Quality & Quantity
https://doi.org/10.1007/s11135-026-02918-4

Pseudo-principal portfolios: a risk-reduction framework for benchmark investing

Mario Maggi[1,2] · **Pierpaolo Uberti**[3]

Received: 29 October 2025 / Accepted: 13 June 2026

Abstract

This paper presents a generalization of principal component analysis aimed at reducing the risk of any given benchmark portfolio. Taking the benchmark investment as the starting point, we propose a method to iteratively construct an orthogonal basis of the return space. These orthogonal vectors, called pseudo-principal portfolios after suitable normalization, are combined with the benchmark through an allocation rule to achieve risk reduction. From a theoretical perspective, we connect this construction to the mean–variance framework and derive geometric properties with meaningful financial implications. From an empirical perspective, we provide in-sample and out-of-sample experiments on different datasets of real financial data to support the effectiveness of the strategy, which combines the original benchmark with the pseudo-principal portfolios to lower risk, measured in terms of return volatility.

Keywords Diversification · Portfolio risk · Benchmark beating · Generalized principal components analysis · Out-of-sample performance

JEL Classification D81 · G1 · G11

Mario Maggi and Pierpaolo Uberti contributed equally to this work.

✉ Mario Maggi
mario.maggi@unipv.it

Pierpaolo Uberti
pierpaolo.uberti@unimib.it

1 Department of Economics and Management, University of Pavia, Pavia, Italy

2 CAMRisk, University of Pavia, Pavia, Italy

3 Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy

Published online: 20 June 2026

1 Introduction

Modern Portfolio Theory dates back to the seminal work of Markowitz (1952), which introduced the well-known mean-variance model. Despite the popularity of Markowitz's framework and its intuitive interpretation in terms of risk and return, it is important to emphasize the central role of risk management relative to return forecasting. Indeed, if an investor could reliably predict even the sign of future returns, there would be no need to hedge against risk, as losses could be avoided. Consequently, when constructing investment portfolios, the primary focus should be on risk control, given the difficulty of forecasting returns accurately, both in terms of magnitude and sign Lee (2011).

In practice, portfolio risk is often evaluated relative to a given benchmark, while absolute risk is usually not considered. As noted in Basak et al. (2006), "*Portfolio theory must address the fact that, in reality, portfolio managers are evaluated relative to a benchmark, and therefore adopt risk management practices to account for the benchmark performance.*" The importance of benchmark comparison in evaluating portfolio performance has been widely discussed, see, for example, Strongin et al. (2000); Basak et al. (2006). The benchmark can be another portfolio strategy or, as is common for fund managers, a financial index that also defines the universe of investment opportunities.

The well-known paper DeMiguel et al. (2009) compares the out-of-sample performance of optimization-based approaches with the so-called naive diversification, i.e., the equally weighted (EW) portfolio, which serves as the benchmark in that context. The out-of-sample empirical results highlight the difficulty of outperforming the EW portfolio in terms of Sharpe ratio, turnover, and other measures within the mean-variance framework, especially when estimation windows are short and proportional transaction costs are considered. This somewhat counterintuitive result has sparked extensive discussion in the literature. Some researchers have investigated theoretical explanations for these findings Malladi and Fabozzi (2017), with model uncertainty identified as a primary cause of the poor out-of-sample performance of optimization-based approaches Pflug et al. (2012). Another stream of research points to the numerical instability of the Markowitz model as the main factor behind its weak out-of-sample results Hirschberger et al. (2010); Best and Hlouskova (2008). Additional studies have explored non-trivial portfolio strategies that can outperform the equally weighted benchmark out-of-sample, either to justify the use of complex optimization models Kritzman et al. (2010), or simply to beat a challenging benchmark, see, among others, Kirby and Ostdiek (2012); Ackermann et al. (2017); Bessler et al. (2017); Fugazza et al. (2015); Hanke et al. (2019); Jiang et al. (2019).

In line with the existing literature, we focus on the risk side of portfolio management. The idea of reducing portfolio risk through the use of uncorrelated, orthogonal portfolios composed of the original assets has been studied previously. On one hand, these approaches exploit simplifications in calculating the efficient frontier when the covariance matrix is diagonal Partovi and Caputo (2004). On the other hand, the potential of using uncorrelated, orthogonal portfolios to enhance diversification has been investigated in the literature, see, for example, Meucci (2009); Lohre et al. (2014). The study of principal component analysis in this context remains an active area of research; see, among others, Lassance et al. (2022); Kelly et al. (2023); Imajo et al. (2021); Bonaccolto and Caporin (2024); Liu et al. (2022), and Giglio and Xiu (2021).

In our proposal, we combine the idea of using orthogonal portfolios to reduce risk with the practical need for portfolio managers to evaluate investments relative to a given benchmark. This combined approach leads to a generalization of principal component analysis. The methodological contribution of the paper lies in introducing a benchmark component, with a clear economic interpretation, as the starting point in the construction of a set of uncorrelated factors. The logical structure of the proposal can be summarized as follows. We consider the situation in which a portfolio manager holds a given portfolio or must refer to a benchmark. We study the problem of modifying this starting position to improve its risk profile and, where possible, its risk-adjusted performance relative to the benchmark. We show how appropriate diversification can reduce the out-of-sample risk of any benchmark portfolio. To achieve this, we propose a generalization of principal component analysis that produces what we call pseudo-principal portfolios. Given n assets, the benchmark portfolio defines one direction in the portfolio return space, a subspace of $\mathbb{R}^n$. Our pseudo-principal component analysis takes the benchmark as the starting vector and constructs an orthonormal basis of $\mathbb{R}^n$; after appropriate scaling, each basis vector can be interpreted as a portfolio. The pseudo-principal portfolios are computed iteratively: at each step, we solve a constrained optimization problem that depends on the vectors (portfolios) found in previous steps. The benchmark is then combined with a subset of these pseudo-principal portfolios to capture diversification benefits.

In in-sample exercises, all pseudo-principal portfolios can be added to incrementally increase diversification. In out-of-sample settings, however, only the first few pseudo-principal portfolios contribute meaningfully to risk reduction. This parallels standard PCA applications, where the leading components capture most of the signal, while later components mostly contain noise (see, e.g., Peres-Neto et al. 2005).

We perform a comprehensive empirical analysis using different databases of real financial data. Both in-sample and out-of-sample experiments confirm the effectiveness of our approach in reducing risk relative to the benchmark portfolio across all considered datasets. To test the robustness of our findings, we evaluate risk reduction using multiple risk measures: standard deviation, value at risk, and conditional value at risk (both computed at the 5% level). Within the framework of risk measures, the positive impact of diversification on risk reduction is reflected in the sub-additivity property (see, for example, Artzner et al. (1999)). Given the generality of our approach, we present empirical results for two standard benchmark portfolio strategies: the equally weighted (EW) portfolio and the equal risk contribution (ERC) portfolio DeMiguel et al. (2009); Maillard et al. (2010).

The paper is organized as follows. Section 2 presents the formal definition of pseudo-principal portfolios and the allocation rule based on their combination. The two subsections of Sect. 3 report the in-sample and out-of-sample results, respectively, and Sect. 4 concludes the paper.

2 Pseudo principal portfolios

Given n risky assets, let x_j denote the percentage of wealth allocated to the j^{th} asset in the portfolio $\mathbf{x}$, and define the portfolio set as $W_n = \left\{ \mathbf{x} = [x_1, \dots, x_n]' \in \mathbb{R}^n : \sum_{j=1}^{n} x_j = 1 \right\}$

. We denote by $\mathbf{x}^0 \in W_n$ the benchmark portfolio. Let $T \geq n$ be a positive integer, and let

$\mathrm{Mat}T \times n(\mathbb{R})$ denote the set of $T \times n$ full-rank real matrices. Moreover, let $\mathbf{1} \in \mathbb{R}^n$ be the column vector with unit entries. Interpreting T as the sample size, a matrix $A \in \mathrm{Mat}T \times n(\mathbb{R})$ collects the returns: each column A^j represents the time series of T returns of asset j. The full-rank assumption on A is standard in empirical applications and ensures that the return of one asset cannot be replicated by a portfolio composed of the remaining $n-1$ assets. When $T \gg n$, this assumption is typically satisfied in practice; consequently, the covariance matrix of returns V is symmetric, positive definite, and non-singular. Using standard notation, we denote transposition by the prime, while $|\cdot|_2$ indicates the Euclidean norm of a vector. The principal notation used throughout the paper is summarized in the following Table 1.

We now introduce the notion of pseudo-principal components and, consequently, pseudo-principal portfolios.

Definition 1 Let $\mathbf{x}^0 \in W_n$ be the benchmark portfolio, then for $i = 1, \ldots, n-1$, the i^{th} **pseudo-principal component** $\overline{\mathbf{x}}^i$ can be recursively obtained as the solution of the optimization problem

$$\begin{aligned} \max_{\overline{\mathbf{x}}^i} \quad & (\overline{\mathbf{x}}^i)' \, V \, \overline{\mathbf{x}}^i \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{x}^0 & \overline{\mathbf{x}}^1 & \cdots & \overline{\mathbf{x}}^{i-1} \end{bmatrix}' V \, \overline{\mathbf{x}}^i = [0] \\ & \|\overline{\mathbf{x}}^i\|_2 = 1 \end{aligned} \tag{1}$$

where $\begin{bmatrix} \mathbf{x}^0 & \overline{\mathbf{x}}^1 \cdots \overline{\mathbf{x}}^{i-1} \end{bmatrix}$ is the matrix whose columns are $\mathbf{x}^0$ and the pseudo-principal components with $k < i$, i.e. the ones obtained at the previous steps of the recursion.

The orthogonality constraint in optimization problem (1) ensures that $[\mathbf{x}^0 \; \overline{\mathbf{x}}^1 \cdots \overline{\mathbf{x}}^{i-1}]$ is a full-rank matrix. Therefore, at the end of the $n-1$ recursions, an orthogonal basis of $\mathbb{R}^n$ is obtained. This basis is an alternative to the PCA basis, since the first vector $\mathbf{x}^0$ is the given benchmark. While PCA can be defined in many different ways, the proposed iterative optimization procedure, when implemented n times without fixing the first vector, produces the standard PCA basis; see, for example, (Jolliffe, 2011, page 6). Therefore, our proposal differs from standard principal component analysis because it is defined starting from a given vector-the benchmark investment-which has a direct financial interpretation. From a geometric perspective, pseudo-principal components share with PCA the property of providing an orthogonal basis for $\mathbb{R}^n$: PCA yields an optimal basis in terms of variance decomposition, while pseudo-principal analysis focuses on variance decomposition relative to a specific starting point. Figure 1 illustrates an example in a three-asset market, allowing for a visual comparison of the basis vectors obtained with the two methods. Two benchmarks are considered: the equally weighted (EW) portfolio and $x^0 = [1.1, -0.2, 0.1]$. As expected, the pseudo-principal components change depending on the chosen starting portfolio.

Table 1 Principal notation used throughout the paper

Notation	
n	Number of risky assets
$\mathbf{x}$	Vector of portfolio weights
$\mathbf{x}_j$	Share invested in asset j, with $j = 1, \ldots, n$
$\mathbf{x}^0$	Benchmark portfolio
V	Covariance matrix of returns
T	Sample size

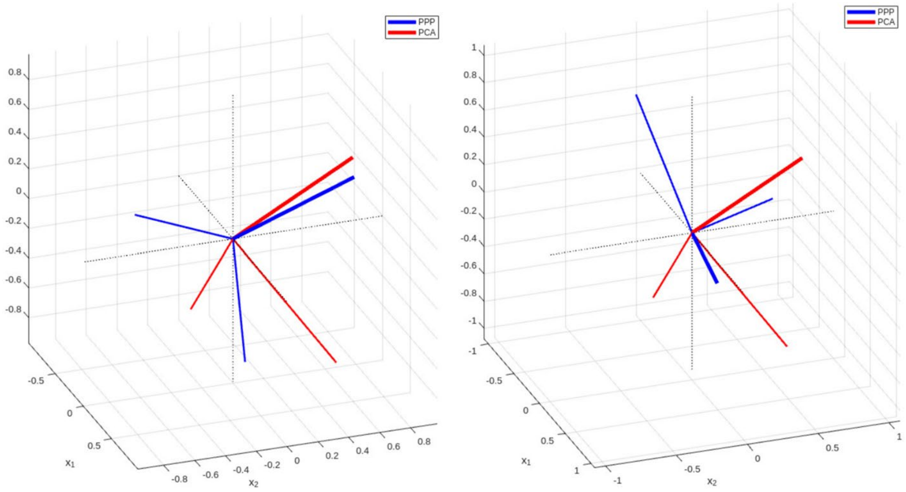

Fig. 1 Unit norm vectors collinear to the principal (red) and the pseudo-principal (blue) components. The benchmark portfolio is $x^0 = x^{\mathrm{EW}}$ on the left plot, and $x^0 = [1.1, -0.2, 0.1]$ on the right plot. The thicker lines indicate the first components

Definition 2 If $\mathbf{1}'\overline{\mathbf{x}}^i \neq 0$, then $\mathbf{x}^i = \dfrac{1}{\mathbf{1}'\overline{\mathbf{x}}^i}\,\overline{\mathbf{x}}^i \in W_n$ is the i^{th} **pseudo-principal portfolio**.

The i^{th} pseudo-principal portfolio is obtained by rescaling the i^{th} pseudo-principal component so that its elements sum to one. This is possible only when the sum of its elements is non-zero.

Remark 1 Definitions 1 and 2 imply the following direct consequences.

1. By construction, the vectors $\mathbf{x}^0, \overline{\mathbf{x}}^1, \ldots, \overline{\mathbf{x}}^{n-1}$ constitute an orthonormal basis of $\mathbb{R}^n$. This follows from the first condition of problem (1).
2. Definition 1 introduces the pseudo-principal components recursively. We observe that, if $\mathbf{x}^0$ is proportional to one of the eigenvectors of $A'A$ – i.e., if it is a principal component –, then $\left\{\mathbf{x}^0, \overline{\mathbf{x}}^1, \ldots, \overline{\mathbf{x}}^{n-1}\right\}$ is the basis of $\mathbb{R}^n$ identified by principal component analysis. Indeed, given the covariance matrix V, the direction of the i-th principal component $\mathbf{w}^i$ is obtained by recursively solving for all $k = 2, \ldots, i$ the optimization problem

$$\begin{aligned} \max_{\mathbf{w}^k} \quad & (\mathbf{x}^k)'\, V\, \overline{\mathbf{w}}^k \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{w}^1 & \cdots & \mathbf{w}^{k-1} \end{bmatrix}' V\, \mathbf{x}^k = [0] \\ & \|\mathbf{w}^k\|_2 = 1 \end{aligned} \tag{2}$$

where the first component ($i = 1$) solves

$$\max_{\|\mathbf{w}^1\|_2 = 1} \left(\mathbf{w}^1\right)' V\, \mathbf{w}^1,$$

i.e, $\mathbf{w}^1$ is the eigenvector of V corresponding to its largest eigenvalue. Therefore, if $\mathbf{x}^0 = \mathbf{w}^i$, for any $i = 1, \ldots n$, the iterations defined in (1) yield the same solutions as problem (2) (only the order of the components may differ). In this sense, the pseudo-principal components can be regarded as a generalization of the principal components.

3. While the assumption on the rank of A guarantees the existence of n pseudo-principal components, the number of pseudo-principal portfolios is at most $\tilde{n} \leq n$. This is because, if $\mathbf{1}'\overline{\mathbf{x}}^i = 0$, the corresponding pseudo-principal portfolio is not defined. As will be noted in the allocation rule proposed next (see Procedures 1 and 2), this situation does not affect either the theoretical or the empirical results. As with standard PCA, the solution of problem (1) is unique up to a change of sign; see Jolliffe (2011).[1] Therefore, if a vector $\mathbf{x}$ is a solution of problem (1), then so is $-\mathbf{x}$. What is particularly interesting in our application is that, in terms of pseudo-principal portfolios, the two solutions coincide, since if $\mathbf{1}'\mathbf{x} \neq 0$, then $\frac{\mathbf{x}}{\mathbf{x}'\mathbf{1}} = \frac{-\mathbf{x}}{-\mathbf{x}'\mathbf{1}}$.
4. We also note that pseudo-principal portfolios can exhibit very extreme positive and negative weights. This suggests that these portfolios should be used judiciously to make diversification practical, without compromising the economic interpretation of the results. For further discussion, see Sect. 3.
5. In problem (1), each $\overline{\mathbf{x}}^i$ is the solution of a constrained optimization problem that requires a numerical approach, as no closed-form solution is available.[2].

In the context of portfolio management, pseudo-principal portfolios can serve as a useful tool when combined with the benchmark portfolio $\mathbf{x}^0$ according to a specified allocation rule. In the following, we distinguish between cases where short positions are allowed (*long-short* portfolios) and cases where they are not (*long-only* portfolios). Define $\mathbf{p}^i$, for $i = 0, \ldots, \tilde{n}$, as the portfolio obtained by combining the benchmark $\mathbf{x}^0$ with the first i pseudo-principal portfolios. For simplicity, we set $\mathbf{p}^0 = \mathbf{x}^0$. We denote $\sigma^2_{\mathbf{x}}$ as the variance of the portfolio $\mathbf{x}$. The entries of a vector are indicated with a subscript, so that, for example, $\left(\mathbf{x}^1\right)_2$ is the second component of $\mathbf{x}^1$.

Procedure 1 (*Allocation Rule, Long-Short*) Given the benchmark portfolio $\mathbf{x}^0$ and its pseudo-principal portfolios $\{\mathbf{x}^i\}$, with $i = 1, \ldots, \tilde{n} \leq n-1$, the i^{th} combined portfolio is

$$\begin{cases} \mathbf{p}^0 = \mathbf{x}^0 \\ \mathbf{p}^i = (1-\alpha_i)\,\mathbf{p}^{i-1} + \alpha_i\,\mathbf{x}^i, \quad \text{for} \quad i = 1, \ldots, \tilde{n} \end{cases}$$

where the weight α_i is

$$\alpha_i = \frac{\left(\left(\mathbf{x}^i\right)' V \mathbf{x}^i\right)^{-1}}{\sum_{j=0}^{i} \left(\left(\mathbf{x}^j\right)' V \mathbf{x}^j\right)^{-1}}, \quad \text{or equivalently} \quad \alpha_i = \frac{\sigma^2_{\mathbf{p}^{i-1}}}{\sigma^2_{\mathbf{p}^{i-1}} + \sigma^2_{\mathbf{x}^i}}, \tag{3}$$

[1] Similarly, when calculating principal components through the eigenvalue-eigenvector problem, each unit-norm eigenvector is unique up to a change of sign.

[2] In the empirical application, the solution of Problem 1 is computed using the MATLAB function fmincon Since no particular numerical issues were encountered, we do not focus on this aspect.

The two formulations in (3) are equivalent. Indeed, since $\sigma^2_{\mathbf{x}^i} = \left(\mathbf{x}^i\right)' V\mathbf{x}^i$, some algebra yields

$$\frac{\left(\left(\mathbf{x}^i\right)' V\mathbf{x}^i\right)^{-1}}{\sum_{j=0}^{i}\left(\left(\mathbf{x}^j\right)' V\mathbf{x}^j\right)^{-1}} = \frac{\left(\mathbf{p}^{i-1}\right)' V\mathbf{p}^{i-1}}{\left(\mathbf{p}^{i-1}\right)' V\mathbf{p}^{i-1} + \left(\mathbf{x}^i\right)' V\mathbf{x}^i}, \quad \text{i.e.} \quad \frac{\sigma^2_{\mathbf{p}^{i-1}}}{\sigma^2_{\mathbf{p}^{i-1}} + \sigma^2_{\mathbf{x}^i}}$$

The first formula in (3) is more intuitive, as it directly reflects the minimum-variance allocation. Additionally, this formulation helps simplify the proof of Proposition 3. However, from a computational perspective, the second formulation is preferable due to its faster computation time.

In the case of a long-only requirement, we propose the following allocation rule (see Procedure 2), designed to select the largest adjustment portfolio that is collinear with the pseudo-principal component while avoiding short positions in the resulting allocation.

Procedure 2 (*Allocation Rule, Long-Only*) Given the benchmark portfolio $\mathbf{x}^0$ and its pseudo-principal portfolios $\{\mathbf{x}^i\}$, with $i = 1, \ldots, \tilde{n} \leq n-1$, the i^{th} combined portfolio is

$$\begin{cases} \mathbf{p}^0 = \mathbf{x}^0 \\ \mathbf{p}^i = (1-\alpha_i)\,\mathbf{p}^{i-1} + \alpha_i\,\mathbf{x}^i, \quad \text{for} \quad i = 1, \ldots, \tilde{n} \end{cases}$$

where $\alpha_i = \min\left\{\frac{\left(\mathbf{p}^{i-1}\right)' V\mathbf{p}^{i-1}}{\left(\mathbf{p}^{i-1}\right)' V\mathbf{p}^{i-1} + \left(\mathbf{x}^i\right)' V\mathbf{x}^i}, \overline{\alpha}_i\right\}$, where the value of $\overline{\alpha}_i$ is chosen to avoid short positions as follows

$$\overline{\alpha}_i = \min_{j \in \left\{j \mid \left(\mathbf{x}^i\right)_j < 0\right\}} \left\{\frac{\left(\mathbf{p}^i\right)_j}{\left(\mathbf{x}^{i-1}\right)_j - \left(\mathbf{p}^i\right)_j}\right\}.$$

The allocation rules in Procedure 1 and 2 are intended for active management investment strategies. Nevertheless, the mathematical formulation permits the application of the approach also in a passive context. The application (see Sect. 3) is performed in an active framework. If we combine $\mathbf{x}^0$ with all $\tilde{n}$ pseudo-principal portfolios in the long-short case, following the allocation rule provided in Procedure 1, we obtain the global minimum-variance portfolio, as proven in Proposition 3. Partovi and Caputo (2004) obtained this result for the basis derived from standard principal component analysis.

Proposition 3 *Given* $\mathbf{x}^0 \in W_n$, *if its* $n-1$ *pseudo-principal portfolios* $\{\mathbf{x}^1, \ldots, \mathbf{x}^{n-1}\}$ *are well defined, then* $\mathbf{p}^{n-1}$ *in the long-short framework corresponds to the global minimum variance portfolio.*

Proof Let us assume that $\left\{\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}\right\}$ are the available assets. The allocation rule of Procedure 1 leads, by backward substitution, to

$$\mathbf{p}^{n-1} = \sum_{j=0}^{n-1} \mathbf{x}^j \frac{\left(\left(\mathbf{x}^i\right)' V\mathbf{x}^i\right)^{-1}}{\sum_{j=0}^{i} \left(\left(\mathbf{x}^j\right)' V\mathbf{x}^j\right)^{-1}}.$$

The covariance matrix W of the portfolios $\left\{\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}\right\}$ is equal to

$$W = X' V X = \begin{bmatrix} \sigma^2_{\mathbf{x}^0} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\mathbf{x}^1} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\mathbf{x}^{n-1}} \end{bmatrix}, \tag{4}$$

where $\sigma^2_{\mathbf{x}^i} = \left(\mathbf{x}^i\right)' V\mathbf{x}^i,\ i = 0, \ldots, n-1$, and $\mathbf{X} = \left[\ \mathbf{x}^0 \cdots \mathbf{x}^{i-1}\ \right]$ is the matrix whose columns are the pseudo-principal portfolios. Then, since W is diagonal, the weights of the global minimum variance portfolio are

$$\frac{W^{-1}\mathbf{1}}{\mathbf{1}'W^{-1}\mathbf{1}} = \left[\ \frac{\left(\left(\mathbf{x}^1\right)' V\mathbf{1}^1\right)^{-1}}{\sum_{j=0}^{i}\left(\left(\mathbf{x}^j\right)' V\mathbf{x}^j\right)^{-1}} \cdots \frac{\left(\left(\mathbf{x}^{n-1}\right)' V\mathbf{x}^{n-1}\right)^{-1}}{\sum_{j=0}^{i}\left(\left(\mathbf{x}^j\right)' V\mathbf{x}^j\right)^{-1}}\ \right]'$$

Considering that the assets $\left\{\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}\right\}$ are portfolios of the original assets, the global minimum variance portfolio is

$$\left[\ \frac{((\mathbf{x}^1)' V\mathbf{1}^1)^{-1}}{\sum_{j=0}^{i}((\mathbf{x}^j)' V\mathbf{x}^j)^{-1}} \cdots \frac{((\mathbf{x}^{n-1})' V\mathbf{x}^{n-1})^{-1}}{\sum_{j=0}^{i}((\mathbf{x}^j)' V\mathbf{x}^j)^{-1}}\ \right] \left[\mathbf{x}^0 \cdots \mathbf{x}^{i-1}\right]' = \mathbf{p}^{n-1}.$$

□

The result in Proposition 3 is derived assuming that all $n-1$ pseudo-principal portfolios are well defined. The following corollary shows that the result remains valid for $\tilde{n} < n-1$; this situation arises when one or more pseudo-principal portfolios are not defined due to a zero sum of the entries in the corresponding pseudo-principal components.

Remark 2 The result of Proposition 3 is essentially due to the orthogonality of the pseudo-principal portfolios in the return space. Therefore, the proposed allocation rule produce the same result (i.e., $\mathbf{p}^{n-1}$ is the global minimum variance portfolio) as long as the set of considered portfolios $\{x^0, x^1, \ldots, x^{n-1}\}$ span the return space and have non correlated returns.

Remark 3 In the long–short context, Procedure 1 can be applied to any given starting benchmark, with no significant consequences or limitations. By contrast, in a long-only context, starting from a highly concentrated benchmark portfolio could drastically reduce the possibilities of diversification through Procedure 2. Indeed, a highly concentrated benchmark is already close to the boundary of the simplex (the set of all long-only portfolios), thereby geometrically reducing opportunities for diversification through combinations of orthogonal portfolios with negative exposures to some assets. Nevertheless, in practice, investors' standard reference benchmarks are broad financial indices, which represent the investment universe.

Corollary 4 *If, for a given $i \in \{1, n-1\}$, $\lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \mathbf{1}'\overline{\mathbf{x}}^i = 0$, and $\mathbf{1}'\overline{\mathbf{x}}^j \neq 0$ for $j < i$, then the allocation rule of Procedure 1 yields a vanishing weight $\alpha_i \to 0$.*

Proof Procedure 1 updates the portfolio by assigning to the pseudo portfolio $\mathbf{x}^i$ the weight $\alpha_i = \frac{\left((\mathbf{x}^i)' V \mathbf{x}^i\right)^{-1}}{\sum_{j=0}^{i} \left((\mathbf{x}^j)' V \mathbf{x}^j\right)^{-1}}$, therefore

$$\lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \alpha_i = \lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \frac{\left(\frac{1}{\left(\mathbf{1}'\overline{\mathbf{x}}^i\right)^2} \left(\overline{\mathbf{x}}^i\right)' V \overline{\mathbf{x}}^i\right)^{-1}}{\sum_{j=0}^{i-1} \left(\frac{1}{\left(\mathbf{1}'\overline{\mathbf{x}}^j\right)^2} \left(\overline{\mathbf{x}}^j\right)' V \overline{\mathbf{x}}^j\right)^{-1} + \left(\frac{1}{\left(\mathbf{1}'\overline{\mathbf{x}}^i\right)^2} (\overline{\mathbf{x}}^i)' V \overline{\mathbf{x}}^i\right)^{-1}},$$

which is null since

$$\lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \left(\frac{1}{\left(\mathbf{1}'\overline{\mathbf{x}}^i\right)^2} \left(\overline{\mathbf{x}}^i\right)' V \overline{\mathbf{x}}^i\right)^{-1} = \lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \left(\mathbf{1}'\overline{\mathbf{x}}^i\right)^2 \left(\left(\overline{\mathbf{x}}^i\right)' V \overline{\mathbf{x}}^i\right)^{-1} = 0,$$

being $\lim_{\mathbf{1}'\overline{\mathbf{x}}^i \to 0} \mathbf{1}'\overline{\mathbf{x}}^i = 0$ and the quadratic form positive definite. □

This shows that when the weights of the i^{th} pseudo-principal portfolio are very large—because the sum of the elements of the corresponding pseudo-principal component is close to zero—x^i does not significantly affect the overall allocation.

Remark 4 In practical applications, a pseudo-principal component whose entries sum exactly to zero is virtually impossible to obtain; typically, for some components, the sum may be very close to zero. Although this can produce pseudo-principal portfolios with impractically large absolute weights, they do not affect the proposed allocation for the reason given in Corollary 4. The example below, as well as the empirical application in Sect. 3, shows that only a small number of pseudo-principal portfolios meaningfully influence the allocation. This occurs because the rule in Definition 1 assigns small weights to pseudo-principal portfolios with large variance, and because the variances of the pseudo-principal portfolios increase rapidly after the first few components. Motivated by this observation and the empirical evidence, we introduce a heuristic (Sect. 3.3) to limit the number of pseudo-principal portfolios added to the benchmark.

The peculiarity of our approach is that, unlike other proposals using principal portfolios (see, among others, Meucci (2009); Lohre et al. (2014)), we specify a benchmark strategy to outperform (or at least to compare with), so that the orthogonality and the resulting diversification are tailored specifically to the benchmark. Compared to Markowitz's standard approach, we exploit the fact that the first few principal components efficiently capture the majority of the variance in the data, and the same holds for pseudo-principal portfolios. This can be interpreted in two ways. First, the pseudo-principal components are calculated via a recursion analogous to that used for principal components. As a result, their ability to capture return variability decreases with the index *i*. Consequently, the first pseudo-principal components describe a large portion of the information about return variability, whereas the later components mainly reflect market noise.

Second, to provide an empirical interpretation, we perform the following exercise. Using the 30-industry Fama–French dataset (see Sect. 3 for details), the pseudo-principal components and portfolios are computed on 1,000 bootstrapped sub-samples of 1,000 observations, using the equally weighted portfolio as the benchmark.[3] For each simulated sample, we compute: *(i)* the absolute value of the sum of the elements of the pseudo-principal components and their variance; *(ii)* the leverage[4] of the pseudo-principal portfolios and their variance. The median, as well as the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of the bootstrapped values, are reported in Fig. 2. It is evident that the first few pseudo-components (approximately 3 to 4 out of 30) have a fairly large sum of weights, resulting in corresponding pseudo-principal portfolios with practically reasonable leverage and variance. As the index i increases, the leverage and variance of the pseudo-principal portfolios rise rapidly. Nevertheless, we note that the allocation rule of Procedure 1 effectively manages this behavior: when large variances are input into formula (3), the resulting weights are negligible, keeping the overall portfolio risk under control. This behavior motivates the introduction of a heuristic rule (see Sect. 3.3) to limit the number of pseudo-components effectively employed. Reducing

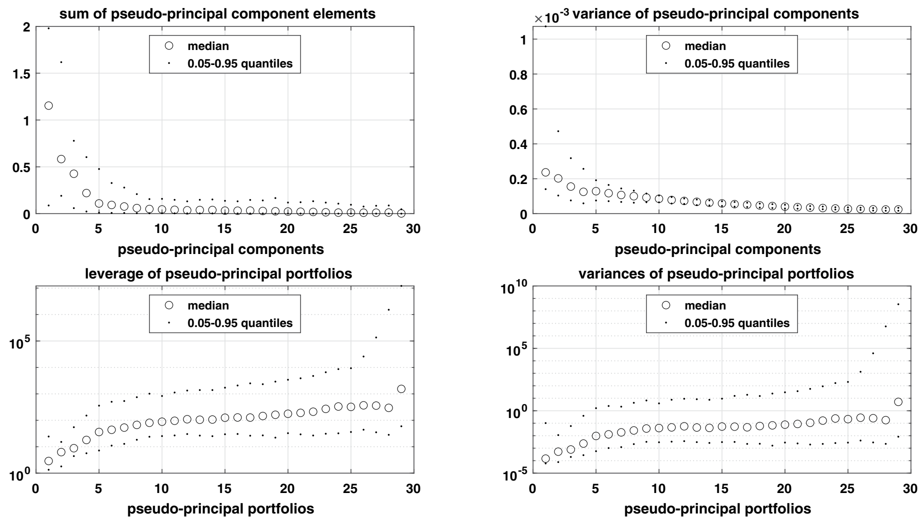

Fig. 2 Absolute value of the sum of the components of the pseudo-principal components (top left) and their variance (top right); leverage of the pseudo-principal portfolios (bottom left) and their variance (bottom right). The results are based on 1,000 bootstrapped subsets of 1,000 observations drawn from the 30-industry Fama–French dataset. The equally weighted portfolio is used as the benchmark. The median, 5$^{\text{th}}$, and 95$^{\text{th}}$ percentiles of the bootstrapped values are shown. Note the use of a log scale for the two bottom plots

[3] The bootstrap procedure was applied to generate a large number of covariance matrices with varying correlation structures. Since the selected period includes significant crises, recoveries, and price surges, it exhibits diverse correlation scenarios. This approach allows us to evaluate the robustness of our findings.

[4] We use the definition of leverage of a portfolio $x \in \mathbb{R}^n$ by Jacobs and Levy (2013): $\Lambda = \left(\sum_{j=1}^{n} |x_j|\right) - 1$

.

the number of considered pseudo-components allows for a practical use of the information contained in the covariance matrix, while avoiding noise in the data. As will become evident in the application, a small number of pseudo-principal portfolios is sufficient to diversify the original strategy, effectively reducing risk and often improving risk-adjusted performance.

Even though the proposed approach is closely related to principal component analysis, it is more effective at reducing risk than portfolios formed from principal components (i.e., portfolios obtained by normalizing the principal components to unit sum). This is due to the anchoring of pseudo-principal components to a starting portfolio, and the subsequent optimization procedure that finds orthogonal components best describing the residual variability. Moreover, PCA portfolios are not tied to a starting benchmark, resulting in the same portfolio set for all investors. It is also possible to use PCA portfolios to adjust a given benchmark, but in this case, the orthogonality is lost. To better illustrate the comparison, we consider the following example. Using the 10-sector Fama-French portfolios introduced in the next section, we compute both PCA and pseudo-principal portfolios for 1,000 random starting benchmarks. We then construct the combined portfolios using the allocation rule in Procedure 1, modifying it when the PCA portfolios lose orthogonality. The sequence of portfolios obtained from pseudo-principal components is compared to the PCA-based portfolios in three possible cases: *(i)* PCA portfolios without a benchmark starting portfolio; *(ii)* PCA portfolios with indices from 1 to 9 (i.e., from 1 to $n-1$) are used in the allocation rule of Procedure 1; *(iii)* PCA portfolios with indices from 2 to 10 (i.e., from 2 to n) are used in the allocation rule of Procedure 1.

As Fig. 3 shows, in all cases the pseudo-principal method is more effective at reducing risk. The higher risk of PCA-based allocations in cases *(i)* and *(ii)* may arise from the fact that PCA aims to decompose variance optimally rather than to minimize it. Consequently, the first principal-component portfolio can be relatively risky, producing – in case *(ii)* – an increase in risk when combined with the starting portfolio (see the hump at $i = 2$ in the second plot of Fig. 3). The slightly higher risk in case *(iii)* may result from PCA portfolios not being orthogonal to the random starting portfolio. We also note that in approach *(i)* the PCA allocation is determined independently of any benchmark, whereas in approaches *(ii)* and *(iii)* the application of PCA may appear arbitrary, as it is imposed on a benchmark without being inherently related to it.

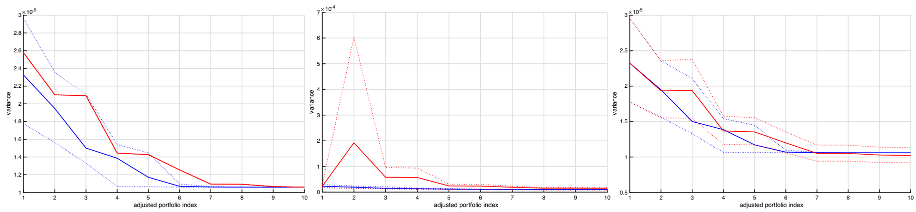

Fig. 3 Comparison of combined portfolios computed using Procedure 1. The horizontal axes indicate the adjusted portfolio index, and the vertical axes show portfolio variances: PCA (red) and pseudo-principal (blue) portfolios. Results are based on 1,000 random starting long portfolios. Continuous lines represent the average variance, while dotted lines indicate the 90% bounds. Left: case *(i)* PCA portfolios without a benchmark starting portfolio. Center: case *(ii)* PCA portfolios with indices 1–9 (1 to $n-1$) used in the allocation rule. Right: case *(iii)* PCA portfolios with indices 2–10 (2 to n) used in the allocation rule

3 Application

This section illustrates how to apply the results described above. These results can be used in risk management. Since pseudo-principal portfolios are uncorrelated with the benchmark portfolio, they are natural candidates to enhance the diversification of the starting position. As benchmarks $\mathbf{x}^0$, we consider the equally weighted (EW) and equal risk contribution (ERC) portfolios. Nevertheless, we emphasize that the procedure is very general and that, qualitatively, the results do not depend on the chosen benchmark. The section is organized in three subsections. Section 3.1 provides a description of the databases used in the empirical application. Section 3.2 presents a simple toy example with a real data portfolio and a small number of constituents $(n = 5)$, allowing us to highlight the main features of pseudo-principal portfolios. This example is performed in-sample to show the theoretical properties of the approach. Section 3.3 presents a more detailed empirical application, highlighting the benefits of the proposed strategy in an out-of-sample framework. We test our results on different real financial datasets to investigate if and how portfolio size affects diversification. This latter application provides relevant insights into the feasibility of applying the proposed methodology in a real-world investment context.

3.1 Data description

For this analysis, we consider the Fama–French US industry portfolios as individual assets. This choice follows the common practice of allocating portfolios across asset classes. Moreover, the number of assets remains relatively small while still providing adequate diversification.

The complete dataset consists of 5 samples, with the number of sectors ranging from 5 to 30. Stock prices are available on Kenneth French's website[5], and are labeled as "Industry Portfolios." The risk-free rate used to compute risk-adjusted performance measures is the US 3-month Treasury rate, available from the Federal Reserve Bank of St. Louis.[6] Daily data span from January 2000 to May 2023. We note that this period includes various market phases, including two of the most severe recent financial crises: the 2008 subprime crisis and the 2020 COVID-19 pandemic.

3.2 In-sample application

This section presents an example highlighting the main features of pseudo-principal portfolios. The assets are the five Fama–French US industry sectors described in Sect. 3.1, sampled from January 2, 2013, to April 1, 2013, with the equally weighted portfolio used as the benchmark.

The covariance matrix of the returns is

[5] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research, accessed July 7, 2023.

[6] See https://fred.stlouisfed.org, accessed July 7, 2023.

$$V = 10^{-4}\begin{bmatrix} 0.1740 & 0.1782 & 0.1630 & 0.1630 & 0.1893 \\ 0.1782 & 0.2905 & 0.2215 & 0.1976 & 0.2758 \\ 0.1630 & 0.2215 & 0.3305 & 0.1749 & 0.2450 \\ 0.1630 & 0.1976 & 0.1749 & 0.2594 & 0.1975 \\ 0.1893 & 0.2758 & 0.2450 & 0.1975 & 0.3200 \end{bmatrix}.$$

The pseudo-principal components are given by the columns of $\bar{X}$, while the matrix X contains the corresponding pseudo-principal portfolios. When translating principal components, which have a purely geometrical meaning, into principal portfolios, it is possible to obtain extreme weights that are difficult to interpret economically and to implement in practice. We address this point below. We emphasize that this issue is common to all approaches based on PCA and arises from two main aspects. First, it frequently occurs that a principal component has an almost vanishing sum of entries, causing numerical issues when rescaling to form a portfolio. Portfolios with a zero sum of weights are known in the literature as arbitrage portfolios. Second, from a geometrical perspective, since the first few principal components inherently capture the majority of the variability in the data, the later principal components—restricted to be orthogonal to the others—often identify particular directions in $\mathbb{R}^n$ characterized by extreme weights.

$$\begin{aligned} \bar{X} &= \begin{bmatrix} 0.4472 & 0.2080 & 0.0573 & 0.8954 & -0.0436 \\ 0.4472 & 0.0349 & -0.5407 & -0.0569 & 0.7116 \\ 0.4472 & -0.6664 & 0.5355 & -0.0313 & 0.0977 \\ 0.4472 & 0.7028 & 0.4911 & -0.3514 & -0.0514 \\ 0.4472 & -0.1323 & -0.4200 & -0.2658 & -0.6925 \end{bmatrix}, \\ X &= \begin{bmatrix} 0.2000 & 1.4142 & 0.4650 & 4.7132 & -1.9993 \\ 0.2000 & 0.2376 & -4.3910 & -0.2993 & 32.6090 \\ 0.2000 & -4.5298 & 4.3487 & -0.1649 & 4.4777 \\ 0.2000 & 4.7774 & 3.9886 & -1.8499 & -2.3569 \\ 0.2000 & -0.8994 & -3.4113 & -1.3991 & -31.7306 \end{bmatrix}. \end{aligned} \tag{5}$$

The variances of the pseudo-principal portfolios are

$$\begin{aligned} &\sigma^2_{\mathbf{x}^0} = 2.15 \times 10^{-5}, \quad \sigma^2_{\mathbf{x}^1} = 5.66 \times 10^{-4}, \quad \sigma^2_{\mathbf{x}^2} = 5.98 \times 10^{-4} \\ &\sigma^2_{\mathbf{x}^3} = 1.03 \times 10^{-4}, \quad \sigma^2_{\mathbf{x}^4} = 5.89 \times 10^{-3}. \end{aligned}$$

Procedure 1 produces the following sequences of weights α_i (%) and the corresponding combined portfolios.

α_0	α_1	α_2	α_3	α_4
100.0000	3.6664	3.3546	16.2692	0.2845

$\mathbf{p}^0 = \mathbf{x}^0$	$\mathbf{p}^1$	$\mathbf{p}^2$	$\mathbf{p}^3$	$\mathbf{p}^4$
0.2000	0.2445	0.2519	0.9777	0.9693
0.2000	0.2014	0.0473	-0.0091	0.0837
0.2000	0.0266	0.1716	0.1168	0.1292
0.2000	0.3678	0.4893	0.1087	0.1017
0.2000	0.1597	0.0399	-0.1942	-0.2839

Despite its simplicity, this example highlights several interesting aspects of our proposal. First, we note that the weights in (5) of the pseudo-principal portfolios are large in absolute value, requiring substantial short positions. This behavior generally increases with i, so the pseudo-principal portfolios would rarely be of practical interest if considered as standalone investment portfolios: in addition to the high leverage they imply, their variance is extremely large. In this case, Procedure 1 generates small weights, resulting in combined portfolios $\mathbf{p}^i$ that appear less extreme. It is evident that the sequence $\{\mathbf{p}^i\}_{i=1,\ldots,n-1}$ converges to the global minimum variance portfolio, as shown in Fig. 4 (left), confirming the findings of Proposition 3. As i increases, the behavior of the pseudo-principal portfolios indicates that the first ones are more manageable in terms of leverage and provide the most significant risk reduction, quickly closing the gap with the efficient frontier. As will become clearer in the application to real data, the perspective of iteratively combining additional orthogonal factors highlights the need to use the information content of the covariance matrix efficiently. As is common in principal component applications, the first few components capture most of the data's volatility, leaving mainly noise in the remaining components. This suggests that only part of the covariance matrix's information content is useful for risk reduction. This intuition is confirmed by the out-of-sample experiment in the next section, which shows that only the first pseudo-principal portfolios contribute meaningfully to risk reduction. Moreover, for practical asset management, the update rule in Procedure 2, under the short-sales restriction, offers an attractive strategy that does not require leverage, achieving

α_0	α_1	α_2	α_3	α_4
100.0000	3.6664	3.3546	2.7727	0.0000
$p_0 = x_0$	p_1	p_2	p_3	p_4
0.2000	0.2445	0.2519	0.3756	0.3756
0.2000	0.2014	0.0473	0.0377	0.0377
0.2000	0.0266	0.1716	0.1622	0.1622
0.2000	0.3678	0.4893	0.4244	0.4244
0.2000	0.1597	0.0399	0.0000	0.0000

The long-only restriction does not allow the full application of the pseudo-principal portfolio adjustment. Consequently, the sequence of portfolios does not reach the global minimum variance portfolio nor the efficient frontier. However, the combined portfolios still reduce risk and, in this example, also improve returns relative to the benchmark portfolio

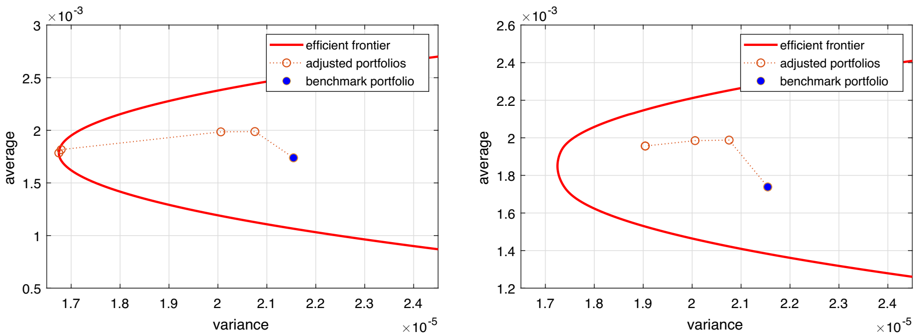

Fig. 4 Representation in the mean–variance plane of the sequence of portfolios obtained by adjusting the benchmark allocation with the pseudo-principal portfolios, following Procedures 1 and 2: no restrictions (left), long-only (right). Continuous lines indicate the portfolio frontiers, and circles represent the individual portfolios

(see Fig. 4, right). In both the long-short and long-only cases, the sequence of combined portfolios moves toward the northwest direction in the mean-variance plane, indicating that risk reduction is accompanied by an increase in return.

3.3 Out-of-sample application

This section presents an empirical application of the pseudo-principal portfolios, consistent with the findings of the previous section. The analysis is conducted in an out-of-sample framework to assess the practical implications for risk management. We implement a long-only rolling investment strategy, using as starting benchmarks $\mathbf{x}^0$ alternatively the equally weighted (EW) and equal risk contribution (ERC) portfolios DeMiguel et al. (2009); Maillard et al. (2010). The means and covariance matrix of asset returns are computed over a sliding window of length w. Portfolios are then constructed and held for a period d, after which they are rebalanced based on the updated data.

We repeat the analysis for each group of industry indices, i.e., for $n \in \{5, 10, 12, 17, 30\}$, using sliding windows of length $w \in \{5n, 10n\}$ and holding periods $d \in \{5, 20\}$ working days. With daily data, holding periods of 5 and 20 working days correspond to weekly and monthly portfolio rebalancing frequencies, respectively, which are commonly used in practice by investors. Although all combinations of sliding windows and holding periods have been implemented, only a subset is reported here for clarity and brevity.

Remark 5 The study of rebalancing frequencies in the application of the investment strategy is beyond the scope of the present paper. For this reason, the experiment is implemented with periodic rebalancing at fixed intervals of one week and one month. We remark that monthly rebalancing represents one of the most frequent scenarios in asset allocation, particularly for medium- to long-term investment horizons. Nevertheless, more active strategies often entail higher frequency portfolio adjustments; consequently, we have also examined weekly rebalancing. Higher frequencies are appropriate for dynamic strategies characterized by more speculative management. Although the proposed instrument could be applied in such cases, we consider its base philosophy to be essentially different and less consistent with such management styles. For this reasons, we think that the analyzed rebalancing frequencies present an acceptable and consistent picture of the potential use of the proposed tool. For general results on the effects of rebalancing strategies on overall performance, we refer to Dichtl et al. (2014) and Dichtl et al. (2016). In the present framework, rebalancing is used primarily for risk reduction. We recognize that the application of an optimization approach (see, among others, Sun et al. (2006); Donohue and Yip (2003), and Antar (2026)) could further improve the performance of the investment strategy, also in terms of returns. The introduction of an endogenous portfolio rebalancing strategy induced by the computation of pseudo-principal portfolios represents a potentially interesting topic for future research.

Given that the contribution of the pseudo-principal portfolios generally decreases with i , we propose to apply Procedures 1 and 2, considering two combined portfolios: *(i)* $\mathbf{p}^1$ the benchmark + the first pseudo-principal portfolio; *(ii)* $\mathbf{p}^*$ the benchmark + a variable number of pseudo-principal portfolios obtained by stopping the recursion when $\alpha_i < 0.01$. This last heuristic rule is justified by the fact that the last pseudo portfolio αs are often negligible. To illustrate this, Fig. 5 shows the number of selected pseudo-principal portfolios, for the

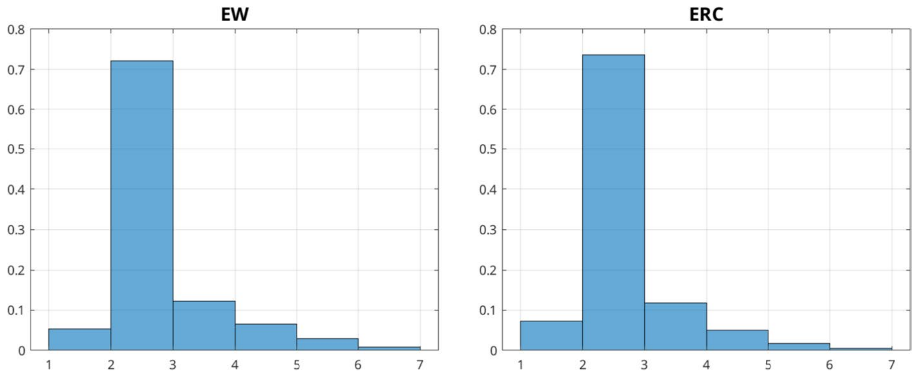

Fig. 5 Frequency distribution of the number of selected pseudo-principal portfolios for the rolling out-of-sample strategy. In the considered case $n = 10, w = 5, d = 5$, the heuristic stopping rule is applied with a threshold of 0.01

case where $n = 10$, $w = 5$, and $d = 5$ (the other cases are omitted for the sake of conciseness). The choice of α_i is arbitrary in our application. We expect investors to set the value of α_i according to their individual preferences and sensitivity when rebalancing portfolio weights.

The overall results of the application are reported in Tables 2 and 3, using the EW and ERC portfolios, respectively, as x^0. For each combination of holding and training periods, we report the return, standard deviation, value at risk, conditional value at risk (both at the 5% significance level), Sharpe ratio, and turnover for the benchmark $\mathbf{x}^0$ and of $\mathbf{p}^1$ and $\mathbf{p}^*$.

Moreover, for each case, we compute the standard deviations and Sharpe ratios over a one-year sliding window. This allows us to measure the frequency of periods in which the combined portfolios $\mathbf{p}^1$ and $\mathbf{p}^*$ outperform the benchmark in terms of both standard deviation and Sharpe ratio. These frequencies are reported in Tables 2 and 3.

Figures 6 and 7 illustrate the impact of the proposed allocation strategy. Notably, for each benchmark and for every combination of training window length and holding period, both $\mathbf{p}^1$ and $\mathbf{p}^*$ reduce risk in nearly 100% of cases, regardless of the risk measure considered. Examining the returns, it is evident that the proposed strategies do not exhibit any systematic positive or negative effect. This is expected, as the strategy focuses solely on risk. Nevertheless, in terms of risk-adjusted performance, measured here by the Sharpe ratio, the results appear promising.

In the majority of scenarios, the proposed strategies improve the Sharpe ratio with a frequency of around 60% or more. Moreover, in the out-of-sample framework, it is confirmed that the most relevant effect is obtained with $\mathbf{p}^1$, which combines only the first pseudo-principal portfolio with the benchmark. While the proposed strategies ($\mathbf{p}^1$ and $\mathbf{p}^*$) consistently outperform the benchmarks in terms of standard deviation, an analysis of higher-order moments reveals a trade-off. In the majority of scenarios, the adjusted portfolios exhibit higher excess kurtosis than the benchmark $\mathbf{x}^0$, indicating a distribution with fatter tails relative to its variance. This increase in kurtosis is likely a mathematical artifact of the significant reduction in the denominator (variance) achieved by the strategies, rather than an increase in absolute tail risk, as confirmed by the improvement of the VaR and C-VaR. The behavior of skewness is more heterogeneous: while it deteriorates (becomes more negative)

Table 2 Out-of-sample performance indicators of the portfolios: the benchmark $\mathbf{x}^0$ is the equally weighted portfolio (EW), the two adjusted portfolios $\mathbf{p}^1$ and $\mathbf{p}^*$ consider the first and a variable number (stopping the recursion when $\alpha_i < 0.01$) of pseudo-principal portfolios, respectively

n		$w = 5n, d = 5$			$w = 10n, d = 5$			$w = 5n, d = 20$			$w = 10n, d = 20$		
x^0		x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*
5	mean (%)	0.0378	0.0388	0.0392	0.0389	0.0399	0.0383	0.0383	0.0396	0.0397	0.0389	0.0414	0.0406
EW	stdev (%)	1.1967	1.0983	1.0925	1.1957	1.0915	1.0872	1.1965	1.1144	1.1062	1.1957	1.0892	1.0840
	$\text{VaR}_{5\%}$ (%)	1.8095	1.6592	1.6467	1.7930	1.6555	1.6348	1.8020	1.6682	1.6515	1.7930	1.6529	1.6529
	$\text{C-VaR}_{5\%}$ (%)	2.8582	2.6246	2.6029	2.8543	2.6105	2.6098	2.8575	2.6537	2.6350	2.8543	2.5977	2.5889
	skewness	−0.2069	−0.2439	−0.2205	−0.2065	−0.2124	−0.2303	−0.2070	−0.1917	−0.1854	−0.2065	−0.1649	−0.1676
	kurtosis	12.5403	12.8252	13.2166	12.6177	13.4614	13.4540	12.5542	14.6988	14.6858	12.6177	13.9604	13.9895
	Sharpe	0.0264	0.0296	0.0301	0.0274	0.0309	0.0295	0.0267	0.0299	0.0303	0.0274	0.0323	0.0318
	freq smaller σ	–	0.9797	0.9829	–	0.9803	0.9717	–	0.9761	0.9977	–	0.9703	0.9590
	freq larger SR	–	0.4857	0.4986	–	0.5704	0.4949	–	0.6677	0.6554	–	0.6225	0.5940
10	mean (%)	0.0397	0.0423	0.0417	0.0388	0.0385	0.0371	0.0397	0.0423	0.0410	0.0385	0.0396	0.0375
EW	stdev (%)	1.1883	1.0648	1.0591	1.1855	1.0677	1.0622	1.1883	1.0901	1.0859	1.1859	1.0744	1.0689
	$\text{VaR}_{5\%}$ (%)	1.7775	1.5717	1.5637	1.7820	1.5713	1.5655	1.7775	1.5859	1.5843	1.7825	1.5534	1.5682
	$\text{C-VaR}_{5\%}$ (%)	2.8482	2.5393	2.5319	2.8437	2.5566	2.5475	2.8482	2.5923	2.5863	2.8474	2.5589	2.5469
	skewness	−0.2417	−0.0573	−0.0626	−0.2447	−0.1593	−0.1500	−0.2417	−0.1742	−0.1759	−0.2451	−0.0904	−0.0813
	kurtosis	13.8771	15.2688	15.5050	13.9990	17.8520	18.1467	13.8771	19.6027	19.8418	14.0039	17.7904	18.0476
	Sharpe	0.0282	0.0339	0.0335	0.0276	0.0304	0.0292	0.0282	0.0331	0.0320	0.0273	0.0312	0.0294
	freq smaller σ	–	1.0	1.0	–	0.9993	1.0	–	0.9503	0.9556	–	1.0	1.0
	freq larger SR	–	0.7261	0.7017	–	0.6541	0.6189	–	0.6694	0.6575	–	0.6263	0.6158
	turnover	0.0128	0.2997	0.3436	0.0126	0.2416	0.2988	0.0267	0.3765	0.4046	0.0261	0.2700	0.3001
12	mean (%)	0.0384	0.0398	0.0380	0.0387	0.0401	0.0386	0.0379	0.0401	0.0392	0.0384	0.0368	0.0357
EW	stdev (%)	1.1978	1.0770	1.0716	1.1980	1.0683	1.0641	1.1985	1.0814	1.0765	1.1988	1.0764	1.0733
	$\text{VaR}_{5\%}$ (%)	1.7983	1.5955	1.5895	1.8000	1.5564	1.5537	1.7988	1.5937	1.5914	1.8037	1.5731	1.5692
	$\text{C-VaR}_{5\%}$ (%)	2.8818	2.5800	2.5701	2.8809	2.5541	2.5521	2.8856	2.5788	2.5710	2.8846	2.5771	2.5763
	skewness	−0.2672	−0.1374	−0.1296	−0.2575	−0.1723	−0.1758	−0.2662	−0.1519	−0.1463	−0.2568	−0.1413	−0.1412
	kurtosis	13.6190	16.3468	16.5931	13.6688	16.3871	16.6245	13.6072	16.2998	16.5040	13.6586	16.4613	16.6365
	Sharpe	0.0269	0.0312	0.0297	0.0273	0.0318	0.0306	0.0265	0.0313	0.0307	0.0270	0.0285	0.0276

Table 2 (continued)

n		$w = 5n, d = 5$			$w = 10n, d = 5$			$w = 5n, d = 20$			$w = 10n, d = 20$		
x^0		x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*
	freq smaller σ	–	1.0	1.0	–	1.0	1.0	–	0.9943	1.0	–	1.0	1.0
	freq larger SR	–	0.6983	0.6356	–	0.7166	0.6978	–	0.6604	0.6315	–	0.5816	0.5562
	turnover	0.0124	0.2601	0.2926	0.0123	0.2259	0.2664	0.0255	0.3285	0.3523	0.0252	0.2363	0.2461
17	mean (%)	0.0432	0.0418	0.0419	0.0426	0.0438	0.0432	0.0432	0.0384	0.0383	0.0426	0.0416	0.0403
EW	stdev (%)	1.2947	1.2001	1.1913	1.3004	1.2059	1.1967	1.2950	1.2078	1.2019	1.3004	1.2155	1.2093
	$\text{VaR}_{5\%}$ (%)	1.9284	1.7594	1.7411	1.9329	1.7719	1.7527	1.9285	1.7539	1.7478	1.9329	1.7579	1.7449
	$\text{C-VaR}_{5\%}$ (%)	3.0933	2.8978	2.8768	3.1096	2.9017	2.8787	3.0933	2.9143	2.8988	3.1096	2.9220	2.9093
	skewness	−0.2899	−0.2228	−0.2224	−0.2902	−0.2756	−0.2646	−0.2900	−0.2423	−0.2390	−0.2902	−0.2393	−0.2351
	kurtosis	12.7737	15.4727	15.5091	12.7332	14.6630	14.9674	12.7718	16.2793	16.5071	12.7332	16.4037	16.5391
	Sharpe	0.0286	0.0297	0.0300	0.0281	0.0313	0.0311	0.0286	0.0267	0.0268	0.0281	0.0293	0.0284
	freq smaller σ	–	0.9932	0.9953	–	0.9931	0.9973	–	0.9777	0.9950	–	0.9521	0.9539
	freq larger SR	–	0.6097	0.6088	–	0.6603	0.6398	–	0.4997	0.4808	–	0.5826	0.5526
	turnover	0.0144	0.2758	0.3157	0.0143	0.2360	0.2769	0.0287	0.3208	0.3402	0.0287	0.2427	0.2699
30	mean (%)	0.0424	0.0395	0.0401	0.0432	0.0417	0.0409	0.0424	0.0384	0.0386	0.0439	0.0408	0.0408
EW	stdev (%)	1.2756	1.1281	1.1219	1.2822	1.1496	1.1411	1.2756	1.1412	1.1373	1.2819	1.1492	1.1406
	$\text{VaR}_{5\%}$ (%)	1.8880	1.6848	1.6532	1.8983	1.7068	1.6839	1.8880	1.6885	1.6833	1.8935	1.6847	1.6792
	$\text{C-VaR}_{5\%}$ (%)	3.0626	2.7418	2.7212	3.0827	2.7943	2.7746	3.0626	2.7677	2.7554	3.0859	2.8009	2.7772
	skewness	−0.3050	−0.3992	−0.3947	−0.3082	−0.3579	−0.3666	−0.3050	−0.3526	−0.3547	−0.3092	−0.3697	−0.3746
	kurtosis	12.5096	13.5028	13.6140	12.5503	13.8082	13.9147	12.5096	14.9806	15.1134	12.5831	14.0733	14.1237
	Sharpe	0.0285	0.0297	0.0304	0.0290	0.0310	0.0306	0.0285	0.0284	0.0287	0.0296	0.0303	0.0306
	freq smaller σ	–	1.0	1.0	–	1.0	1.0	–	1.0	1.0	–	0.9996	1.0
	freq larger SR	–	0.5589	0.5608	–	0.5964	0.5930	–	0.5486	0.5405	–	0.5729	0.5538
	turnover	0.0162	0.1412	0.1601	0.0158	0.1418	0.1547	0.0326	0.1460	0.1564	0.0319	0.1136	0.1204

The results concern all the combinations of various settings: number of assets $n \in \{5, 10, 12, 17, 30\}$; estimating window length $w \in \{5n, 10n\}$, holding period $d \in \{5, 20\}$ days. The reported indicators are the average, standard deviation, $\text{VaR}_{\{5\%\}}$, $\text{C-VaR}_{\{5\%\}}$, skewness, kurtosis, Sharpe ratio of the returns; moreover, the table shows the frequencies of smaller standard deviation and larger Sharpe ratio, with respect to the benchmark, computed on a year-length sliding window; finally, the turnover of the portfolio is presented

Table 3 Out-of-sample performance indicators of the portfolios: the benchmark $\mathbf{x}^0$ is the equal risk contribution portfolio (ERC), the two adjusted portfolios $\mathbf{p}^1$ and $\mathbf{p}^*$ consider the first and a variable number (stopping the recursion when $\alpha_i < 0.01$) of pseudo-principal portfolios, respectively

n		$w=5n, d=5$			$w=10n, d=5$			$w=5n, d=20$			$w=10n, d=20$		
x^0		x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*
5	mean (%)	0.0388	0.0389	0.0389	0.0402	0.0401	0.0398	0.0396	0.0400	0.0408	0.0404	0.0414	0.0411
ERC	stdev (%)	1.1507	1.0868	1.0821	1.1513	1.0823	1.0785	1.1524	1.1021	1.0965	1.1532	1.0829	1.0792
	$\text{VaR}_{5\%}$ (%)	1.7251	1.6464	1.6392	1.7261	1.6383	1.6248	1.7337	1.6524	1.6349	1.7244	1.6475	1.6486
	C-$\text{VaR}_{5\%}$ (%)	2.7507	2.5958	2.5815	2.7495	2.5805	2.5760	2.7540	2.6261	2.6127	2.7533	2.5803	2.5719
	skewness	−0.2211	−0.2328	−0.2274	−0.2200	−0.2022	−0.1858	−0.2219	−0.2000	−0.1996	−0.2120	−0.1676	−0.1669
	kurtosis	13.1043	12.9389	13.1102	13.1497	13.4906	13.6265	13.1002	14.2396	14.2898	13.1954	13.7764	13.7623
	Sharpe	0.0282	0.0300	0.0302	0.0295	0.0313	0.0311	0.0289	0.0306	0.0315	0.0297	0.0325	0.0323
	freq smaller σ	–	0.9535	0.9626	–	0.9708	0.9524	–	0.9526	0.9551	–	0.9649	0.9365
	freq larger SR	–	0.5098	0.4726	–	0.5620	0.5267	–	0.6465	0.6873	–	0.6039	0.5783
	turnover	0.0597	0.3042	0.3738	0.0303	0.2226	0.2991	0.1215	0.4561	0.5322	0.0680	0.3196	0.3838
10	mean (%)	0.0398	0.0413	0.0408	0.0390	0.0380	0.0373	0.0403	0.0420	0.0417	0.0389	0.0386	0.0378
ERC	stdev (%)	1.1213	1.0474	1.0433	1.1224	1.0475	1.0443	1.1240	1.0631	1.0610	1.1241	1.0585	1.0562
	$\text{VaR}_{5\%}$ (%)	1.6437	1.5351	1.5348	1.6614	1.5382	1.5382	1.6510	1.5491	1.5394	1.6588	1.5292	1.5290
	C-$\text{VaR}_{5\%}$ (%)	2.6980	2.5059	2.5007	2.7000	2.5108	2.5045	2.7020	2.5354	2.5307	2.7065	2.5329	2.5263
	skewness	−0.2518	−0.1183	−0.1265	−0.2368	−0.1387	−0.1350	−0.2244	−0.1598	−0.1581	−0.2205	−0.1129	−0.1073
	kurtosis	15.3574	15.8243	15.9303	15.5508	17.3118	17.4766	15.3380	18.2137	18.3400	15.5392	17.8718	17.9632
	Sharpe	0.0299	0.0335	0.0331	0.0293	0.0304	0.0299	0.0303	0.0337	0.0334	0.0291	0.0308	0.0300
	freq smaller σ	–	1.0	1.0	–	1.0	1.0	–	0.9816	0.9902	–	0.9993	0.9998
	freq larger SR	–	0.6736	0.6375	–	0.6431	0.6140	–	0.6717	0.6529	–	0.6285	0.5941
	turnover	0.0449	0.2153	0.2502	0.0245	0.1766	0.2209	0.0988	0.3093	0.3368	0.0535	0.2119	0.2450
12	mean (%)	0.0388	0.0405	0.0393	0.0389	0.0389	0.0385	0.0382	0.0395	0.0385	0.0387	0.0370	0.0362
ERC	stdev (%)	1.1320	1.0603	1.0573	1.1358	1.0549	1.0519	1.1346	1.0650	1.0632	1.1373	1.0601	1.0580
	$\text{VaR}_{5\%}$ (%)	1.6749	1.5566	1.5582	1.7016	1.5344	1.5344	1.6973	1.5579	1.5541	1.7027	1.5591	1.5456
	C-$\text{VaR}_{5\%}$ (%)	2.7316	2.5413	2.5382	2.7357	2.5262	2.5191	2.7378	2.5475	2.5445	2.7403	2.5429	2.5401
	skewness	−0.2764	−0.2029	−0.2050	−0.2525	−0.1625	−0.1601	−0.2551	−0.1442	−0.1421	−0.2376	−0.1537	−0.1528
	kurtosis	14.9585	16.2981	16.4324	14.9672	16.5470	16.6898	14.9005	16.7475	16.8359	14.9461	16.7250	16.8386

Table 3 (continued)

n		$w=5n, d=5$			$w=10n, d=5$			$w=5n, d=20$			$w=10n, d=20$		
x^0		x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*
	Sharpe	0.0288	0.0324	0.0313	0.0289	0.0311	0.0308	0.0282	0.0312	0.0304	0.0287	0.0292	0.0285
	freq smaller σ	–	0.9995	1.0	–	1.0	1.0	–	0.9894	0.9941	–	1.0	1.0
	freq larger SR	–	0.7177	0.6904	–	0.6717	0.6748	–	0.6516	0.6272	–	0.6036	0.5738
	turnover	0.0376	0.1955	0.2268	0.0214	0.1559	0.1872	0.0808	0.2703	0.2999	0.0469	0.1768	0.1961
17	mean (%)	0.0429	0.0406	0.0402	0.0424	0.0442	0.0433	0.0428	0.0392	0.0386	0.0425	0.0431	0.0425
ERC	stdev (%)	1.1980	1.1472	1.1442	1.2065	1.1559	1.1507	1.1991	1.1535	1.1506	1.2070	1.1622	1.1591
	$\text{VaR}_{5\%}$ (%)	1.7637	1.6948	1.6913	1.7896	1.6947	1.6788	1.7630	1.6826	1.6796	1.7878	1.6890	1.6836
	$\text{C-VaR}_{5\%}$ (%)	2.8829	2.7768	2.7695	2.9027	2.7825	2.7704	2.8820	2.7832	2.7753	2.9016	2.7922	2.7831
	skewness	−0.3065	−0.2651	−0.2612	−0.2952	−0.2458	−0.2346	−0.2861	−0.2490	−0.2490	−0.2831	−0.2283	−0.2221
	kurtosis	14.3314	15.5322	15.6554	14.1089	15.4996	15.5559	14.2446	16.3081	16.4445	14.0120	16.4257	16.5549
	Sharpe	0.0306	0.0300	0.0298	0.0302	0.0330	0.0324	0.0306	0.0287	0.0282	0.0303	0.0319	0.0315
	freq smaller σ	–	0.9703	0.9744	–	0.9526	0.9558	–	0.9540	0.9558	–	0.9519	0.9523
	freq larger SR	–	0.5816	0.5594	–	0.6954	0.6542	–	0.4877	0.4651	–	0.6239	0.6076
	turnover	0.0325	0.1788	0.2087	0.0204	0.1576	0.1953	0.0696	0.2337	0.2596	0.0452	0.1757	0.2042
30	mean(%)	0.0417	0.0401	0.0398	0.0425	0.0419	0.0416	0.0415	0.0395	0.0396	0.0432	0.0415	0.0414
ERC	stdev(%)	1.1732	1.0937	1.0889	1.1838	1.1082	1.1027	1.1740	1.1024	1.0982	1.1841	1.1130	1.1079
	$\text{VaR}_{5\%}$ (%)	1.7350	1.6152	1.6158	1.7411	1.6384	1.6233	1.7365	1.6199	1.6267	1.7363	1.6379	1.6257
	$\text{C-VaR}_{5\%}$ (%)	2.8302	2.6514	2.6405	2.8557	2.6863	2.6717	2.8305	2.6662	2.6568	2.8600	2.7012	2.6878
	skewness	−0.3065	−0.2651	−0.2612	−0.2952	−0.2458	−0.2346	−0.2861	−0.2490	−0.2490	−0.2831	−0.2283	−0.2221
	kurtosis	14.3314	15.5322	15.6554	14.1089	15.4996	15.5559	14.2446	16.3081	16.4445	14.0120	16.4257	16.5549
	Sharpe	0.0304	0.0311	0.0310	0.0308	0.0324	0.0323	0.0303	0.0304	0.0306	0.0315	0.0320	0.0320
	freqsmallers	–	1.0	1.0	–	1.0	1.0	–	1.0	1.0	–	1.0	1.0
	freqlargerSR	–	0.5578	0.5702	–	0.6072	0.6014	–	0.5664	0.5708	–	0.5793	0.5742
	turnover	0.0246	0.0863	0.1051	0.0182	0.0826	0.0995	0.0532	0.1129	0.1292	0.0389	0.0853	0.0971

The results concern all the combinations of various settings: number of assets $n \in \{5, 10, 12, 17, 30\}$; estimating window length $w \in \{5n, 10n\}$, holding period $d \in \{5, 20\}$ days. The reported indicators are the average, standard deviation, $\text{VaR}_{\{5\%\}}$, $\text{C-VaR}_{\{5\%\}}$, skewness, kurtosis, Sharpe ratio of the returns; moreover, the table shows the frequencies of smaller standard deviation and larger Sharpe ratio, with respect to the benchmark, computed on a year-length sliding window; finally, the turnover of the portfolio is presented

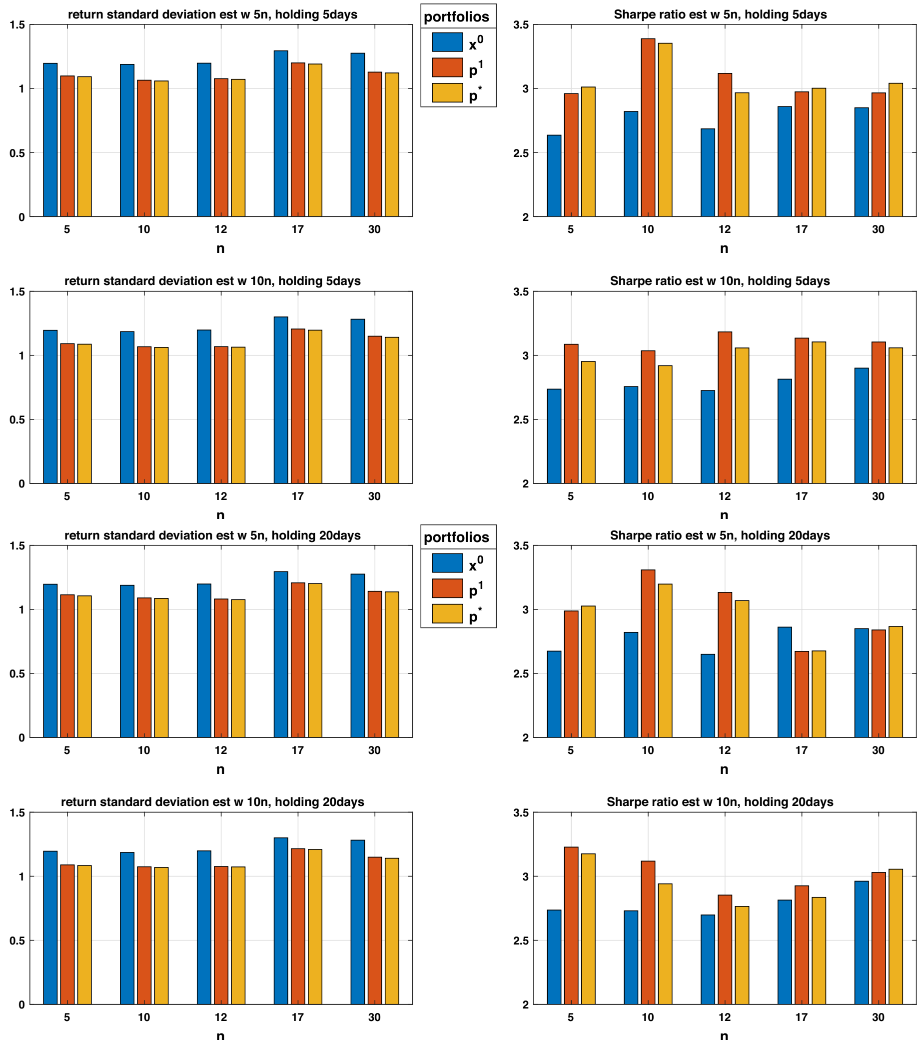

Fig. 6 Standard deviation (left) and Sharpe ratio (right) of the benchmark portfolio $\mathbf{x}^0$ (blue), and the two adjusted portfolios $\mathbf{p}^1$ (orange) and $\mathbf{p}^*$ (yellow), when $\mathbf{x}^0$ is the equally weighted portfolio. Bars are grouped by the number of sectors (indicated on the horizontal axis). Each plot concerns a given combination of estimation window w and holding period d, as indicated in its title

in specific EW configurations (e.g., Table 2), $n = 30$), the proposed strategies frequently demonstrate improved asymmetry (less negative skewness) compared to the ERC benchmark (see Table 3), particularly for $n \geq 10$.

In Figs. 6 and 7 (left column), it is evident that there is a substantial jump when moving from x^0 to p^1, while the difference between p^1 and p^* is usually small. This indicates that the contribution of subsequent pseudo-principal portfolios to risk reduction quickly

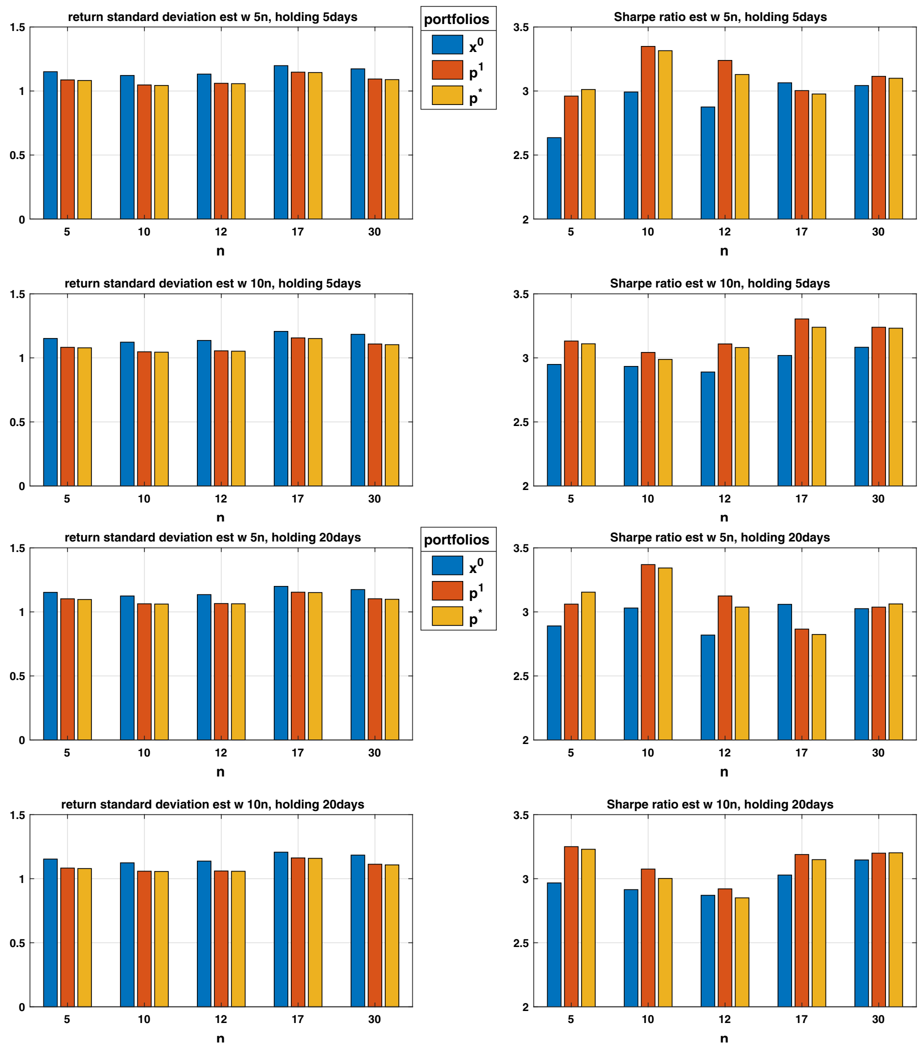

Fig. 7 Standard deviation (left) and Sharpe ratio (right) of the benchmark portfolio $\mathbf{x}^0$ (blue), and the two adjusted portfolios $\mathbf{p}^1$ (orange) and $\mathbf{p}^*$ (yellow), when $\mathbf{x}^0$ is the equal risk contribution portfolio. Bars are grouped by the number of sectors (indicated on the horizontal axis). Each plot concerns a given combination of estimation window w and holding period d, as indicated in its title

becomes negligible. Moreover, comparing $\mathbf{p}^1$ with $\mathbf{p}^*$, adding further pseudo-principal portfolios appears to only marginally reduce risk while negatively affecting returns, often resulting in lower Sharpe ratios. This evidence is in line with standard applications of principal component analysis. When applying PCA or similar techniques, such as the one described in the present paper, the main practical issue concerns the choice of how many components should be considered, since the contribution of each additional component is, by construction, smaller than that of the previous one. Within our experiment and for the specific benchmarks used in the application, the empirical evidence strongly supports combining the benchmark with only the first pseudo-principal portfolio. It appears that the majority of the useful information is captured by the first pseudo-principal portfolio. Tables 2 and 3 also report the turnovers of the strategies, to provide a direct indication of the potential impact of transaction costs. Let us emphasize that transaction costs mainly affect expected returns. Naturally, the magnitude of these costs is a key factor when choosing between a passive and an active investment strategy. As expected, the level of turnover is small for the two selected benchmarks, which represent very classical, stable, and almost passive investments, while it is significantly higher for $\mathbf{p}^0$ and $\mathbf{p}^1$. Owing to the generality of the proposed approach, if an active investment strategy is chosen as the benchmark x^0, for example, an optimization-based strategy, then transaction costs would be comparable for both the benchmark strategy and the combined portfolios. To provide a more precise picture of the effects of transaction costs, Table 4 presents an example in which proportional transaction costs of 10 basis points are applied to the proposed strategies, in the case of EW starting portfolios.[7] Table 4 confirms the limited impact of transaction costs on risk reduction.

One empirical consideration emerging from our experiment is that the proposed approach seems to work better for portfolios with a small number of constituents. This may depend on the allocation rule itself or on the data; we recall that, although standard in this context, the assets considered are portfolios themselves. As a consequence, when n is large, the EW and the ERC portfolios, which by construction include all available assets, are already strongly diversified. A further possible explanation of the empirical evidence is that portfolios with a large number of assets are already well diversified simply because of their size. Therefore, the proposed procedures yield weaker results—although still consistently present—in terms of risk reduction. Conversely, when the number of assets n is small, the diversification effect due to portfolio size is less pronounced, leaving more room to increase diversification through the proposed geometric procedure.

3.3.1 Computation time

The computation time, as reported in Table 5, is reasonable and does not constitute a significant barrier to practical applications. Moreover, it should be noted that this method is intended for strategic or medium-frequency portfolio allocation performed offline, rather than for high-frequency trading. Therefore, the computational requirements are appropriate for its intended purpose.[8] However, it is worth noting that the computation time increases more rapidly than the problem size, with the exception of the CPU time for the entire back-test, which scales approximately linearly with size. Interestingly, for portfolios involving 10 to 12 assets, the computation time remains relatively constant. We hypothesize that this behavior stems from inherent variability in computational performance.

[7] For the sake of conciseness, we omit other cases (different levels of transaction costs and different benchmarks), which are available upon request.

[8] In some cases, the CPU time is larger than the clock time because MATLAB sums CPU time across parallel threads.

Table 4 Out-of-sample performance indicators of the portfolios, in case of transaction costs of 10 basis points: the benchmark $\mathbf{x}^0$ is the equally weighted portfolio (EW), the two adjusted portfolios $\mathbf{p}^1$ and $\mathbf{p}^*$ consider the first and a variable number (stopping the recursion when $\alpha_i < 0.01$) of pseudo-principal portfolios, respectively

n		$w = 5n, d = 5$			$w = 10n, d = 5$			$w = 5n, d = 20$			$w = 10n, d = 20$		
$\mathbf{x}^0$		x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*	x^0	p^1	p^*
5	mean (%)	0.0376	0.0317	0.0304	0.0388	0.0344	0.0311	0.0382	0.0371	0.0369	0.0389	0.0396	0.0385
EW	stdev (%)	1.1966	1.0983	1.0924	1.1957	1.0914	1.0873	1.1965	1.1143	1.1061	1.1957	1.0892	1.0840
	$\text{VaR}_{5\%}$ (%)	1.8097	1.6633	1.6573	1.7934	1.6676	1.6555	1.8020	1.6698	1.6577	1.7930	1.6555	1.6564
	$\text{C-VaR}_{5\%}$ (%)	2.8584	2.6312	2.6104	2.8546	2.6159	2.6168	2.8576	2.6561	2.6377	2.8544	2.5998	2.5909
	Sharpe	0.0262	0.0231	0.0221	0.0272	0.0259	0.0229	0.0267	0.0277	0.0277	0.0273	0.0307	0.0298
10	mean (%)	0.0395	0.0363	0.0349	0.0386	0.0337	0.0312	0.0396	0.0404	0.0389	0.0383	0.0383	0.0360
EW	stdev (%)	1.1883	1.0646	1.0589	1.1855	1.0676	1.0619	1.1883	1.0900	1.0859	1.1859	1.0743	1.0688
	$\text{VaR}_{5\%}$ (%)	1.7775	1.5849	1.5714	1.7820	1.5726	1.5659	1.7775	1.5869	1.5846	1.7825	1.5581	1.5682
	$\text{C-VaR}_{5\%}$ (%)	2.8485	2.5461	2.5388	2.8441	2.5612	2.5524	2.8484	2.5941	2.5883	2.8476	2.5601	2.5480
	Sharpe	0.0280	0.0283	0.0271	0.0274	0.0259	0.0236	0.0281	0.0314	0.0301	0.0272	0.0299	0.0280
12	mean (%)	0.0381	0.0346	0.0322	0.0385	0.0356	0.0333	0.0378	0.0384	0.0375	0.0383	0.0356	0.0345
EW	stdev (%)	1.1978	1.0769	1.0715	1.1980	1.0681	1.0638	1.1985	1.0815	1.0765	1.1987	1.0764	1.0732
	$\text{VaR}_{5\%}$ (%)	1.7983	1.6006	1.5943	1.8000	1.5623	1.5602	1.7988	1.5937	1.5914	1.8037	1.5781	1.5731
	$\text{C-VaR}_{5\%}$ (%)	2.8822	2.5848	2.5753	2.8812	2.5571	2.5553	2.8857	2.5803	2.5724	2.8848	2.5781	2.5774
	Sharpe	0.0267	0.0264	0.0242	0.0271	0.0276	0.0256	0.0264	0.0298	0.0291	0.0269	0.0274	0.0265
17	mean (%)	0.0429	0.0363	0.0356	0.0423	0.0391	0.0376	0.0431	0.0368	0.0366	0.0424	0.0404	0.0389
EW	stdev (%)	1.2947	1.2002	1.1914	1.3004	1.2058	1.1966	1.2950	1.2077	1.2018	1.3004	1.2156	1.2094
	$\text{VaR}_{5\%}$ (%)	1.9284	1.7809	1.7612	1.9329	1.7730	1.7599	1.9285	1.7566	1.7508	1.9329	1.7604	1.7487
	$\text{C-VaR}_{5\%}$ (%)	3.0936	2.9039	2.8831	3.1099	2.9063	2.8835	3.0934	2.9160	2.9004	3.1097	2.9234	2.9107
	Sharpe	0.0284	0.0252	0.0247	0.0279	0.0274	0.0264	0.0285	0.0254	0.0254	0.0280	0.0283	0.0272
30	mean (%)	0.0421	0.0367	0.0369	0.0428	0.0388	0.0378	0.0422	0.0377	0.0379	0.0437	0.0402	0.0402
EW	stdev (%)	1.2756	1.1279	1.1217	1.2822	1.1496	1.1410	1.2756	1.1411	1.1372	1.2819	1.1492	1.1406
	$\text{VaR}_{5\%}$ (%)	1.8885	1.6848	1.6564	1.8983	1.7068	1.6839	1.8880	1.6885	1.6833	1.8935	1.6847	1.6792
	$\text{C-VaR}_{5\%}$ (%)	3.0630	2.7442	2.7238	3.0831	2.7971	2.7774	3.0628	2.7684	2.7562	3.0860	2.8019	2.7782
	Sharpe	0.0283	0.0272	0.0276	0.0288	0.0286	0.0279	0.0284	0.0278	0.0280	0.0295	0.0298	0.0300

The results concern all the combinations of various settings: number of assets $n \in \{5, 10, 12, 17, 30\}$; estimating window length $w \in \{5n, 10n\}$, holding period $d \in \{5, 20\}$ days. The reported indicators are the average, standard deviation, $\text{VaR}_{\{5\%\}}$, $\text{C-VaR}_{\{5\%\}}$, Sharpe ratio of the returns

Table 5 Computation times in seconds measured on a PC equipped with an AMD Ryzen 5 5500U with 16 GB RAM, Debian 13, and running MATLAB R2023a, parallel computing with 6 processes. No GPU was used

Portfolio size	Single run		Complete backtest	
	clock time	cpu time	Clock time	cpu time
5	0.0607	0.2200	3.7417	0.4520
10	0.1658	0.2080	12.2586	1.0030
12	0.2105	0.2860	12.6826	0.7940
17	0.4132	0.6070	41.4097	1.6720
30	3.7100	12.5620	112.4984	2.7790

For each number of assets, we report the time for a single run (computation of pseudo principal portfolios) and for the complete backtest using the settings described in Sect. 3.3. To reduce the impact of the computation time variability, the reported values are the average times over 10 repetitions of the same task. The CPU times are larger because the time for each parallel process is aggregated

4 Conclusions

In this paper, we define the pseudo-principal portfolios as a suitable generalization of principal portfolios, which can be calculated through principal component analysis. The advantage of our definition is that, given any benchmark portfolio, the pseudo-principal portfolios constitute an orthogonal basis of the return space. This geometric property allows us to define a general strategy that combines the benchmark portfolio with its pseudo-principal portfolios in order to exploit orthogonality and, consequently, achieve risk reduction through diversification. The theoretical section of the paper presents the formalization of the idea, the proofs of several relevant properties, and some economic interpretations that translate the geometric intuition underlying the approach. The empirical evidence from experiments on real financial data supports the effectiveness of the proposed strategy in systematically reducing the risk of a benchmark portfolio for all the considered risk measures: standard deviation, Value at Risk (VaR), and expected shortfall. The results in terms of return improvement are ambiguous, but this outcome was predictable, since a diversification-based strategy is expected to primarily affect the risk dimension. Nevertheless, in the majority of the out-of-sample experiments, the composed portfolios outperform the benchmark also in terms of risk-adjusted performance. In our opinion, this research indirectly highlights one of the possible reasons behind the poor out-of-sample performance of the mean–variance approach. As shown in the paper, only the first few pseudo-principal portfolios contain useful information to achieve effective risk reduction in practice, while the last pseudo-principal portfolios mainly capture noise. Classical mean–variance optimization, instead, relies on the entire covariance matrix, thereby incorporating all the noise present in the data into the optimization process. Future research will focus on an extensive application of the procedure across diverse benchmarks, while more thoroughly incorporating transaction costs—which are currently only briefly outlined—to assess the overall reliability of the proposed method. To this extent, the comparison between different rebalancing policies, including a passive allocation, could be interesting. It is also worth noting that, from an operational perspective, the computational times do not pose a significant barrier to the application of the proposed procedures. Future research will investigate the applicability of the proposed approach to multi-asset class allocation problems and dynamic allocation models. Furthermore, the possibility of mechanically implementing the allocation procedure suggests potential interest in integration within robo-advisory and smart beta frameworks.

Author contributions The authors contributed equally to the research and the manuscript preparation.

Funding Open access funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement. This research received no external funding.

Data availability The data used for this research can be downloaded for free at the indicated website. The code used for the analysis is available upon request to the authors.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Ackermann, F., Pohl, W., Schmedders, K.: Optimal and naive diversification in currency markets. Manage. Sci. **63**(10), 3347–3360 (2017)

Antar, M.: Blue gold dynamics – a multi-scale correlation analysis of the behavior of water-themed etfs. Rev. Acc. Financ. 1–20 (2026). https://doi.org/10.1108/RAF-04-2025-0145

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Financ. **9**(3), 203–228 (1999)

Basak, S., Shapiro, A., Teplá, L.: Risk management with benchmarking. Manage. Sci. **52**(4), 542–557 (2006)

Bessler, W., Opfer, H., Wolff, D.: Multi-asset portfolio optimization and out-of-sample performance: an evaluation of black-litterman, mean-variance, and naïve diversification approaches. Eur. J. Finance **23**(1), 1–30 (2017)

Best, M.J., Hlouskova, J.: Quadratic programming with transaction costs. Comput. Operat. Res. **35**(1), 18–33 (2008)

Bonaccolto, G., Caporin, M.: On the ordering of dynamic principal components and the implications for portfolio analysis. Ann. Oper. Res. 1–28 (2024)

DeMiguel, V., Garlappi, L., Uppal, R.: Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? Rev. Financ. Stud. **22**(5), 1915–1953 (2009)

Dichtl, H., Drobetz, W., Wambach, M.: Where is the value added of rebalancing? a systematic comparison of alternative rebalancing strategies. Fin. Markets. Portfolio Mgmt. **28**(3), 209–231 (2014)

Dichtl, H., Drobetz, W., Wambach, M.: Testing rebalancing strategies for stock-bond portfolios across different asset allocations. Appl. Econ. **48**(9), 772–788 (2016)

Donohue, C., Yip, K.: Optimal portfolio rebalancing with transaction costs. J. Portf. Manag. **29**(4), 49 (2003)

Fugazza, C., Guidolin, M., Nicodano, G.: Equally weighted versus long-run optimal portfolios. Eur. Financ. Manag. **21**(4), 742–789 (2015)

Giglio, S., Xiu, D.: Asset pricing with omitted factors. J. Polit. Econ. **129**(7), 1947–1990 (2021)

Hanke, B., Keswani, A., Quigley, G., Stolin, D., Zagonov, M.: The equal-weight tilt in managed portfolios. Econ. Lett. **182**, 59–63 (2019)

Hirschberger, M., Qi, Y., Steuer, R.E.: Large-scale mv efficient frontier computation via a procedure of parametric quadratic programming. Eur. J. Oper. Res. **204**(3), 581–588 (2010)

Imajo, K., Minami, K., Ito, K., Nakagawa, K.: Deep portfolio optimization via distributional prediction of residual factors. In: Proceedings of the aaai conference on artificial intelligence, 35: 213–222 (2021)

Jacobs, B.I., Levy, K.N.: Leverage aversion, efficient frontiers, and the efficient region. J. Portf. Manag. **39**(3), 54–64 (2013)

Jiang, C., Du, J., An, Y.: Combining the minimum-variance and equally-weighted portfolios: Can portfolio performance be improved? Econ. Model. **80**, 260–274 (2019)
Jolliffe, I.: Principal component analysis. International encyclopedia of statistical science, pp. 1094–1096. Springer (2011)
Kelly, B., Malamud, S., Pedersen, L.H.: Principal portfolios. J. Financ. **78**(1), 347–387 (2023)
Kirby, C., Ostdiek, B.: It's all in the timing: simple active portfolio strategies that outperform naive diversification. J. Financ. Quant. Anal. **47**(2), 437–467 (2012)
Kritzman, M., Page, S., Turkington, D.: In defense of optimization: the fallacy of 1/n. Financ. Anal. J. **66**(2), 31–39 (2010)
Lassance, N., DeMiguel, V., Vrins, F.: Optimal portfolio diversification via independent component analysis. Oper. Res. **70**(1), 55–72 (2022)
Lee, W.: Risk-based asset allocation: a new answer to an old question? Inst. Investor J. Umbrella **37**(4), 11–28 (2011)
Liu, Y., Tsyvinski, A., Wu, X.: Common risk factors in cryptocurrency. J. Financ. **77**(2), 1133–1177 (2022)
Lohre, H., Opfer, H., Orszag, G.: Diversifying risk parity. J. Risk **16**(5), 53–79 (2014)
Maillard, S., Roncalli, T., Teïletche, J.: The properties of equally weighted risk contribution portfolios. J. Portf. Manag. **36**(4), 60–70 (2010)
Malladi, R., Fabozzi, F.J.: Equal-weighted strategy: Why it outperforms value-weighted strategies? theory and evidence. J. Asset Manag. **18**, 188–208 (2017)
Markowitz, H.: Portfolio selection. J. Finance **7**(1), 77–91 (1952)
Meucci, A.: Managing diversification. Risk, pp. 74–79. (2009)
Partovi, M.H., Caputo, M.: Principal portfolios: recasting the efficient frontier. Econ. Bull. **7**(3), 1–10 (2004)
Peres-Neto, P.R., Jackson, D.A., Somers, K.M.: How many principal components? stopping rules for determining the number of non-trivial axes revisited. Comput. Stat. Data Anal. **49**(4), 974–997 (2005)
Pflug, G.C., Pichler, A., Wozabal, D.: The 1/n investment strategy is optimal under high model ambiguity. J. Bank. Finance **36**(2), 410–417 (2012)
Strongin, S., Petsch, M., Sharenow, G.: Beating benchmarks. J. Portf. Manag. **26**(4), 11–27 (2000)
Sun, W., Fan, A., Chen, L., Schouwenaars, T., Albota, M.A.: Optimal rebalancing for institutional portfolios. J. Portf. Manag. **32**(2), 33 (2006)