization
HMOPSO	proposed version of MOPSO using a hybrid constraint-handling technique
HMOPSO*	HMOPSO modification to solve cardinality constrained portfolio optimization problems
NS-MOPSO	non-dominated sorting multi-objective particle swarm optimization
h	Herfindahl index of risk contributions
HV	hypervolume
CR	contribution rate
SR	Sharpe ratio
Ω	Omega ratio
MDD	maximum drawdown
$VaR_{10\%}$	value-at-risk at the 10% significance level
$CVaR_{10\%}$	conditional value-at-risk at the 10% significance level
DI	average of the diversification index

References

- Anagnostopoulos, K.P., Mamanis, G., 2010. A portfolio optimization model with three objectives and discrete variables. *Comput. Oper. Res.* 37 (7), 1285–1297.
- Anagnostopoulos, K.P., Mamanis, G., 2011. The mean-variance cardinality constrained portfolio optimization problem: an experimental evaluation of five multiobjective evolutionary algorithms. *Expert Syst. Appl.* 38 (11), 14208–14217.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. *Math. Financ.* 9 (3), 203–228.
- Auger, A., Bader, J., Brockhoff, D., Zitzler, E., 2009. Theory of the hypervolume indicator: optimal μ -distributions and the choice of the reference point. In: *Proceedings of the Tenth ACM SIGEVO Workshop on Foundations of Genetic Algorithms*. ACM, pp. 87–102.
- Bai, X., Scheinberg, K., Tutuncu, R., 2016. Least-squares approach to risk parity in portfolio selection. *Quant. Financ.* 16 (3), 357–376.
- Beasley, J.E., 2013. Portfolio optimisation: models and solution approaches. In: *Theory Driven by Influential Applications*. INFORMS, pp. 201–221.
- Bertsimas, D., Shioda, R., 2009. Algorithm for cardinality-constrained quadratic optimization. *Comput. Optim. Appl.* 43 (1), 1–22.
- Braga, M.D., 2015. Risk parity versus other μ -free strategies: a comparison in a triple view. *Invest. Manag. Financ. Innov.* 12 (2), 277–289.
- Bruni, R., Cesarone, F., Scozzari, A., Tardella, F., 2016. Real-world datasets for portfolio selection and solutions of some stochastic dominance portfolio models. *Data Brief* 8, 858–862.
- Cao, Y., Smucker, B.J., Robinson, T.J., 2015. On using the hypervolume indicator to compare pareto fronts: applications to multi-criteria optimal experimental design. *J. Stat. Plan. Inference* 160, 60–74.
- de Carvalho, R.L., Lu, X., Moulin, P., 2012. Demystifying equity risk-based strategies: a simple alpha plus beta description. *J. Portf. Manag.* 38 (3), 56–70.
- de Carvalho, R.L., Lu, X., Moulin, P., 2014. An integrated risk-budgeting approach for multi-strategy equity portfolios. *J. Asset Manag.* 15 (1), 24–47.
- Chang, T.-J., Meade, N., Beasley, J.E., Sharaiha, Y.M., 2000. Heuristics for cardinality constrained portfolio optimisation. *Comput. Oper. Res.* 27 (13), 1271–1302.
- Chaves, D., Hsu, J., Li, F., Shakernia, O., 2011. Risk parity portfolio vs. other asset allocation heuristic portfolios. *J. Invest.* 20 (1), 108–118.
- Coello, C.A.C., Pulido, G.T., Lechuga, M.S., 2004. Handling multiple objectives with particle swarm optimization. *IEEE Trans. Evol. Comput.* 8 (3), 256–279.
- Crama, Y., Schyns, M., 2003. Simulated annealing for complex portfolio selection problems. *Eur. J. Oper. Res.* 150 (3), 546–571.
- Cura, T., 2009. Particle swarm optimization approach to portfolio optimization. *Nonlinear Anal. Real World Appl.* 10 (4), 2396–2406.
- Deb, K., 2000. An efficient constraint handling method for genetic algorithms. *Comput. Methods Appl. Mech. Eng.* 186 (2), 311–338.
- DeMiguel, V., Garlappi, L., Uppal, R., 2007. Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? *Rev. Financ. Stud.* 22 (5), 1915–1953.
- Eakins, S.G., Stansell, S., 2007. An examination of alternative portfolio rebalancing strategies applied to sector funds. *J. Asset Manag.* 8 (1), 1–8.

- Eberhart, R., Kennedy, J., 1995. A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, MHS'95. IEEE, pp. 39–43.
- Eiben, A.E., Smit, S.K., 2011. Parameter tuning for configuring and analyzing evolutionary algorithms. *Swarm Evol. Comput.* 1 (1), 19–31.
- Ertenlice, O., Kalayci, C.B., 2018. A survey of swarm intelligence for portfolio optimization: algorithms and applications. *Swarm Evol. Comput.* 39, 36–52.
- Feng, Y., Palomar, D.P., 2015. SCRIP: successive convex optimization methods for risk parity portfolio design. *IEEE Trans. Signal Process.* 63 (19), 5285–5300.
- Feng, Y., Palomar, D.P., 2016. Portfolio optimization with asset selection and risk parity control. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 6585–6589.
- Fonseca, C.M., Knowles, J.D., Thiele, L., Zitzler, E., 2005. A tutorial on the performance assessment of stochastic multiobjective optimizers. In: Proceedings of the Third International Conference on Evolutionary Multi-Criterion Optimization (EMO), 216, p. 240.
- Griveau-Billion, T., Richard, J.-C., Roncalli, T., 2013. A fast algorithm for computing high-dimensional risk parity portfolios. available at SSRN: <https://ssrn.com/abstract=2325255>.
- Guerard Jr, J.B., 2009. Handbook of portfolio construction: contemporary applications of Markowitz techniques. Springer Science & Business Media.
- Hochreiter, R., 2015. An evolutionary optimization approach to risk parity portfolio selection. In: Proceedings of the European Conference on the Applications of Evolutionary Computation. Springer, pp. 279–288.
- Ishibuchi, H., Doi, K., Nojima, Y., 2017. On the effect of normalization in MOEA/D for multi-objective and many-objective optimization. *Complex Intell. Syst.* 3 (4), 279–294.
- 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. *Quant. Financ.* 1 (5), 489–501.
- Kaya, H., Lee, W., 2012. Demystifying risk parity. available at SSRN: <https://ssrn.com/abstract=1987770>.
- Keating, C., Shadwick, W.F., 2002. A universal performance measure. *J. Perform. Meas.* 6 (3), 59–84.
- Kolm, P.N., Tütüncü, R., Fabozzi, F.J., 2014. 60 years of portfolio optimization: practical challenges and current trends. *Eur. J. Oper. Research* 234 (2), 356–371.
- Li, D., Sun, X., Wang, J., 2006. Optimal lot solution to cardinality constrained mean-variance formulation for portfolio selection. *Math. Financ.* 16 (1), 83–101.
- Li, M., Yang, S., Liu, X., 2015. A performance comparison indicator for pareto front approximations in many-objective optimization. In: Proceedings of the Annual Conference on Genetic and Evolutionary Computation. ACM, pp. 703–710.
- Liagkouras, K., Metaxiotis, K., 2015. Efficient portfolio construction with the use of multiobjective evolutionary algorithms: best practices and performance metrics. *Int. J. Inf. Technol. Decis. Mak.* 14 (3), 535–564.
- Liagkouras, K., Metaxiotis, K., 2017. Examining the effect of different configuration issues of the multiobjective evolutionary algorithms on the efficient frontier formulation for the constrained portfolio optimization problem. *J. Oper. Res. Soc.* 1–23.
- Liang, J.J., Qu, B.-Y., 2013. Large-scale portfolio optimization using multiobjective dynamic multi-swarm particle swarm optimizer. In: Proceedings of the IEEE Symposium on Swarm Intelligence (SIS). IEEE, pp. 1–6.
- Lizárraga, G.L., 2009. On the Evaluation of the Quality of Non-dominated Sets. Centro de Investigación en Matemáticas, AC (CIMAT) Ph. D. thesis.
- Lwin, K., Qu, R., Kendall, G., 2014. A learning-guided multi-objective evolutionary algorithm for constrained portfolio optimization. *Appl. Soft Comput.* 24, 757–772.
- Maillard, S., Roncalli, T., Teïletche, J., 2010. The properties of equally weighted risk contribution portfolios. *J. Portf. Manag.* 36 (4), 60–70.
- Markowitz, H., 1952. Portfolio selection. *J. Financ.* 7 (1), 77–91.
- Meghwani, S.S., Thakur, M., 2017. Multi-criteria algorithms for portfolio optimization under practical constraints. *Swarm Evol. Comput.* 37, 104–125.
- Mishra, S.K., Panda, G., Majhi, R., 2014. A comparative performance assessment of a set of multiobjective algorithms for constrained portfolio assets selection. *Swarm Evol. Comput.* 16, 38–51.
- Moral-Escudero, R., Ruiz-Torrubiano, R., Suárez, A., 2006. Selection of optimal investment portfolios with cardinality constraints. In: Proceedings of the IEEE Congress on Evolutionary Computation, CEC. IEEE, pp. 2382–2388.
- Mutunge, P., Haugland, D., 2018. Minimizing the tracking error of cardinality constrained portfolios. *Comput. Oper. Res.* 90, 33–41.
- Qian, E., 2006. On the financial interpretation of risk contribution: risk budgets do add up. *J. Invest. Manag.* 4 (4), 41–51.
- Rockafellar, R.T., Uryasev, S., et al., 2000. Optimization of conditional value-at-risk. *J. Risk* 2, 21–42.
- Roncalli, T., 2013. Introduction to Risk Parity and Budgeting. CRC Press.
- Roncalli, T., 2014. Introducing expected returns into risk parity portfolios: a new framework for asset allocation. available at SSRN: <https://ssrn.com/abstract=2321309>.
- Sharpe, W.F., 1966. Mutual fund performance. *J. Bus.* 39 (1), 119–138.
- Shaw, D.X., Liu, S., Kopman, L., 2008. Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. *Optim. Methods Softw.* 23 (3), 411–420.
- Siu, F., 2014. Risk parity strategies for equity portfolio management: can an asset-class strategy translate to equities? *J. Indexes* 18–25.
- Spinu, F., 2013. An algorithm for computing risk parity weights. available at SSRN: <https://ssrn.com/abstract=2297383>.
- Streichert, F., Ulmer, H., Zell, A., 2004. Evaluating a hybrid encoding and three crossover operators on the constrained portfolio selection problem. In: Proceedings of the Congress on Evolutionary Computation, CEC, 1. IEEE, pp. 932–939.
- Tseng, P., 2001. Convergence of a block coordinate descent method for nondifferentiable minimization. *J. Optim. Theory Appl.* 109 (3), 475–494.
- Vijayalakshmi Pai, G., Michel, T., 2012. Metaheuristic optimization of marginal risk constrained long-short portfolios. *J. Artif. Intell. Soft Comput. Res.* 2 (3), 259–274.
- Woerheide, W., Persson, D., 1993. An index of portfolio diversification. *Financ. Serv. Rev.* 2 (2), 73–85.
- Woldeesenbet, Y.G., Yen, G.G., Tessema, B.G., 2009. Constraint handling in multiobjective evolutionary optimization. *IEEE Trans. Evol. Comput.* 13 (3), 514–525.
- Woodside-Oriakhi, M., Lucas, C., Beasley, J.E., 2011. Heuristic algorithms for the cardinality constrained efficient frontier. *Eur. J. Oper. Res.* 213 (3), 538–550.
- Xu, F., Chen, W., 2006. Stochastic portfolio selection based on velocity limited particle swarm optimization. In: Proceedings of the Sixth World Congress on Intelligent Control and Automation, WCICA, 1. IEEE, pp. 3599–3603.
- Yaakob, S.B., Watada, J., 2010. A hybrid particle swarm optimization approach to mixed integer quadratic programming for portfolio selection problems. *Int. J. Simul. Syst. Sci. Technol.* 11 (5).
- Zhang, H., Rangaiah, G., 2012. An efficient constraint handling method with integrated differential evolution for numerical and engineering optimization. *Comput. Chem. Eng.* 37, 74–88.
- Zitzler, E., Brockhoff, D., Thiele, L., 2007. The hypervolume indicator revisited: on the design of pareto-compliant indicators via weighted integration. In: Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization. Springer, pp. 862–876.
- Zitzler, E., Thiele, L., 1998. Multiobjective optimization using evolutionary algorithms: a comparative case study. In: Proceedings of the International Conference on Parallel Problem Solving from Nature. Springer, pp. 292–301.