



# A constrained swarm optimization algorithm for large-scale long-run investments using Sharpe ratio-based performance measures

Massimiliano Kaucic<sup>1</sup> · Filippo Piccotto<sup>1</sup> · Gabriele Sbaiz<sup>1</sup>

Received: 5 December 2022 / Accepted: 12 October 2023  
© The Author(s) 2023

## Abstract

We study large-scale portfolio optimization problems in which the aim is to maximize a multi-moment performance measure extending the Sharpe ratio. More specifically, we consider the adjusted for skewness Sharpe ratio, which incorporates the third moment of the returns distribution, and the adjusted for skewness and kurtosis Sharpe ratio, which exploits in addition the fourth moment. Further, we account for two types of real-world trading constraints. On the one hand, we impose stock market restrictions through cardinality, buy-in thresholds, and budget constraints. On the other hand, a turnover threshold restricts the total allowed amount of trades in the rebalancing phases. To deal with these asset allocation models, we embed a novel hybrid constraint-handling procedure into an improved dynamic level-based learning swarm optimizer. A repair operator maps candidate solutions onto the set characterized by the first type of constraints. Then, an adaptive  $\ell_1$ -exact penalty function manages turnover violations. The focus of the paper is to highlight the importance of including higher-order moments in the performance measures for long-run investments, in particular when the market is turbulent. We carry out empirical tests on two worldwide sets of assets to illustrate the scalability and effectiveness of the proposed strategies, and to evaluate the performance of our investments compared to the strategy maximizing the Sharpe ratio.

**Keywords** Multi-moment portfolio management · Sharpe ratio · Large-scale optimization · Dynamic level-based learning swarm optimizer · Hybrid constraint-handling

---

✉ Massimiliano Kaucic  
massimiliano.kaucic@deams.units.it

Filippo Piccotto  
filippo.piccotto@phd.units.it

Gabriele Sbaiz  
gabriele.sbaiz@deams.units.it

<sup>1</sup> Department of Economics, Business, Mathematics and Statistics, University of Trieste, Trieste, Italy

**Mathematics Subject Classification** Primary 90C59 · Secondary 91G10 · 68W50 · 90C26

## 1 Introduction

In the portfolio selection problem, two phases are usually involved. The first one is devoted to the selection of promising assets, the second one focuses on the allocation of capital among them. In the modern portfolio theory, the so-called mean-variance analysis developed by Markowitz (1952) represents a milestone paper, and it has gained widespread acceptance as a practical tool for portfolio optimization among researchers and practitioners (Guerard 2009). In this model the mean return of a portfolio represents the profit measure, while the portfolio variance is the risk. Accordingly, 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 return. The set of optimal mean-variance tradeoffs in the risk-return space forms the efficient frontier. To guide the choice of an investor among these efficient portfolios, Sharpe (1994) has introduced a performance measure defined as the ratio between the excess return of an investment with respect to a risk-free asset and its standard deviation. Sharpe ratios with higher values correspond to more promising alternatives. In this manner, only the first two moments of the returns distribution are involved in the portfolio selection problem. However, the issue of whether higher moments should be considered to properly represent the investors' behaviour has been widely debated in literature and it is still open. The pioneering studies by Arditti (1967) and Samuelson (1970) have pointed out the importance of the third-order central moment of the returns distribution in the portfolio allocation process. In the same direction, Scott and Horvath (1980) have observed that a positive preference for skewness and a negative preference for kurtosis, known as prudence and temperance respectively (Kimball 1990), properly explain the behaviour of investors. Thus, to take advantage of the potential upside return represented by positive skewness and the possible benefit given by small kurtosis, in the mean-variance framework, two novel performance measures have been recently proposed in literature, namely the adjusted for skewness Sharpe ratio (Zakamuline and Koekebakker 2009) and the adjusted for skewness and kurtosis Sharpe ratio (Pézier and White 2008). The former multiplies the classical Sharpe index by a factor linked to the portfolio skewness, while the latter extends the Sharpe ratio by including a factor that incorporates both skewness and kurtosis.

These two Sharpe ratio-based performance measures have been used for stock performance evaluation (Nagy and Benedek 2021) but, to the best of our knowledge, they have not been applied as objective functions into the portfolio optimization process. Therefore, we propose portfolio selection strategies which maximize the aforementioned performance measures, considering four types of real-world constraints. A cardinality constraint is used to manage the portfolio size, buy-in threshold constraints ensure that all the available capital is invested. To characterize the investment profile, we introduce a set of bound constraints and a turnover threshold. In this manner, thanks to the former constraints, we avoid both the concentration of

money in a few assets and its splitting into too many assets. Further, due to the turnover bound, we limit the possibility of portfolio changes over time. The proposed asset allocation problem is analyzed from the perspective of an institutional investor who operates in equity markets with hundreds or thousands of constituents and selects a restricted pool of stocks to build up an active portfolio.

The introduction of cardinality constraints in the portfolio design leads to optimization problems for which finding optimal solutions becomes computationally challenging (Moral-Escudero et al. 2006). For this reason, in recent years, swarm optimization algorithms, inspired by the self-organizing interaction among agents, have become popular for this topic. In particular, the particle swarm optimization (PSO) algorithm has shown a good capability in solving small and mid-size portfolio allocation models (Cura 2009; Zhu et al. 2011; Kaucic et al. 2020; Corazza et al. 2021). This algorithm, mimicking the swarm behaviour of social animals such as bird flocking, gathers the information about good solutions through the swarm, and floats in the whole search space to find the global solution of the problem (Wang et al. 2018). However, PSO does not work efficiently in solving high-dimensional optimization problems, due to the so-called curse of dimensionality (Gilli and Schumann 2012; Oldewage et al. 2020). To overcome this issue, Yang et al. (2018) have developed a variant of the level-based learning swarm optimizer (LLSO), which exhibits superiority in achieving higher quality solutions with respect to other competitors in the literature for large scale optimization problems. The algorithm is based on the teaching paradigm where individuals are divided into levels according to their fitness and they are treated differently. The most performing candidates are stored in higher levels and guide the learning of the other particles in the swarm.

In this paper, we adopt a dynamic variant of the LLSO algorithm to solve our portfolio optimization problems, with a specific clamping and reversing procedure for the particles update rule, to improve the exploration efficiency. Moreover, since the LLSO is blind to the constraints, we equip it with a novel hybrid constraint-handling technique which works as follows. In order to deal with the cardinality constraint we use a projection operator, that selects the largest components of the candidate solutions and sets equal to zero the remaining ones. With this technique, we can relax the cardinality equality condition proposed in Kaucic and Piccotto (2022) as an inequality, assuming that the number of stocks included in the portfolio is lower than or equal to a fixed threshold. Furthermore, through this process, we transform the original mixed-integer optimization problem into a problem involving only real variables. Then, buy-in thresholds and budget constraints are handled using the repair operator proposed in Meghwani and Thakur (2017). Finally, to control the turnover constraint, a  $\ell_1$ -exact penalty function method is adopted, as in Corazza et al. (2021).

Summing up, the contribution of this work to the current literature is threefold. On the one hand, it is the first time that the adjusted for skewness Sharpe ratio and the adjusted for skewness and kurtosis Sharpe ratio measures are employed in the portfolio optimization problem. On the other hand, regarding the algorithmic novelties, we propose an improved variant of the LLSO equipped with a novel ad-hoc constraint-handling procedure which involves a repair operator as well as an  $\ell_1$ -exact penalty function strategy. Moreover, from a practical point of view, we study the

robustness of the proposed multi-moment strategies by comparing their profitability in a long-run setting, with almost 14 years of observations, involving also the recent phases of large market fluctuation due to the COVID-19 pandemic and the Ukrainian crisis.

The remainder of the paper is organized as follows. In the next section we review some literature related to our work. In Sect. 3 we introduce the investment framework and the objective functions involved. Section 4 presents the improved LLSO algorithm with the novel hybrid constraint-handling technique. Section 5 is devoted to the experimental analysis while the conclusions and future works are reported in Sect. 6.

## 2 Related works

In this section, we first survey the foremost contributions appeared in literature on the multi-moment formulations of the portfolio optimization problem. Since a complete review of the most recent population-based heuristics for large-scale optimization problem can be found in Omidvar et al. (2022a) and Omidvar et al. (2022b), we focus solely on the papers related to our study and concerning the PSO improvements.

### 2.1 Multi-moment portfolio optimization models

In order to highlight the critical effects of prudence and temperance on investment decisions, and to provide a more complete characterization of investor preferences, many authors have revised the mean-variance framework by incorporating the third and fourth moments when constructing a portfolio (see Jurczenko and Maillet 2006 and references therein). Lai (1991) has introduced a new term in the objective function for including skewness into the asset allocation problem. Konno et al. (1993) have developed a model in which they maximize the third-order moment given a threshold for portfolio expected return and for the variance. In addition, Liu et al. (2003) have employed also a transaction costs constraint. Lai et al. (2006) have proposed a multi-objective portfolio optimization problem which involves the first four central moments of the portfolio return distributions.

The direct optimization of the third and the fourth moment terms in the problem formulation has been shown to be computationally demanding, due to the difficulty in obtaining reliable estimators for the co-skewness and co-kurtosis matrices, as highlighted in Kim et al. (2014). In some recent contributions, Chaigneau and Eeckhoudt (2020) and Gao et al. (2022) have extended the mean-variance framework by considering alternative measures which incorporate prudence and temperance in risk exposure. Finally, several experimental investigations in the field of behavioural finance have been carried out in the last years, pointing out the importance of prudence, temperance, and higher-order preferences on the investors' behaviour (see, among others, Colasante and Riccetti 2020, 2021).

## 2.2 PSO enhancements for large-scale optimization

In literature, there are two major algorithmic approaches to solve the curse of dimensionality for PSO, namely decomposition-based and non decomposition-based approaches. The first type of procedures separate a high-dimensional problem into several small-dimensional instances using a divide-and-conquer strategy to reduce the dimensionality. In this direction, Van der Bergh and Engelbrecht (2004) and Li and Yao (2012) have proposed cooperative co-evolutionary particle swarm optimization algorithm, which randomly divides the decision variables into subgroups and then uses PSO to optimize each subgroup separately. The second type of algorithms directly optimizes all the variables at the same time, employing a learning mechanism to properly balance diversity and convergence. In this context, several learning strategies have been proposed recently. For instance, the competitive swarm optimizer (CSO, Cheng and Jin 2015a) compares two randomly chosen particles, and then the superior particle guides the update of the inferior one. Inspired by social animal behaviors, Cheng and Jin (2015b) have proposed the social learning particle swarm optimizer (SL-PSO) which first sorts particles by fitness and then worse individuals learn from the better ones. Following the teaching concept that teachers should treat students in accordance with their abilities, Yang et al. (2018) have developed the so-called level-based learning swarm optimizer (LLSO). The optimal compromise between exploration and exploitation of this learning technique guarantees more accurate solutions than the above cited CSO and SL-PSO. For this reason, several extensions of LLSO have been developed. The dynamic LLSO (Yang et al. 2018) dynamically adjusts the number of groups in which the particle swarm is divided based on the performance of the algorithm over time. Due to the oversensitivity of the standard LLSO to the parameter setting, Song et al. (2021) have proposed an adaptive variant in which the evolution state of the swarm is adjusted on the information given by the swarm aggregation indicator. Similarly, the reinforcement learning level-based particle swarm optimization algorithm by Wang et al. (2022) introduces a reinforcement learning strategy to control the number of levels and to improve the search efficiency.

## 3 Investment framework

In this study, we consider a frictionless market where short selling is not allowed and all investors act as price takers. The investable universe is represented by  $n$  risky assets. A portfolio is denoted by the vector of its assets weights  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ . In our dynamic setting, portfolio weights are periodically rebalanced, with an investment horizon of length  $h$ . We observe the market over a time window  $\mathcal{T} = \{0, 1, \dots, T\}$  and we adopt the following two-steps scheme for the investment strategy:

1. at time  $T$  the optimal portfolio composition is determined using a scenario-based approach;
2. the same portfolio composition is retained until time  $T + h$ , assuming that the stocks selected are still available at  $T + h$ .

The prices of the  $n$  risky assets are available for the time window  $\mathcal{T}$ . Then, we define the observed price of asset  $i$  at time  $t$ ,  $t \in \mathcal{T}$  and  $i = 1, \dots, n$  with  $p_{i,t}$ , and the realized rate of return at time  $t$ , with  $t \geq 1$ , as  $r_{i,t} = \frac{p_{i,t}}{p_{i,t-1}} - 1$ . Based on this information, we aim to identify the optimal allocation at time  $T$  that guarantees the best performance at time  $T + h$ .

Let now  $(\Omega, \mathcal{F}, P)$  be the probability space on which we assume the random variables are defined. We denote by:

- $R_i^{(h)}$  the random variable representing the rate of return of asset  $i$  at future time  $T + h$ , with expected value  $\mu_i$ ;
- $R_x^{(h)} = \sum_{i=1}^n x_i R_i^{(h)}$  the random variable that expresses the rate of return of portfolio  $\mathbf{x}$  at future time  $T + h$ .

We repeat this investment procedure over time, updating the observation window  $\mathcal{T}$  by eliminating the  $h$  oldest observations and adding the most recent ones. For simplicity of notation, we will indicate the random rate of return of portfolio  $\mathbf{x}$  at  $T + h$ ,  $R_x^{(h)}$ , by  $R_x$ .

### 3.1 Sharpe ratio-based performance measures

In this section, we will introduce the three performance measures that we will consider in our portfolio optimization problems.

#### 3.1.1 Sharpe ratio

In the asset allocation problem proposed by Markowitz (1952), portfolio risk is represented by the volatility, given by  $\sigma(R_x) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j}$ , where  $(C)_{ij} = c_{ij}$  is the covariance between stocks  $i$  and  $j$ , with  $i, j = 1, \dots, n$ . In this model, the portfolio choice is made solely with respect to the expected rate of return  $\mu$  of the portfolio  $\mathbf{x}$  and its risk, where a large portfolio volatility is perceived as damaging by the investors. As mentioned in the introduction, the Sharpe ratio is then defined as

$$SR(\mathbf{x}) = \frac{\mu - r_f}{\sigma(R_x)} \tag{1}$$

where  $r_f$  is a risk-free rate. This performance measure can be interpreted as the compensation earned by the investor per unit of risk. Thus, higher values of  $SR$  indicate more promising portfolios and are preferred by rational investors.

Even if it has an easy interpretation, the Sharpe ratio presents some pitfalls. The first is the choice of the risk-free rate used to define the excess rate of return. The

debate among scholars and practitioners is still open. For instance, Hitaj and Zambruno (2016) consider  $r_f = 0$  as a reasonable value. Similarly, Amédée-Manesme and Barthélémy (2022) set exogenously  $r_f = 2\%$ . Alternatively, Deguest et al. (2022) suggest the use of 1 month or 3 month maturity US Treasury Bills. Since these sovereign bonds exhibited values close to zero during a large part of the investment period analyzed in this paper, in the empirical part we follow Hitaj and Zambruno (2016) and set  $r_f = 0$ .

Another problem related to this indicator is its incoherence with preference relations in periods of market downturns, when the expected excess return of the portfolio is negative. In these cases, a Sharpe ratio-oriented agent could select portfolios with higher volatility. To overcome this issue, we consider only the first moment, that means we prefer, among two portfolios with negative expected excess rate of return, the one with the less expected loss.

### 3.1.2 Adjusted for skewness Sharpe ratio

As observed in the introduction, a possible drawback of the Sharpe ratio is that it uses only the first two moments of the portfolio returns distribution, and does not consider the potential upside return represented by positive skewness. For this reason, Zakamuline and Koekebakker (2009) have derived an adjustment to the Sharpe ratio for skewness, which depends on the investor’s utility function.

The proposed alternative performance measure is called adjusted for skewness Sharpe ratio (ASSR), and is given by

$$ASSR_b(\mathbf{x}) = SR(\mathbf{x}) \sqrt{1 + b \frac{S_3(R_{\mathbf{x}})}{3} SR(\mathbf{x})} \tag{2}$$

where  $b$  expresses the individual’s relative preference to the third moment of the returns distribution, and  $S_3(R_{\mathbf{x}})$  is the skewness, defined by  $E \left[ \left( \frac{R_{\mathbf{x}} - \mu}{\sigma(R_{\mathbf{x}})} \right)^3 \right]$ .

The properties of this performance measure have been investigated by Cheridito and Kromer (2013). In particular, we can note that the  $ASSR_b$  preserves the standard Sharpe ratio for zero skewness, while it is higher if the skewness is positive. However, in order to compute the  $ASSR_b$  one needs to determinate the value of  $b$ , which depends on the choice of the utility function. Thus, this performance measure is not unique for all investors, but it is rather an individual performance measure. In this work, we will set  $b = 1$ , meaning that we consider an investor with exponential utility, and we will indicate the above quoted performance measure by  $ASSR$ .

### 3.1.3 Adjusted for skewness and kurtosis Sharpe ratio

In order to account for both skewness and kurtosis, Pézier and White (2008) have proposed the so-called adjusted for skewness and kurtosis Sharpe ratio (AKSR), that is

$$AKSR(\mathbf{x}) = SR(\mathbf{x}) \left[ 1 + \frac{S_3(R_{\mathbf{x}})}{3!} SR(\mathbf{x}) - \left( \frac{K_4(R_{\mathbf{x}}) - 3}{4!} \right) SR(\mathbf{x})^2 \right] \tag{3}$$

where  $K_4(R_{\mathbf{x}})$  is the kurtosis, defined by  $E \left[ \left( \frac{R_{\mathbf{x}} - \mu}{\sigma(R_{\mathbf{x}})} \right)^4 \right]$ .

According to this performance measure, an investor prefers portfolios with higher skewness and dislikes portfolios with kurtosis values higher than 3, meaning that leptokurtic portfolio distributions are penalized in order to avoid extreme events.

### 3.2 Portfolio optimization model

Now, we introduce the family of constraints used to define the set of admissible portfolios.

- 1) *Budget constraint.* We require that all the available capital is invested. In terms of portfolio weights this translates into

$$\sum_{i=1}^n x_i = 1. \tag{4}$$

- 2) *Cardinality constraint.* We assume that the portfolio includes up to  $K$  stocks out of the  $n$  available, where  $K < n$  is a predefined number. To model the inclusion or the exclusion of the  $i$ -th asset in the portfolio, an auxiliary variable  $\delta_i$  is defined as follows

$$\delta_i = \begin{cases} 1, & \text{if asset } i \text{ is included} \\ 0, & \text{otherwise} \end{cases}$$

for  $i = 1, \dots, n$ .

The resulting vector of selected assets is  $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ . We can write the cardinality constraint as

$$\sum_{i=1}^n \delta_i \leq K. \tag{5}$$

- 3) *Box constraints.* To avoid extreme positions and foster diversification, we introduce a maximum and a minimum limit for the wealth allocation in the  $i$ -th stock included in the portfolio. Let  $l_i$  and  $u_i$  be respectively the lower and the upper bound for the weight of the  $i$ -th asset, with  $0 < l_i < u_i \leq 1$ , then we can write the box constraints as

$$\delta_i l_i \leq x_i \leq \delta_i u_i, \quad i = 1, \dots, n. \tag{6}$$

Note that if an asset is not included in the portfolio, no capital is invested on it.

- 4) *Turnover constraint.* In every rebalancing phase, the portfolio composition used in the previous investment window, denoted by  $\mathbf{x}_0$ , is updated. Let  $\mathbf{x}$  be the vector of



weights in the rebalanced portfolio and  $\tilde{\mathbf{x}}_0$  be the vector of re-normalized weights associated to  $\mathbf{x}_0$  (see Shen et al. 2014), which is calculated component-wise as

$$\tilde{x}_{0,i} = \frac{x_{0,i}(r_i + 1)}{\sum_{j=1}^n x_{0,j}(r_j + 1)}$$

and  $r_i, i = 1, \dots, n$  is the rate of return of the  $i$ -th stock in the portfolio at the moment of the rebalancing phase. Then, the portfolio turnover constraint is given by

$$\sum_{i=1}^n |x_i - \tilde{x}_{0,i}| \leq TR \tag{7}$$

where  $TR$  denotes the maximum turnover rate, which lies between 0 and 1. It can be noted that if  $TR = 0$  rebalancing is not allowed, and as  $TR$  increases more trades are allowed.

We indicate with  $\mathcal{X}$  the feasible set comprising the pairs  $(\delta, \mathbf{x})$  that satisfy (4), (5), (6) and (7).

Summing up, our asset allocation model can be written as

$$\begin{aligned} \max \quad & \Phi(\mathbf{x}) \\ \text{s.t.} \quad & (\delta, \mathbf{x}) \in \mathcal{X} \end{aligned}$$

where  $\Phi(\mathbf{x})$  is one of the three Sharpe ratio-based performance measures introduced in Sect. 3.1. As it is customary in the programming literature, we transform this maximization problem into the equivalent minimization instance

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & (\delta, \mathbf{x}) \in \mathcal{X} \end{aligned} \tag{8}$$

where  $f(\mathbf{x}) = -\Phi(\mathbf{x})$ .

### 3.2.1 Scenario generation technique

To solve the optimization problem (8), we use the following scenario-based generation technique. To estimate the values of the considered performance measures for a given portfolio  $\mathbf{x}$ , we need to simulate the distribution of the  $h$ -step ahead rate of return  $R_{\mathbf{x}}$ . To this end, we consider the historical rates of return of the  $n$  risky assets realized on the time window  $[0, T]$ . We assume that historical observations are good proxies for the future rates of return. Then, we define a scenario as the set of the joint realizations of the rates of return for the  $n$  assets in a given time period. Due to the good performance of block bootstrapping techniques to preserve correlations between time series (see, for instance, Guastaroba et al. 2009), we adopt the so-called stationary bootstrap (Politis and Romano 1994). This technique considers a random block size, that is a set of consecutive scenarios with variable length, in order to bring some robustness with respect to the standard block bootstrap, which uses fixed block size. The procedure works as follows. First, we select the optimal

average block size  $B^*$  by the procedure developed in Politis and White (2004). Then, we extract randomly from the observed data frame a block of length  $B^*$ . We repeat this exercise until the extracted sample reaches the desired size  $h$ , adjusting the last block length if the procedure exceeds the desired number of periods.

After having a bootstrap sample of  $h$  rates of return for each asset  $i$ , denoted as  $\widehat{R}_{i,t'}$ , with  $i = 1, \dots, n$  and  $t' = 1, \dots, h$ , we calculate the simulated  $h$ -step ahead rate of return of the  $i$ -th asset as  $\widehat{R}_i = \prod_{t'=1}^h (1 + \widehat{R}_{i,t'})$  and the simulated  $h$ -step ahead rate of return of portfolio  $\mathbf{x}$  as  $\widehat{R}_x = \sum_{i=1}^n x_i \widehat{R}_i$ . We repeat this procedure  $S$  times to have an estimate of the empirical distribution of  $R_x$ . With an abuse of notation, let  $\widehat{R}_x(s)$  be the  $s$ -th simulation of the  $h$ -step ahead rate of return of portfolio  $\mathbf{x}$ . We can then calculate all the quantities used to evaluate the Sharpe ratio-based performance measures. Indeed, the sample mean of the empirical distribution is given by

$$\widehat{\mu} = \frac{1}{S} \sum_{s=1}^S \widehat{R}_x(s)$$

and the sample standard deviation is

$$\widehat{\sigma} = \sqrt{\frac{1}{S-1} \sum_{s=1}^S (\widehat{R}_x(s) - \widehat{\mu})^2}.$$

Similarly, we estimate the skewness and kurtosis as follows

$$\widehat{S}_3(R_x) = \frac{1}{S-1} \sum_{s=1}^S \left( \frac{\widehat{R}_x(s) - \widehat{\mu}}{\widehat{\sigma}} \right)^3$$

and

$$\widehat{K}_4(R_x) = \frac{1}{S-1} \sum_{s=1}^S \left( \frac{\widehat{R}_x(s) - \widehat{\mu}}{\widehat{\sigma}} \right)^4.$$

## 4 Optimization algorithm

After introducing the LLSO paradigm, we present the dynamic LLSO with the proposed improvements as well as the hybrid constraint handling technique.

### 4.1 Level-based learning swarm optimizer

The LLSO algorithm, developed by Yang et al. (2018), evolves a swarm  $\mathcal{P}$  of  $NP$  candidate solutions using the so-called level-based population strategy. This process operates in accordance with the following two steps.

1. First, at each generation  $g$  the individuals in  $\mathcal{P}$  are sorted ascending based on their fitness and grouped into  $NL_g$  levels. Each level contains  $LP_g = \lfloor \frac{NP}{NL_g} \rfloor$  particles and in the last one, there are  $\lfloor \frac{NP}{NL_g} \rfloor + NP \% NL_g$  candidate solutions.<sup>1</sup> Better individuals belong to higher levels, and a higher level corresponds to a smaller level index. Therefore, we denote with  $L_1$  the best level and with  $L_{NL}$  the worst one.
2. Individuals belonging to the first level  $L_1$  are not updated and directly enter into the next generation, because they represent the most valuable information conveyed in the swarm at the current generation. On the contrary, the  $p$ -th particle in level  $L_l$ , denoted by  $\mathbf{x}^{l,p}(g)$ , where  $l = 3, \dots, NL_g$  and  $p = 1, \dots, LP_g$ , is allowed to learn from two particles  $\mathbf{x}^{l_1,p_1}(g)$  and  $\mathbf{x}^{l_2,p_2}(g)$ . These two individuals are randomly extracted from two different higher levels  $L_{l_1}$  and  $L_{l_2}$ , with  $l_1 < l_2$ , and  $p_1, p_2$  randomly chosen from  $\{1, \dots, LP_g\}$ . For  $l = 2$ , we sample two particles from  $L_1$  in such a way that  $\mathbf{x}^{l_1,p_1}(g)$  is better than  $\mathbf{x}^{l_2,p_2}(g)$  in terms of fitness function. Thus, the update rule for particle  $\mathbf{x}^{l,p}(g)$  is given by

$$\mathbf{v}^{l,p}(g + 1) = r_1 \mathbf{v}^{l,p}(g) + r_2 (\mathbf{x}^{l_1,p_1}(g) - \mathbf{x}^{l,p}(g)) + \psi r_3 (\mathbf{x}^{l_2,p_2}(g) - \mathbf{x}^{l,p}(g)) \tag{9}$$

$$\mathbf{x}^{l,p}(g + 1) = \mathbf{x}^{l,p}(g) + \mathbf{v}^{l,p}(g + 1) \tag{10}$$

for  $i = 1, \dots, n$ , where  $\mathbf{v}^{l,p}(g)$  is the so-called velocity of particle  $p$  in level  $L_l$  at generation  $g$ , and  $r_1, r_2, r_3$  are three real numbers randomly generated within  $[0, 1]$ . The initial velocities, at generation 0, are all set equal to the zero vector, that is  $\mathbf{v}^{l,p}(0) = \mathbf{0}$ . The parameter  $\psi \in [0, 1]$  controls the influence of the less performing exemplar  $\mathbf{x}^{l_2,p_2}(g)$  on  $\mathbf{v}^{l,p}(g + 1)$ .

The algorithm repeats these two steps until a maximum number of generations,  $MAX_{GEN}$ , is reached.

### 4.1.1 Dynamic LLSO with clamping and reversion

Following the suggestions in Yang et al. (2018), we adopt a dynamic setting for the number of levels  $NL_g$  by designing a pool  $S = \{l_1, \dots, l_s\}$  containing  $s$  different candidate integers. Then, at each generation  $g$ , the algorithm selects one of the elements of  $S$  based on their probabilities. At the end of the generation, the performance of the algorithm with the current level number is recorded, in order to update the probability of this level number for the next generation selection. To compute the probabilities of the elements of  $S$ , a record list  $Y_s = \{\gamma_1, \dots, \gamma_s\}$  is defined. Initially, each  $\gamma_i$  is set equal to 1. Then, the element  $\gamma_i$  corresponding to the level number  $l_i$  in the current generation is updated as follows

$$\gamma_i = \frac{|F - \tilde{F}|}{|F|} \tag{11}$$

<sup>1</sup> We denote by  $\lfloor z \rfloor$  the floor of  $z$  and by  $z \% y$  the rest of the division of  $z$  by  $y$ .

where  $F$  is the global best fitness of the last generation, and  $\tilde{F}$  is the global best fitness of the current generation. Then, the  $i$ -th element of the probability vector  $P_s = \{p_1, \dots, p_s\}$  is computed as

$$p_i = \frac{e^{7 \cdot \gamma_i}}{\sum_{j=1}^s e^{7 \cdot \gamma_j}} \tag{12}$$

with  $i = 1, \dots, s$ . At each generation, based on  $P_s$ , an integer from the pool  $S$  is selected as the level number following a roulette wheel scheme.

From now on, we will remove the dependency on the generation  $g$  if it will be clear from context.

Based on the results of a preliminary analysis, we introduce a ‘‘clamping and reversion’’ procedure to increase the exploration capability of the LLSO. The clamping mechanism is applied component-wise to the velocity vector in (9) as follows

$$v_i^{l,p} = \min\{\max\{v_i^{l,p}, v_i^{\min}\}, v_i^{\max}\} \tag{13}$$

where  $v_i^{\min}$  and  $v_i^{\max}$  are the minimum and the maximum velocity allowed for component  $i$ , with  $i = 1, \dots, n$ . In the experimental part, recalling (6), we set the maximum velocity as  $v_i^{\max} = u_i$  and the minimum velocity as  $v_i^{\min} = -v_i^{\max}$ . Moreover, when  $v_i^{l,p}$  and  $x_i^{l,p}$  are both negative, we reverse and scale  $v_i^{l,p}$  as follows:

$$v_i^{l,p} = -r_4 \cdot v_i^{l,p} \tag{14}$$

where  $r_4$  is a random number uniformly generated in  $[0, 1]$ .

### 4.2 Hybrid constraint-handling procedure

In the so-called construction phase, admissible portfolios have to satisfy cardinality, buy-in threshold, budget and turnover constraints. However, the LLSO update procedure is blind to these constraints. To overcome this issue, we equip the solver with a hybrid constraint-handling procedure.

First, in order to assure the cardinality requirement, for each candidate solution  $\mathbf{x}^p$ , the  $K$  largest components enter the corresponding portfolio, while zero weights are assigned to the remaining  $n - K$ . In this manner, we implicitly remove the binary variables from problem (8).

To guarantee the feasibility with respect to the bound constraints (6), we consider the following projection

$$x_i^p = \min\{\max\{x_i^p, l_i\}, u_i\} \tag{15}$$

where  $p = 1, \dots, NP$  and  $i \in I_K^p = \{i = 1, \dots, n : x_i^p > 0\}$ . Note that  $|I_K^p| \leq K$ . Then, we use the repair transformations developed in Meghwhani and Thakur (2017) to also satisfy the budget constraint (4). More precisely, for each  $p = 1, \dots, NP$ , assuming that  $l_i$  and  $u_i$  are such that  $\sum_{i \in I_K^p} l_i < 1$  and  $\sum_{i \in I_K^p} u_i > 1$ , we adjust the candidate solution  $\mathbf{x}^p$  component-wise:

$$x_i^p = \begin{cases} l_i + \frac{(x_i^p - l_i)}{\sum_{j \in I_k^p} (x_j^p - l_j)} \left( 1 - \sum_{j \in I_k^p} l_j \right), & \text{if } \sum_{j \in I_k^p} x_j^p > 1 \\ x_i^p, & \text{if } \sum_{j \in I_k^p} x_j^p = 1 \\ u_i - \frac{(u_i - x_i^p)}{\sum_{j \in I_k^p} (u_j - x_j^p)} \left( \sum_{j \in I_k^p} u_j - 1 \right), & \text{if } \sum_{j \in I_k^p} x_j^p < 1 \end{cases} \tag{16}$$

for all  $i \in I_k^p$ . As proved in Meghwani and Thakur (2017), solutions transformed through (16) fulfill at the same time budget and box constraints.

Finally, to handle the turnover constraint, we use the  $\ell_1$ -exact penalty function approach as in Corazza et al. (2021). In particular, we define the constraint violation of (7) as

$$CV = \max \left\{ \sum_{i=1}^n |x_i - \tilde{x}_{0,i}| - TR, 0 \right\}. \tag{17}$$

Then, we introduce the  $\ell_1$ -exact penalty function

$$F_{\ell_1}(\mathbf{x}, \varepsilon_0(g), \varepsilon_1(g)) = f(\mathbf{x}) + \frac{\varepsilon_1(g)}{\varepsilon_0(g)} CV \tag{18}$$

where  $\varepsilon_0$  and  $\varepsilon_1$  are two positive real numbers, defined adaptively at each generation  $g$ . Initially,  $\varepsilon_0(0) = 10^{-4}$  and  $\varepsilon_1(0) = 1$  in order to privilege feasible solutions. This parameters are then updated by checking the decrease of the objective function  $f(\mathbf{x})$  and the violation of the constraints. More precisely, every 5 iterations  $\varepsilon_0(g)$  is updated according to the rule

$$\varepsilon_0(g + 1) = \begin{cases} \min\{3 \cdot \varepsilon_0(g), 1\} & \text{if } f(\mathbf{x}(g)) \geq f(\mathbf{x}(g - 1)) \\ \max\{0.6 \cdot \varepsilon_0(g), 10^{-15}\} & \text{if } f(\mathbf{x}(g)) < 0.9 \cdot f(\mathbf{x}(g - 1)) \\ \varepsilon_0(g) & \text{otherwise} \end{cases} \tag{19}$$

while, every 10 iterations,  $\varepsilon_1(g)$  is updated following the scheme

$$\varepsilon_1(g + 1) = \begin{cases} \min\{2 \cdot \varepsilon_1(g), 10^4\} & \text{if } CV(g) > 0.95 \cdot CV(g - 1) \\ \max\{0.5 \cdot \varepsilon_1(g), 10^{-4}\} & \text{if } CV(g) < 0.9 \cdot CV(g - 1) \\ \varepsilon_1(g) & \text{otherwise.} \end{cases} \tag{20}$$

With this strategy, we privilege optimality of solutions possibly at the expenses of their feasibility, due to the fact that  $\varepsilon_0(g + 1)$  in (19) is increasing in  $F_{\ell_1}(\mathbf{x}, \varepsilon_0(g + 1), \varepsilon_1(g + 1))$  when the function value  $f(\mathbf{x}(g))$  increases. Moreover, to favour feasibility of solutions possibly at the expenses of their optimality, the

penalty parameter  $\varepsilon_1(g+1)$  in (20) increases when the relative constraint violation in the  $g$ -th generation increases with respect to the previous one.

Using the penalty approach, the constrained optimization problem (8) reduces to an unconstrained one, in which we minimize  $F_{\ell_1}$  and thus it can be solved by the proposed LLSO variant.

### 4.3 Initialisation strategy

Due to the complexity of the problem and to the fact that the search space grows exponentially with the dimension, the common strategies of seeking a search space coverage by initializing the particles uniformly throughout the space are inefficient (see van Zyl and Engelbrecht 2015). In particular, for our portfolio optimization problems, the presence of the turnover constraint exacerbates even more the initialization phase. To address this issue, we directly initialize the candidate solutions in a neighbourhood of  $\mathbf{x}_0$  as proposed by Kaucic et al. (2023). A brief description of the procedure follows. Let  $d_i^{\min}$  and  $d_i^{\max}$  be the minimum and the maximum allowed weight changes for  $\mathbf{x}_{0,i}$  respectively, with  $i = 1, \dots, n$ . Let  $D^p$  denote the total portfolio weight allowed to be re-allocated in  $\mathbf{x}_0$  for defining the  $p$ -th candidate solution  $\mathbf{x}^p(0)$  at generation 0, with  $p = 1, \dots, NP$ . Then, for each  $p$ ,

1. we randomly select  $D^p$  within  $\left[0, \frac{TR}{2}\right]$ ;
2. we select a subset  $J^-$  of  $K'$  assets from the assets with positive weight in  $\mathbf{x}_0$ , so that

$$x_j^p(0) = x_{0,j} - d_j, \text{ for } j \in J^-$$

where  $d_j$  is randomly sampled in  $\left[d_j^{\min}, d_j^{\max}\right]$  in such a way that  $\sum_{j \in J^-} d_j = D^p$ , and  $x_j^p(0) = 0$  or  $l_j \leq x_j^p(0) \leq u_j$ ;

3. we select a subset  $J^+$  of  $K''$  assets from the assets with zero weight in  $\mathbf{x}_0$ , with  $K'' \leq K'$ , so that

$$x_j^p(0) = x_{0,j} + d_j, \text{ for } j \in J^+$$

where  $d_j$  is randomly sampled in  $\left[d_j^{\min}, d_j^{\max}\right]$  in such a way that  $\sum_{j \in J^+} d_j = D^p$ , and  $l_j \leq x_j^p(0) \leq u_j$ ;

4. for  $j \in I \setminus (J^- \cup J^+)$ , we set  $x_j^p(0) = x_{0,j}$ .

The portfolios assembled using this scheme satisfy cardinality, buy-in threshold and turnover constraints. This initialization strategy encourages the swarm to focus on exploitation rather than exploration, and allows to identify promising solutions, even in problems with high dimension and small feasible regions.

**Table 1** Preliminary analysis on the normality assumption of the assets in the two data sets

Data set name	$ S_3  > 0.5$	$K_4 > 3$	J–B rejections
Pacific	19%	88%	46%
World	22%	89%	53%

The first column reports the name of the data set considered, while the corresponding percentage of assets with high skewness and large kurtosis is given in the second and third columns, respectively. The last column displays the rejections percentage of the null hypothesis of normality for the Jarque–Bera (J–B) test at 5% significance level

## 5 Computational analysis

### 5.1 Data description and portfolio parameters

We select two equity investment universes that differ for the number of constituents and for the geographic area, to highlight the scalability of the proposed portfolio strategies. The first data set, called Pacific, consists of 323 assets selected among the constituents of the MSCI Pacific Index at 28/07/2022. For the second data set, called shortly World, we consider 1229 assets listed in the MSCI World Index at 28/07/2022.

We obtain the daily rates of return and the market values for each asset from Bloomberg, covering the period 01/01/2008 to 28/07/2022 for a total of 3803 observations. For comparison purposes, we build up an auxiliary market value-weighted benchmark for each data set.

Table 1 reports a preliminary analysis concerning normality assumption for the time series of rates of return in the two data sets. In both cases, the number of assets exhibiting high skewness is around 20%, while those with large kurtosis is close to 90%. The Jarque-Bera (J-B) test rejects the null hypothesis of normality at the 5% significance level in almost half of both samples.

The goal of our computational analysis is twofold. On the one hand, we aim at pointing out the capabilities of the proposed dynamic LLSO algorithm in comparison to a state-of-the-art solver, which has been ad-hoc developed to solve cardinality-constrained portfolio optimization problems (see Corazza et al. 2021). On the other hand, we study the profitability of the proposed investment strategies focusing on the impact of both portfolio size and amount of trades.

To this end, we consider an investment plan with monthly portfolio rebalancing. The out-of-sample window is given by 126 months, covering the period from 02/01/2012 to 28/07/2022. For each month in the out-of-sample window, we generate 1000 scenarios of monthly rates of return for the assets in each data set by using the stationary bootstrap technique introduced above. The procedure employs an in-sample window of 1000 days, which is updated monthly by including the daily rates of return of the last month and by removing the information about the oldest month.

For the analysis, we set the buy-in thresholds  $l_i$  and  $u_i$  equal to 0.001 and 0.2 respectively, according to Kaucic and Piccotto (2022). We express the cardinality parameter  $K$  as a fraction  $K_\%$  of the number of assets in a given data set, that is

$K = \lfloor K_{\%} \cdot n \rfloor$ , and pick up  $K_{\%}$  in the set  $\{15\%, 30\%, 50\%\}$ . The turnover rate  $TR$  is in  $\{10\%, 20\%, 40\%\}$ . For each data set and each out-of-sample month, an instance of Problem (8) is then obtained by fixing one of the three Sharpe-based performance measures, a value of  $K_{\%}$  and  $TR$ , for a total of 27 alternative investment schemes.

Finally, the role of the trades is analyzed ex-post through the cost function introduced in Beraldi et al. (2021), with an initial wealth  $W_0 = 10,000,000\$$ .

## 5.2 Algorithm performance evaluation

In this subsection, we compare the performance of the proposed dynamic LLSO algorithm with respect to the PSO developed by Corazza et al. (2021). The latter has been showed to tackle efficiently non-smooth portfolio optimization problems with real-world constraints.

For each data set, the test suite consists of the 27 instances of Problem (8) previously introduced and specified at three dates randomly drawn from the out-of-sample window. For each sampled date and each portfolio optimization problem, we compute the initial portfolio without including the turnover constraint. The next out-of-sample month, we optimize the portfolio weights, accounting for the rebalancing constraint. It is worth noticing that we select three different dates for the evaluations in order to avoid the possible time dependence of the results.

The parameter settings for the considered algorithms follow the suggestions in the reference papers. More specifically, for the dynamic LLSO we set  $\psi = 0.4$  and the set of candidate level numbers  $S = \{4, 6, 8, 10, 20, 50\}$ , as in Yang et al. (2018). For the PSO variant by Corazza et al. (2021), we consider  $\omega_{min} = 0.4$ ,  $\omega_{max} = 0.9$ ,  $c_{1,min} = c_{2,min} = 0.5$ , and  $c_{1,max} = c_{2,max} = 2.5$ . To guarantee a fair comparison, we set for both solvers the maximum number of generations  $MAX_{GEN} = 1000$ , and the swarm size  $NP$  equal to 300 for the Pacific data set and 500 for the World one. To obtain more robust results, we run 30 times each test instance. The analysis have been implemented in MATLAB 2023a and carried out on a 3.3 GHz Intel Core i9-7900X workstation with 16 GB of RAM. To prove the efficiency of the proposed algorithm, comparisons are made in terms of run time, capability to identify optimal solutions which satisfy the constraints, and accuracy in solving the optimization problems.

Due to the negligible impact of the turnover rate levels on the results, in the following we show only the findings related to  $TR = 20\%$ . Tables 2 and 3 display the average computational time on 30 runs for the dynamic LLSO and the PSO. For both methods, we can observe that the results remain relatively stable when transitioning from one date to another, regardless of the cardinality threshold. However, the outcomes depend on the employed objective function and the number of involved decision variables. In our analysis the PSO exhibit a slightly lower average computational time. This difference can be attributed to the proposed hybrid constraint-handling technique.

Tables 4 and 5 show the average percentage of feasible solutions provided by the two solvers over the 30 runs at the final generation. Analyzing the results of the Pacific data set, it is evident that our LLSO is able to identify feasible solutions



**Table 2** Average computational time in seconds on 30 runs for the two compared algorithms, for the Pacific data set and the three ex-post dates, with  $TR = 20\%$  and increasing values of  $K_{\%}$

		Date 1			Date 2			Date 3		
<i>Pacific data set - LLSO</i>										
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR	
15%	97.98	162.23	233.07	103.16	164.91	237.26	94.77	148.86	214.20	
30%	100.62	165.15	236.22	103.96	168.93	239.78	91.26	151.94	217.56	
50%	100.83	165.83	235.90	104.14	168.95	240.60	91.52	152.00	217.91	
<i>Pacific data set - PSO</i>										
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR	
15%	80.51	125.99	193.27	82.882	127.37	195.91	79.041	119.62	181.70	
30%	80.54	126.89	195.39	81.827	128.40	197.57	78.70	119.83	182.73	
50%	80.55	127.40	195.57	81.604	127.90	196.56	79.08	120.50	182.88	

**Table 3** Average computational time in seconds on 30 runs for the two compared algorithms, for the World data set and the three ex-post dates, with  $TR = 20\%$  and increasing values of  $K_{\%}$

		Date 1			Date 2			Date 3		
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR	
<i>World data set - LLSO</i>										
15%	307.78	357.05	418.34	305.47	361.52	421.76	284.03	334.24	387.68	
30%	307.32	363.84	424.71	307.87	365.03	426.48	285.68	337.22	391.22	
50%	307.52	365.38	424.22	308.85	365.99	426.91	286.77	337.59	392.12	
<i>World data set - PSO</i>										
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR	
15%	281.87	314.34	357.79	278.45	305.44	345.67	278.90	310.53	348.13	
30%	281.39	315.74	358.73	273.24	304.48	345.08	278.90	310.73	355.27	
50%	282.65	314.61	356.33	273.07	304.66	345.48	277.97	309.80	347.59	

in almost all the cases. Conversely, the PSO algorithm struggles to properly handle constraints in the first test date, performing accurately in the other ones. Furthermore, the PSO is heavily influenced by the complexity of the objective function and the increase of portfolio size. As the number of decision variables increases, moving to the World case, the difference between the algorithms becomes even more pronounced. Specifically, for the cardinality threshold of 50%, the dynamic LLSO consistently finds optimal solutions, while the PSO fails to identify feasible portfolios in any of the 30 runs.

Moreover, we validate the capabilities of the proposed LLSO over the PSO in solving our test problems through a non-parametric statistical test. In particular, we focus on the average of the best values of the  $\ell_1$ -exact penalty function (18), and we conduct a Wilcoxon signed-rank test to determine if there is a significant difference in the distributions of the values obtained by the two algorithms (Derrac et al. 2011).

**Table 4** Percentage of optimal solutions that satisfy all the constraints, over the 30 runs for the two compared algorithms for the Pacific data set and for the three ex-post dates, with  $TR = 20\%$  and increasing values of  $K_{\%}$ 

$K_{\%}$	Date 1			Date 2			Date 3		
	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR
<i>Percentage of feasible solutions found (%) - Pacific data set - LLSO</i>									
15%	100	100	100	100	98.89	98.89	98.89	97.778	100
30%	100	100	100	100	98.89	100	98.89	100	98.89
50%	100	100	100	100	100	100	100	97.78	98.89
<i>Percentage of feasible solutions found (%) - Pacific data set - PSO</i>									
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR
15%	33.33	33.33	0	100	100	100	100	100	100
30%	33.33	33.33	0	98.89	96.67	96.67	96.67	93.33	86.67
50%	0	0	0	96.67	96.67	95.56	61.11	64.44	55.56

**Table 5** Percentage of optimal solutions that satisfy all the constraints, over the 30 runs for the two compared algorithms for the World data set and for the three ex-post dates, with  $TR = 20\%$  and increasing values of  $K_{\%}$ 

$K_{\%}$	Date 1			Date 2			Date 3		
	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR
<i>Percentage of feasible solutions found (%) - World data set - LLSO</i>									
15%	98.89	100	98.89	98.89	95.56	98.89	97.78	100	98.89
30%	100	100	100	97.78	96.67	93.33	98.89	96.67	96.67
50%	100	100	100	100	100	96.67	100	100	98.89
<i>Percentage of feasible solutions found (%) - World data set - PSO</i>									
$K_{\%}$	SR	ASSR	AKSR	SR	ASSR	AKSR	SR	ASSR	AKSR
15%	64.44	33.33	32.22	66.67	33.33	62.22	65.56	61.11	56.67
30%	33.33	33.33	0	33.33	33.33	32.22	0	0	0
50%	0	0	0	0	0	0	0	0	0

The results are given in Table 6, where  $R^+$  is the sum of ranks for the problems in which LLSO outperformed PSO, and  $R^-$  denotes the sum of the ranks for the opposite. Based on the  $p$ -values reported on the last column of this table, we conclude that our dynamic LLSO outperforms its competitor at the 5% significance level in all the case studies.

### 5.3 Long-run sensitivity analysis

#### 5.3.1 Ex-post performance metrics

In this subsection, we present the ex-post performance measures that we will use to assess the profitability of the proposed investment strategies. Let  $r_{p,t}^{out}$  be the

**Table 6** Wilcoxon signed-rank test results for the two data sets at different dates

Data set	Ex-post date	$R^+$	$R^-$	$p$ -value
Pacific	Date 1	378	0	$2.97 \cdot 10^{-6}$
Pacific	Date 2	378	0	$2.97 \cdot 10^{-6}$
Pacific	Date 3	273	105	0.0224
World	Date 1	357	21	$2.85 \cdot 10^{-5}$
World	Date 2	375	3	$4.16 \cdot 10^{-6}$
World	Date 3	376	2	$3.72 \cdot 10^{-6}$

The null hypothesis assumes that the distributions of the best values of the dynamic LLSO and the PSO by Corazza et al. (2021) are equal, meaning there is no significant difference between the solvers' performance. The alternative hypothesis suggests that the dynamic LLSO consistently outperforms the compared algorithm

ex-post portfolio rate of return at the month of the out-of-sample window. We compute the net wealth at time  $t$  as

$$W_t = (W_{t-1} - c_t) \left( 1 + r_{p,t}^{out} \right) \tag{21}$$

where  $c_t$  represents the transaction costs associated to the rebalancing at time  $t$ . Given the optimal portfolio at time  $t$ ,  $\mathbf{x}_t$ , and the re-normalized portfolio at time  $t - 1$ ,  $\tilde{\mathbf{x}}_{t-1}$ , we define the portfolio cost as the sum of the trading costs of each constituent, that is

$$c_t = \sum_{i=1}^n c(ts_{i,t})$$

where  $ts_{i,t} = W_{t-1} |x_{t,i} - \tilde{x}_{t-1,i}|$  is the so-called trade segment for asset  $i$  at time  $t$ , with  $i = 1, \dots, n$  and  $t = 1, \dots, 126$ , and  $c(\cdot)$  is given by

$$c(ts_{i,t}) = \begin{cases} 0 & ts_{i,t} = 0 \\ 40 & 0 < ts_{i,t} < 8000 \\ 0.05 \times ts_{i,t} & 8000 \leq ts_{i,t} < 50,000 \\ 0.04 \times ts_{i,t} & 50,000 \leq ts_{i,t} < 100,000 \\ 0.025 \times ts_{i,t} & 100,000 \leq ts_{i,t} < 200,000 \\ 400 & ts_{i,t} \geq 200,000. \end{cases} \tag{22}$$

According to (22), we divide the traded monetary amount into non-overlapping intervals and apply a different cost percentage on the interval in which the traded capital lies. This transaction costs structure reflects the main configurations proposed nowadays by financial brokers (Beraldi et al. 2021).

Next, we evaluate the attractiveness of the proposed asset allocation strategies through the so-called compound annual growth rate (shortly *CAGR*), which is calculated as

$$CAGR = \left( \frac{W_{T_{out}}}{W_0} \right)^{\frac{12}{T_{out}}} - 1 \quad (23)$$

where  $T_{out}$  denotes the length of the out-of-sample window, while  $W_0$  and  $W_{T_{out}}$  represent the initial wealth and the wealth at the end of the investment period, respectively.

We also compute the monthly ex-post average rate of return and the ex-post standard deviation

$$\mu^{out} = \frac{1}{T_{out}} \sum_{t=1}^{T_{out}} r_{p,t}^{out} \quad (24)$$

$$\sigma^{out} = \sqrt{\frac{1}{T_{out} - 1} \sum_{t=1}^{T_{out}} (r_{p,t}^{out} - \mu^{out})^2}. \quad (25)$$

Furthermore, to analyze more precisely the distribution of the ex-post returns, we calculate the ex-post skewness and kurtosis

$$S_3^{out} = \frac{1}{T_{out} - 1} \sum_{t=1}^{T_{out}} \left( \frac{r_{p,t}^{out} - \mu^{out}}{\sigma^{out}} \right)^3 \quad (26)$$

$$K_4^{out} = \frac{1}{T_{out} - 1} \sum_{t=1}^{T_{out}} \left( \frac{r_{p,t}^{out} - \mu^{out}}{\sigma^{out}} \right)^4. \quad (27)$$

To evaluate the capability of a strategy to avoid high losses, we introduce the draw-down measure, which is defined as follows (Chekhlov et al. 2005)

$$DD_t = \min \left\{ 0, \frac{W_t - W_{peak}}{W_{peak}} \right\} \quad t = 1, \dots, T_{out} \quad (28)$$

where  $W_{peak}$  is the maximum amount of wealth reached by the strategy until time  $t$ . Particularly, we focus on the mean, standard deviation, and maximum value of the drawdown measure over time.

Finally, we measure the effect of the costs on the available capital in the out-of-sample period as in Kaucic et al. (2020) by

$$\Lambda_{\%} = \frac{1}{T_{out}} \sum_{t=1}^{T_{out}} \frac{c_t}{W_{t-1}} \cdot 100. \quad (29)$$

### 5.3.2 Results for the two data sets

In this subsection we present the results of the ex-post analysis. We start by highlighting the relation between portfolio size, allowed trades in the rebalancing phases, and the ex-post performance. Tables 7 and 8 report the findings for the three proposed portfolio models employing the Pacific and the World data sets, respectively. In the last row, we also include the results for the associated market value-weighted benchmark. Overall, we can observe that the impact of costs, expressed as percentage of the total wealth, does not exceed 0.057% for the Pacific data set, whereas it attains 0.15% in the World case study. Therefore, the influence of transaction costs seems relatively marginal for the smaller data set and becomes more pronounced for the larger data set. Furthermore, it is worth noting that increasing the cardinality and turnover thresholds leads to an expected rise in costs. Instead, if we fix  $TR$  and vary  $K_{\%}$ , we can notice the quasi-linear growth of costs. On the contrary, with  $K_{\%}$  fixed, the cost increase is not proportional to the twofold increase of the turnover levels. This is coherent with the transaction cost structure defined in (22).

Regarding the Pacific data set, the ASSR and AKSR-based investment plans show similar results in the ex-post analysis for all the choices of  $K_{\%}$  and  $TR$ , with the AKSR-based strategy with  $K_{\%} = 15\%$  and  $TR = 10\%$  which outperforms the other tested alternatives.

For the World data set, maximizing the ASSR with  $K_{\%} = 15\%$  and  $TR = 10\%$  gives promising results in terms of ex-post returns, CAGR, and drawdown measures. However, this investment strategy lacks of robustness, as it underperforms significantly in all other parameter configurations compared to both the AKSR and SR strategies. As in the Pacific case, the AKSR-based model appears to be the most resilient and viable choice for the long-term investments.

Summing up, the optimal combination of the parameters is  $K_{\%} = 15\%$  as cardinality threshold and  $TR = 10\%$  as turnover rate. The comparison with the buy-and-hold strategy with the auxiliary benchmarks points out that the proposed models, in both the investable universes, outperform the benchmark in terms of profits, generating a higher ex-post net wealth on the long-run. In addition, the multi-moment strategies show less volatility and lower drawdown than the benchmark for the Pacific data set. However, for the World data set, the ASSR and AKSR-based investments present higher standard deviations and drawdowns.

### 5.3.3 Pre- and post-COVID analysis

Figure 1 reports the evolution of the net wealth for the best three Sharpe ratio-based portfolio strategies in each data set. The plot confirms the insights previously evidenced. All these investment strategies perform better than the auxiliary benchmarks. In the Pacific case study, we can observe that the SR investment plan is the one with the best performance until February 2020, date of the pandemic outbreak, at which it realizes a very large drawdown. After this period, only the AKSR strategy seems to be able to gain a great advantage of the market fluctuations in the last years. Similarly, in the World data set, the three models have a comparable net

**Table 7** Results of the long-run sensitivity analysis for the three Sharpe ratio-based optimization strategies with the Pacific data set

Pacific data set											
$K_{\%}$	$TR$	$\mu^{out}$	$\sigma^{out}$	$S_3^{out}$	$K_4^{out}$	Mean $DD_t$	std $DD_t$	Max $DD_t$	CAGR (%)	$\Lambda_{\%}$	
<i>Sharpe ratio</i>											
15%	10%	0.0155	0.0524	-0.3363	3.8968	-0.0401	0.0604	-0.3221	18.1213	0.0168	
15%	20%	0.0130	0.0510	-0.4092	3.9483	-0.0464	0.0630	-0.3312	14.6364	0.0246	
15%	40%	0.0129	0.0527	-0.2663	4.2228	-0.0469	0.0660	-0.3500	14.2350	0.0364	
30%	10%	0.0144	0.0511	-0.3447	4.0604	-0.0413	0.0615	-0.3402	16.5397	0.0258	
30%	20%	0.0126	0.0495	-0.6337	3.8347	-0.0456	0.0688	-0.3646	14.0217	0.0346	
30%	40%	0.0123	0.0509	-0.3258	3.6120	-0.0520	0.0697	-0.3407	13.4481	0.0482	
50%	10%	0.0141	0.0517	-0.3131	4.3767	-0.0401	0.0632	-0.3462	15.9818	0.0362	
50%	20%	0.0132	0.0508	-0.3919	3.6058	-0.0470	0.0655	-0.3339	14.5913	0.0448	
50%	40%	0.0140	0.0521	-0.3356	3.8593	-0.0498	0.0705	-0.3556	15.4638	0.0576	
<i>Adjusted for skewness Sharpe ratio</i>											
15%	10%	0.0151	0.0530	-0.2048	4.1423	-0.0413	0.0605	-0.3186	17.4500	0.0173	
15%	20%	0.0147	0.0526	-0.3414	3.6750	-0.0438	0.0602	-0.3200	16.8631	0.0227	
15%	40%	0.0135	0.0526	-0.2439	4.2715	-0.0451	0.0660	-0.3476	15.0479	0.0350	
30%	10%	0.0151	0.0523	-0.2658	3.8936	-0.0398	0.0596	-0.3309	17.5081	0.0245	
30%	20%	0.0136	0.0505	-0.5488	3.6586	-0.0450	0.0648	-0.3387	15.4300	0.0324	
30%	40%	0.0124	0.0513	-0.4045	3.6817	-0.0508	0.0697	-0.3527	13.4513	0.0466	
50%	10%	0.0150	0.0532	-0.1995	4.3461	-0.0397	0.0618	-0.3391	17.1096	0.0351	
50%	20%	0.0137	0.0522	-0.3742	3.7466	-0.0469	0.0673	-0.3503	15.2951	0.0422	
50%	40%	0.0140	0.0532	-0.2455	4.0119	-0.0515	0.0709	-0.3538	15.3767	0.0569	
<i>Adjusted for skewness and kurtosis Sharpe ratio</i>											
15%	10%	0.0159	0.0531	-0.1289	4.2466	-0.0391	0.0581	-0.3081	18.6239	0.0170	
15%	20%	0.0143	0.0526	-0.3342	3.6512	-0.0437	0.0608	-0.3180	16.3558	0.0231	
15%	40%	0.0141	0.0526	-0.3061	4.1253	-0.0440	0.0653	-0.3433	15.9187	0.0343	
30%	10%	0.0153	0.0525	-0.2646	4.0030	-0.0406	0.0607	-0.3351	17.6604	0.0248	
30%	20%	0.0139	0.0504	-0.5317	3.6388	-0.0442	0.0645	-0.3407	15.7871	0.0324	
30%	40%	0.0131	0.0520	-0.4019	3.6642	-0.0524	0.0721	-0.3573	14.4739	0.0449	
50%	10%	0.0151	0.0526	-0.2434	4.2324	-0.0403	0.0626	-0.3423	17.2423	0.0346	
50%	20%	0.0140	0.0517	-0.4260	3.7441	-0.0467	0.0677	-0.3502	15.6687	0.0416	
50%	40%	0.0148	0.0535	-0.2355	3.9622	-0.0508	0.0705	-0.3521	16.5384	0.0544	
<i>Benchmark</i>											
-	-	0.0074	0.0541	-0.1912	8.7070	-0.0794	0.0817	-0.3261	7.3157	-	

In the first two columns we report the value of the fraction  $K_{\%}$  of assets making up the portfolio and the value of the turnover rate,  $TR$ , respectively. The other columns show the results of the ex-post metrics presented in Sect. 5.3.1

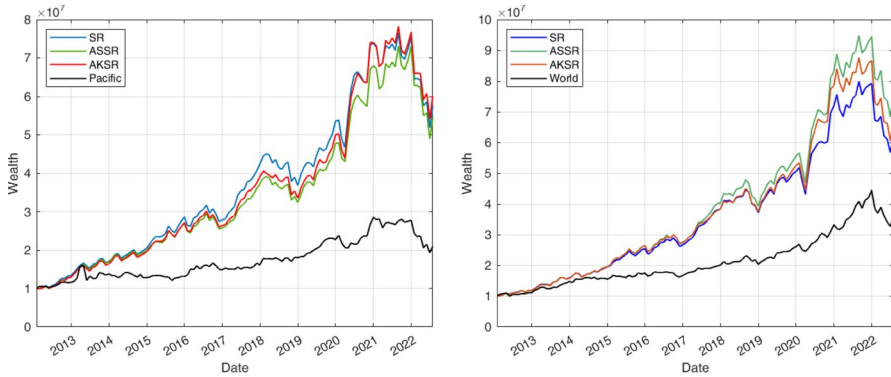
wealth evolution until February 2020. Then, the two multi-moment strategies perform better than the SR one, with the ASSR being the most profitable.

In Table 9 we highlight the different behaviours of the auxiliary markets and the proposed strategies before the above quoted date, namely pre-COVID period,

**Table 8** Results of the long-run sensitivity analysis for the three Sharpe ratio-based optimization strategies with the World data set. In the first two columns we report the value of the fraction  $K_{\%}$  of assets making up the portfolio and the value of the turnover rate,  $TR$ , respectively. The other columns show the results of the ex-post metrics presented in Sect. 5.3.1

World data set										
$K_{\%}$	$TR$	$\mu^{out}$	$\sigma^{out}$	$S_3^{out}$	$K_4^{out}$	Mean $DD_t$	std $DD_t$	Max $DD_t$	CAGR (%)	$\Lambda_{\%}$
<i>Sharpe ratio</i>										
15%	10%	0.0161	0.0466	-0.2773	4.9703	-0.0291	0.0551	-0.2912	18.9178	0.0425
15%	20%	0.0140	0.0430	-0.7310	4.6233	-0.0285	0.0511	-0.2584	16.1381	0.0553
15%	40%	0.0148	0.0438	-0.3794	4.9213	-0.0291	0.0516	-0.2557	16.8147	0.0751
30%	10%	0.0150	0.0451	-0.4116	5.0116	-0.0285	0.0564	-0.3093	17.0248	0.0762
30%	20%	0.0142	0.0442	-0.5375	5.0115	-0.0280	0.0529	-0.2722	15.8571	0.0870
30%	40%	0.0153	0.0446	-0.2956	4.4928	-0.0264	0.0474	-0.2501	17.2135	0.1003
50%	10%	0.0156	0.0463	-0.2359	5.5967	-0.0272	0.0575	-0.3195	17.1385	0.1194
50%	20%	0.0155	0.0453	-0.1842	5.2368	-0.0258	0.0505	-0.2742	17.0534	0.1247
50%	40%	0.0145	0.0449	-0.1576	5.0796	-0.0281	0.0505	-0.2673	15.4072	0.1454
<i>Adjusted for skewness Sharpe ratio</i>										
15%	10%	0.0177	0.0497	-0.1205	4.7322	-0.0290	0.0534	-0.2791	21.0693	0.0402
15%	20%	0.0138	0.0439	-0.4892	4.3779	-0.0306	0.0509	-0.2521	15.7652	0.0556
15%	40%	0.0149	0.0460	-0.1817	4.9571	-0.0302	0.0524	-0.2579	16.9092	0.0712
30%	10%	0.0156	0.0467	-0.4320	4.8426	-0.0297	0.0575	-0.3079	17.7964	0.0746
30%	20%	0.0142	0.0452	-0.4714	4.7642	-0.0294	0.0539	-0.2745	15.8234	0.0828
30%	40%	0.0152	0.0453	-0.2490	4.5258	-0.0264	0.0468	-0.2472	16.9971	0.0959
50%	10%	0.0154	0.0469	-0.1386	5.3080	-0.0279	0.0566	-0.3129	16.7302	0.1260
50%	20%	0.0152	0.0467	-0.0542	5.3984	-0.0267	0.0524	-0.2905	16.5561	0.1274
50%	40%	0.0138	0.0464	-0.0221	5.2357	-0.0290	0.0513	-0.2788	14.3094	0.1490
<i>Adjusted for skewness and kurtosis Sharpe ratio</i>										
15%	10%	0.0167	0.0497	-0.2142	4.8412	-0.0305	0.0575	-0.3090	19.6955	0.0406
15%	20%	0.0146	0.0455	-0.5668	4.9556	-0.0318	0.0556	-0.2766	16.8653	0.0525
15%	40%	0.0151	0.0472	-0.1642	5.1791	-0.0316	0.0539	-0.2549	17.1044	0.0726
30%	10%	0.0159	0.0465	-0.3997	4.8263	-0.0297	0.0570	-0.3009	18.2756	0.0750
30%	20%	0.0154	0.0471	-0.2610	5.0832	-0.0290	0.0537	-0.2740	17.4193	0.0804
30%	40%	0.0160	0.0474	-0.0953	4.8810	-0.0269	0.0490	-0.2624	18.0243	0.0957
50%	10%	0.0151	0.0462	-0.2559	5.2358	-0.0276	0.0570	-0.3136	16.5143	0.1228
50%	20%	0.0154	0.0473	-0.0451	5.1902	-0.0276	0.0549	-0.3064	16.7104	0.1298
50%	40%	0.0149	0.0475	-0.0182	5.3073	-0.0287	0.0536	-0.2936	15.8154	0.1449
<i>Benchmark</i>										
-	-	0.0108	0.0381	-0.6071	3.8057	-0.0273	-0.2626	0.0541	12.8355	-

and after that date, namely post-COVID period. For both data sets, the distributions of the rates of return in the pre-COVID period have a negative skewness while post-COVID rates of return show symmetric distributions and positive skewness. Conversely, regarding the benchmarks, the distribution of the rates of



**Fig. 1** Evolution of the net wealth for the best three Sharpe ratio-based portfolio strategies ( $K_{\%} = 15\%$  and  $TR = 10\%$ ) in comparison to the market value weighted benchmark using the assets in the two data sets

**Table 9** Out-of-sample statistics for the distribution of the ex-post rates of return of the three compared strategies with  $K_{\%} = 15\%$  and  $TR = 10\%$  and the benchmark for the two data sets

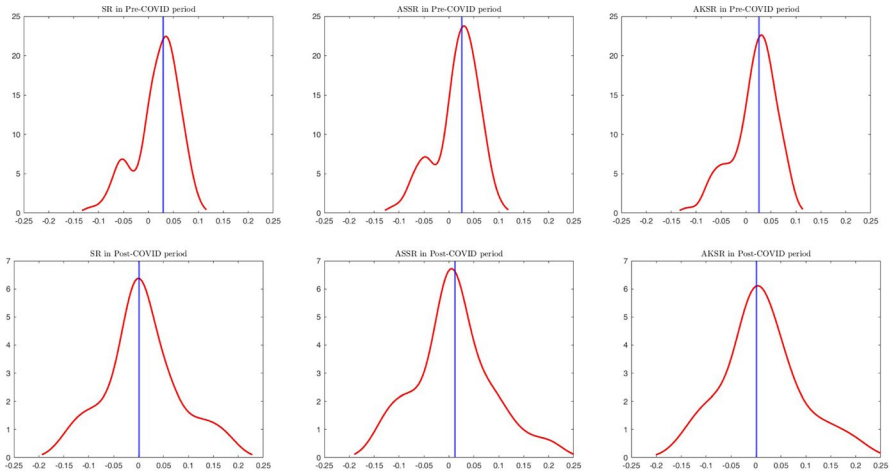
Strategy	Pre-COVID				Post-COVID			
	$\mu^{out}$	$\sigma^{out}$	$S_3^{out}$	$K_4^{out}$	$\mu^{out}$	$\sigma^{out}$	$S_3^{out}$	$K_4^{out}$
<i>Pacific data set</i>								
SR	0.0177	0.0439	-0.8178	3.2845	0.0083	0.0749	0.2446	3.3132
ASSR	0.0165	0.0437	-0.7641	3.1747	0.0103	0.0773	0.3176	3.4553
AKSR	0.0170	0.0435	-0.7728	3.2698	0.0122	0.0780	0.3933	3.4518
Benchmark	0.0095	0.0526	-0.0977	11.7410	0.0004	0.0602	-0.3545	3.5086
<i>World data set</i>								
SR	0.0174	0.0367	-0.7443	3.5505	0.0116	0.0711	0.1419	3.6941
ASSR	0.0183	0.0388	-0.6842	3.4463	0.0155	0.0766	0.2310	3.3632
AKSR	0.0177	0.0380	-0.6780	3.3421	0.0134	0.0780	0.1222	3.2992
Benchmark	0.0100	0.0324	-0.4454	3.6395	0.0135	0.0542	-0.7804	3.0798

Results refer to the period before and after 28 February 2020, the watershed date for the COVID-19 pandemic outburst

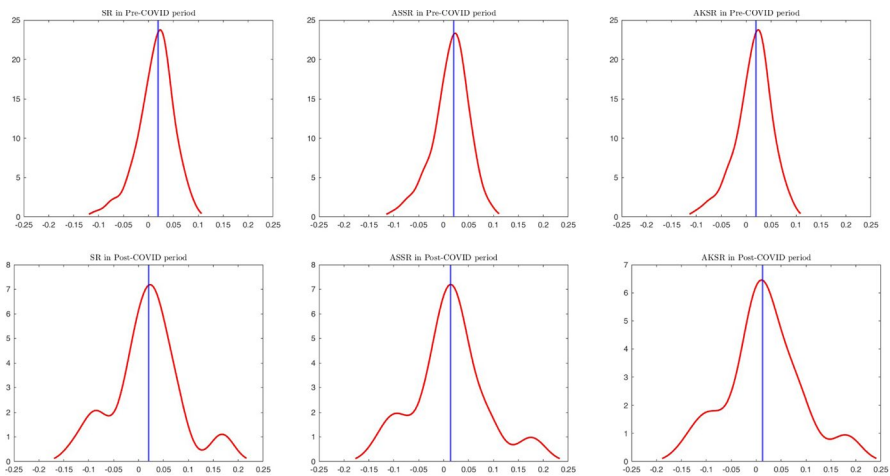
return presents a lower mean with negative skewness in both market phases. We summarize the results in Figs. 2 and 3 in which, in both data sets, we can detect a positive value for the ex-post mean of the rates of return with a reduced dispersion in the pre-COVID period, while in the post-COVID epoch we can discover a lower ex-post mean of the rates of return with more dispersion and fat-tails.

These findings substantiate that extending the Sharpe ratio model by incorporating higher-order moments can yield financial performance benefits, particularly during periods characterized by market instability.





**Fig. 2** Density functions of the ex-post rates of return of the three strategies proposed in Sect. 3.1 for the Pacific data set in the pre-COVID period (in the first row) and post-COVID epoch (in the second row). The vertical blue line represents the median of the distribution (colour figure online)



**Fig. 3** Density functions of the ex-post rates of return of the three strategies proposed in Sect. 3.1 for the World data set in the pre-COVID period (in the first row) and post-COVID epoch (in the second row). The vertical blue line represents the median of the distribution (colour figure online)

## 6 Conclusions and future works

In this study, we have proposed a comparison of three long-run investment strategies based on Sharpe ratio type performance measures on large-scale global market indices. In particular, we have considered the standard Sharpe ratio and two extensions which involve higher-order moments of the returns distribution. Furthermore, we

have included four real-world constraints, namely the cardinality, buy-in threshold, budget and turnover constraints, in order to provide a complete control on the portfolio composition. To solve this family of optimization problems, we have developed a novel swarm optimization algorithm equipped with an ad-hoc constraint handling technique combining the global convergence properties of the  $\ell_1$ -exact penalty functions with a repair operator.

The empirical findings are obtained on two large-scale data sets of the Pacific and World areas, which include several hundreds of stocks, covering the last 14 years. We have performed a sensitivity analysis for the portfolio size and for the limit of the trades magnitude, in order to identify the best combination of these parameters in terms of ex-post performance and management cost. Results show that portfolios with a reduced number of constituents (15% of the investment pool) and with a two-sided turnover up to 10% provide, on both data sets, better profits which are stable over time.

A more detailed analysis reveals that the inclusion of higher-order moments in the performance measures produces superior results in terms of net wealth with respect to the benchmark and to the portfolio optimized through the standard Sharpe ratio. This is more evident after the pandemic outbreak of 2020, where more market fluctuations are present.

In the future works, on the one hand, we plan to extend the above experimental analysis to other markets and, on the other hand, we are interested in the possibility to introduce similar multi-moment performance measures in a passive investment framework.

**Acknowledgements** The first and third authors are members of the INdAM (Italian Institute for Advanced Mathematics) group.

**Author contributions** All authors contributed equally to this work.

**Funding** Open access funding provided by Università degli Studi di Trieste within the CRUI-CARE Agreement.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Amédée-Manesme C, Barthélémy F (2022) Proper use of the modified Sharpe ratios in performance measurement: rearranging the Cornish Fisher expansion. *Ann Oper Res* 313:691–712. <https://doi.org/10.1007/s10479-020-03858-4>
- Arditti FD (1967) Risk and the required return on equity. *J Finance* 22(1):19–36. <https://doi.org/10.1111/j.1540-6261.1967.tb01651.x>
- Beraldi P, Violi A, Ferrara M, Ciancio C, Pansera BA (2021) Dealing with complex transaction costs in portfolio management. *Ann Oper Res* 299(1):7–22. <https://doi.org/10.1007/s10479-019-03210-5>
- Chaigneau P, Eeckhoudt I (2020) Downside risk-neutral probabilities. *Econ Theory Bull* 8:65–77. <https://doi.org/10.1007/s40505-019-00165-5>
- Chekhlov A, Uryasev S, Zabarankin M (2005) Drawdown measure in portfolio optimization. *Int J Theor Appl Finance* 8(1):13–58. <https://doi.org/10.1142/S0219024905002767>
- Cheng R, Jin Y (2015) A competitive swarm optimizer for large scale optimization. *IEEE Trans Cybern* 45(2):191–204. <https://doi.org/10.1109/TCYB.2014.2322602>
- Cheng R, Jin Y (2015) A social learning particle swarm optimization algorithm for scalable optimization. *Inf Sci* 291:43–60. <https://doi.org/10.1016/j.ins.2014.08.039>
- Cheridito P, Kromer E (2013) Reward-risk ratios. *J Investig Strategy* 3(1):3–18. <https://doi.org/10.21314/JOIS.2013.022>
- Colasante A, Riccetti L (2020) Risk aversion, prudence and temperance: It is a matter of gap between moments. *J Behav Exp Finance* 25:100262. <https://doi.org/10.1016/j.jbef.2019.100262>
- Colasante A, Riccetti L (2021) Financial and non-financial risk attitudes: What does it matter? *J Behav Exp Finance* 30:100494. <https://doi.org/10.1016/j.jbef.2021.100494>
- Corazza M, di Tollo G, Fasano G, Pesenti R (2021) A novel hybrid PSO-based metaheuristic for costly portfolio selection problems. *Ann Oper Res* 304:109–137. <https://doi.org/10.1007/s10479-021-04075-3>
- Cura T (2009) Particle swarm optimization approach to portfolio optimization. *Nonlinear Anal Real World Appl* 10(4):2396–2406. <https://doi.org/10.1016/j.nonrwa.2008.04.023>
- Deguest R, Martellini L, Meucci A (2022) Risk parity and beyond—from asset allocation to risk allocation decisions. *J Portf Manag* 48(4):108–135. <https://doi.org/10.3905/jpm.2022.1.340>
- Derrac J, Garcia S, Molina D, Herrera F (2011) A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm Evol Comput* 1(1):3–18. <https://doi.org/10.1016/j.swevo.2011.02.002>
- Gao J, Gao X, Liu X, Wang Z (2022) Detecting prudence and temperance in risk exposure: the hybrid variance framework. *J Risk* 24(5):75–88. <https://doi.org/10.21314/JOR.2022.034>
- Gilli M, Schumann E (2012) Large-scale portfolio optimization with heuristics. In: “Advanced statistical methods for the analysis of large data-sets”. Springer, Berlin, Heidelberg, pp 181–192 [https://doi.org/10.1007/978-3-642-21037-2\\_17](https://doi.org/10.1007/978-3-642-21037-2_17)
- Guastaroba G, Mansini R, Speranza MG (2009) On the effectiveness of scenario generation techniques in single-period portfolio optimization. *Eur J Oper Res* 192(2):500–511. <https://doi.org/10.1016/j.ejor.2007.09.042>
- Guerard JB (2009) *Handbook of portfolio construction: contemporary applications of Markowitz techniques*. Springer Science & Business Media, Berlin
- Hitaj A, Zambruno G (2016) Are Smart Beta strategies suitable for hedge fund portfolios? *Rev Financial Econ* 29:37–51. <https://doi.org/10.1016/j.rfe.2016.03.001>
- Jurczenko E, Maillat B (2006) *Multi-moment asset allocation and pricing moments*. John Wiley & Sons, New Jersey
- Kaucic M, Barbini F, Camerota FJ (2020) Polynomial goal programming and particle swarm optimization for enhanced indexation. *Soft Comput* 24:8535–8551. <https://doi.org/10.1007/s00500-019-04378-5>
- Kaucic M, Piccotto F (2022) A level-based learning swarm optimizer with a hybrid constraint-handling technique for large-scale portfolio selection problems. *IEEE Congr Evol Comput* 2022:1–8. <https://doi.org/10.1109/CEC55065.2022.9870358>
- Kaucic M, Piccotto F, Sbaiz G, Valentinuz G (2023) A hybrid level-based learning swarm algorithm with mutation operator for solving large-scale cardinality-constrained portfolio optimization problems. *Inf Sci* 634:321–339. <https://doi.org/10.1016/j.ins.2023.03.115>

- Kim WC, Fabozzi FJ, Cheridito P, Fox C (2014) Controlling portfolio skewness and kurtosis without directly optimizing third and fourth moments. *Econ Lett* 122(2):154–158. <https://doi.org/10.1016/j.econlet.2013.11.024>
- Kimball MS (1990) Precautionary saving in the small and in the large. *Econometrica* 58(1):53–73. <https://doi.org/10.2307/2938334>
- Konno H, Shirakawa H, Yamazaki H (1993) A mean-absolute deviation-skewness portfolio optimization model. *Ann Oper Res* 45:205–220. <https://doi.org/10.1007/BF02282050>
- Lai TY (1991) Portfolio selection with skewness: a multiple-objective approach. *Rev Quant Financial Acc* 1:293–305. <https://doi.org/10.1007/BF02408382>
- Lai KK, Yu L, Wang S (2006) Mean-variance-skewness-kurtosis-based portfolio optimization. In: 2006 International symposium on computer and computer sciences, pp 292–297. <https://doi.org/10.1109/IMSCCS.2006.239>
- Li X, Yao X (2012) Cooperatively coevolving particle swarms for large scale optimization. *IEEE Trans Evol Comput* 16(2):210–224. <https://doi.org/10.1109/TEVC.2011.2112662>
- Liu S, Wang SY, Qiu W (2003) Mean-variance-skewness model for portfolio selection with transaction costs. *Int J Syst Sci* 34(4):255–262. <https://doi.org/10.1080/0020772031000158492>
- Markowitz H (1952) Portfolio selection. *J Finance* 7(1):77–91. <https://doi.org/10.2307/2975974>
- Meghwani SS, Thakur M (2017) Multi-criteria algorithms for portfolio optimization under practical constraints. *Swarm Evol Comput* 37:104–125. <https://doi.org/10.1016/j.swevo.2017.06.005>
- Moral-Escudero R, Ruiz-Torrubiano R, Suarez A (2006) Selection of optimal investment portfolios with cardinality constraints. In: IEEE conference on evolution and computing, pp 2382–2388. <https://doi.org/10.1109/CEC.2006.1688603>
- Nagy BZ, Benedek B (2021) Higher co-moments and adjusted Sharpe ratios for cryptocurrencies. *Finance Res Lett* 39:1–7. <https://doi.org/10.1016/j.frl.2020.101543>
- Oldewage ET, Engelbrecht AP, Cleghorn CW (2020) Movement patterns of a particle swarm in high dimensional spaces. *Inf Sci* 512:1043–1062. <https://doi.org/10.1016/j.ins.2019.09.057>
- Omidvar MN, Li X, Yao X (2022) A review of population-based metaheuristics for large-scale black-box global optimization-Part I. *IEEE Trans Evol Comput* 26(5):802–822. <https://doi.org/10.1109/TEVC.2021.3130838>
- Omidvar MN, Li X, Yao X (2022) A review of population-based metaheuristics for large-scale black-box global optimization-Part II. *IEEE Trans Evol Comput* 26(5):823–843. <https://doi.org/10.1109/TEVC.2021.3130835>
- Pézier J, White A (2008) The relative merits of alternative investments in passive portfolios. *J Altern Invest* 10(4):37–49. <https://doi.org/10.3905/jai.2008.705531>
- Politis D, Romano JP (1994) The stationary bootstrap. *J Am Stat Assoc* 89(428):1303–1313. <https://doi.org/10.1080/01621459.1994.10476870>
- Politis D, White H (2004) Automatic block-length selection for the dependent bootstrap. *Econom Rev* 23(1):53–70. <https://doi.org/10.1081/ETC-120028836>
- Samuelson PA (1970) The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Rev Econ Stud* 37(4):537–542. <https://doi.org/10.2307/2296483>
- Scott RC, Horvath PA (1980) On the direction of preference for moments of higher order than the variance. *J Finance* 35(4):915–919. <https://doi.org/10.2307/2327209>
- Sharpe WF (1994) The Sharpe ratio. *J Portf Manag* 21(1):45–58. <https://doi.org/10.3905/jpm.1994.409501>
- Shen W, Wang J, Ma S (2014) Doubly regularized portfolio with risk minimization. *Proc AAAI Conf Artif Int* 28(1):1286–1292. <https://doi.org/10.1609/aaai.v28i1.8906>
- Song G-W, Yang Q, Gao X-D, Ma Y-Y, Lu Z-Y, Zhang J (2021) An adaptive level-based learning swarm optimizer for large-scale optimization. *IEEE Trans Syst Man Cybern*. <https://doi.org/10.1109/SMC52423.2021.9658644>
- Van der Bergh F, Engelbrecht AP (2004) A cooperative approach to particle swarm optimization. *IEEE Trans Evol Comput* 8(3):225–239. <https://doi.org/10.1109/TEVC.2004.826069>
- van Zyl ET, Engelbrecht AP (2015) A subspace-based method for PSO initialization. In: IEEE symposium series on computational intelligence, pp 226–233. <https://doi.org/10.1109/SSCI.2015.42>
- Wang D, Tan D, Liu L (2018) Particle swarm optimization algorithm: an overview. *Soft Comput* 22(2):387–408. <https://doi.org/10.1007/s00500-016-2474-6>
- Wang F, Wang X, Sun S (2022) A reinforcement learning level-based particle swarm optimization algorithm for large-scale optimization. *Inf Sci* 602:298–312. <https://doi.org/10.1016/j.ins.2022.04.053>

- Yang Q, Chen W-N, Da Deng J, Li Y, Gu T, Zhang J (2018) A level-based learning swarm optimizer for large-scale optimization. *IEEE Trans Evol Comput* 22(4):578–594. <https://doi.org/10.1109/TEVC.2017.2743016>
- Zakamuline V, Koekebakker S (2009) Portfolio performance evaluation with generalized Sharpe ratios: beyond the mean and variance. *J Bank Finance* 33(7):1242–1254. <https://doi.org/10.1016/j.jbankfin.2009.01.005>
- Zhu H, Wang Y, Wang K, Chen Y (2011) Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. *Expert Syst Appl* 38(8):10161–10169. <https://doi.org/10.1016/j.eswa.2011.02.075>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.