Improving portfolios global performance using a cleaned and robust covariance matrix estimate

Emmanuelle Jay, Thibault Soler, Eugénie Terreaux, Jean-Philippe Ovarlez, Frédéric Pascal, Philippe de Peretti, Christophe Chorro

► To cite this version:

Emmanuelle Jay, Thibault Soler, Eugénie Terreaux, Jean-Philippe Ovarlez, Frédéric Pascal, et al.. Improving portfolios global performance using a cleaned and robust covariance matrix estimate. 2019. halshs-02354596

HAL Id: halshs-02354596
https://shs.hal.science/halshs-02354596

Submitted on 7 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

CES
Centre d'Économie de la Sorbonne
UMR 8174

Improving portfolios global performance using a cleaned and robust covariance matrix estimate

Emmanuelle JAY, Thibault SOLER, Eugénie TERREAUX,
Jean-Philippe OVARLEZ, Frédéric PASCAL,
Philippe De PERETTI, Christophe CHORRO

2019.22

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
https://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Improving portfolios global performance using a cleaned and robust covariance matrix estimate

Emmanuelle Jay · Thibault Soler · Eugénie Terreaux · Jean-Philippe Ovarlez · Frédéric Pascal · Philippe De Peretti · Christophe Chorro

Abstract This paper presents how the most recent improvements made on covariance matrix estimation and model order selection can be applied to the portfolio optimization problem. The particular case of the Maximum Variety Portfolio is treated but the same improvements apply also in the other optimization problems such as the Minimum Variance Portfolio. We assume that the most important information (or the latent factors) are embedded in correlated Elliptical Symmetric noise extending classical Gaussian assumptions. We propose here to focus on a recent method of model order selection allowing to efficiently estimate the subspace of main factors describing the market. This non-standard model order selection problem is solved through Random Matrix Theory and robust covariance matrix estimation. Moreover we extend the method to non-homogeneous assets returns.The proposed procedure will be explained through synthetic data and be applied and compared with standard techniques on real market data showing promising improvements.

Keywords Robust Covariance Matrix Estimation · Model Order Selection · Random Matrix Theory · Portfolio Optimization · Financial Time Series · Multi-Factor Model · Elliptical Symmetric Noise · Maximum Variety Portfolio

JEL: C5, G11

E. Jay (a), (b) - E-mail: ejay@quanted.fr ·
T. Soler (b), (c) - E-mail: tsoler@fideas.fr ·
E. Terreaux (d) - E-mail: eugenie.terreaux@onera.fr ·
J.P. Ovarlez (d) - E-mail: jean-philippe.ovarlez@onera.fr ·
F. Pascal (e) - E-mail: frederic.pascal@centralesupelec.fr ·
P. De Peretti (c) - E-mail: Philippe.de-Peretti@univ-paris1.fr ·
C. Chorro (c) - E-mail: christophe.chorro@univ-paris1.fr ·
(a) Quanted & Europlace Institute of Finance, Palais Brongniart, 28 place de la Bourse, 75002 Paris, France
(b) Fideas Capital, 21 avenue de l'Opéra, 75001 Paris, France
(c) Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, MSE, 106-112 bd de l'Hôpital, 75647 Paris Cedex 13, France
(d) DEMR, ONERA, Université Paris-Saclay, F-91123, Palaiseau, France
(e) L2S, CentraleSupélec, Université Paris-Saclay, F-91190 Gif-sur-Yvette, France

1 Introduction

Portfolio allocation is often associated with the mean-variance framework fathered by Markowitz in the 50's [34]. This framework designs the allocation process as an optimization problem where the portfolio weights are such that the expected return of the portfolio is maximized for a given level of portfolio risk. In practice this needs to estimate both expected returns and covariance matrix leading to estimation errors, particularly important for expected returns. This partly explains why many studies concentrate on allocation process relying solely on the covariance estimation such as the Global Minimum Variance Portfolio or the Equally Risk Contribution Portfolio [10], [32].

Another way to reduce the overall risk of a portfolio is to diversify the risks of its assets and to look for the assets weights that maximize a diversification indicator such as the variety (or diversification) ratio [8, 9], only involving the covariance matrix of the assets returns as well.

The frequently used covariance estimator is the Sample Covariance Matrix (SCM), optimal under the Normal assumption. Nevertheless, financial time series of returns might exhibit outliers related to abnormal returns leading to estimation errors larger than expected. The field of robust covariance estimation under non-Gaussian distributions [51], [35] intends to deal with this problem especially when N, the number of samples, is larger than m, the size of the observations vector. When $N < m$, the covariance matrix estimate is not invertible and regularization approaches are required. Some authors have proposed hybrid robust shrinkage covariance matrix estimates [7], [41], [1], building estimators upon Tyler's robust M-estimator [51] and Ledoit-Wolf's shrinkage approach [30].

Recent works [7], [14], [41], [54] based on Random Matrix Theory (RMT) have therefore considered robust estimation in the m, N regime.
In [54], the Global Minimum Variance Portfolio is studied and the authors show that applying an adapted estimation methodology based on the Shrinkage-Tyler M-estimator leads to achieving superior performance over many other competing methods. Another way to mitigate covariance matrix estimation errors is to filter the noisy part of the data. In financial applications, several empirical evidence militate in favour of the existence of multiple sources of risks challenging the CAPM single market factor assumption [47]. Multi-factor models have therefore emerged based either on statistical factors or on observable factors [17, 18, 20, 45], and are designed to capture the effects of the systematic risks borne by the common factors. In this setup, the covariance matrix estimate of the assets depends solely of the systematic part of the risk, as in [17]. Statistical multi-factor models are also very interesting tools. Instead of choosing the factors among many others and from empirical studies, the factors are determined from the assets universe, using statistical methods. Whereas the principal component analysis may fail in distinguishing informative factors from the noisy ones, RMT helps identifying a solution to filter noise as in [26, 27, 43, 44] by correcting the eigenvalues of the covariance matrix, thanks to the upper bound of the Marčenko-Pastur distribution [33]. This method called "Eigenvalue clipping" provides competitive out-of-sample results [3], even though the single market factor still prevails in the described cleaning method that is not completely satisfactory. Other recent works [4, 5, 28] deal with the class of Rotational Invariant Estimators (RIE) that use all of the information on both eigenvectors and eigenvalues of the covariance matrix. The methodology proposed in [4] leads to portfolios having a lower volatility that those obtained when using SCM, Ledoit & Wolf (LW) and Eigenvalue clipping methods.

The application here proposes to mix several approaches: the assets returns are modelled as a multi-factor model embedded in correlated elliptical and symmetric noise and the final covariance estimate will be computed on the "signal only" part of the observations, separable from the "noise part" thanks to the results found in [48–50, 52]. We also extend the results presented in [23] by considering that the assets returns might be non-homogeneously distributed.

The article is constructed as follows: section 2 introduces the selected methods of portfolio allocation for this paper: the Maximum Variety (or VarMax) portfolio and the Minimum Variance (or MinVar) portfolio. Section 3 presents the classical model and its assumptions under consideration. Section 4 explains how to solve the problem jointly with RMT and the robust estimation theory which allow to design a consistent estimate of the number K of informative factors. Section 5 shows some results obtained on experimental financial data highlighting the efficiency of the proposed method with regards to the conventional ones. Conclusion in section 6 closes this paper.

Notations Matrices are in bold and capital, vectors in bold. $Tr(\mathbf{X})$ is the trace of the matrix $\mathbf{X}$. $\|\mathbf{X}\|$ stands for the spectral norm. For any matrix $\mathbf{A}$, $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. For any $m-$vector $\mathbf{x}$, $\mathscr{L} : \mathbf{x} \mapsto \mathscr{L}(\mathbf{x})$ is defined as the $m \times m$ symmetric and real-valued matrix obtained through the Toeplitz operator: $([\mathscr{L}(\mathbf{x})]_{i,j}) = x_{|i-j|+1}$. For any matrix $\mathbf{A}$ of size $m \times m$, $\mathscr{T}(\mathbf{A})$ represents the matrix $\mathscr{L}(\check{\mathbf{a}})$ where $\check{\mathbf{a}}$ is a vector for which each component $\check{\mathbf{a}}_{i,1 \leq i \leq m}$ is the sum of the elements contained in the i-th diagonal of $\mathbf{A}$ divided by m. Then we have : $\check{a}_i = (\sum_{j=i}^{m} a_{j,j-i+1})/m$, with $a_{i,j}$ the element (i,j) of matrix $\mathbf{A}$. The notation "bp" stands for basis point and one basis point is equal to 0.01%.

2 Portfolio allocation

Portfolio allocation is a widely studied problem. Depending on the investment objective, the resulting portfolio allocation differs. In this section two allocation methods are described: the Maximum Variety process and the Global Minimum Variance one. Both of them depend on a single parameter that is the covariance matrix of the asset returns. In practice, the minimum variance portfolio is known to lead to low diversified but performing portfolios whereas the Maximum Variety process leads to well diversified (by construction) but less performing portfolios.

2.1 Maximum Variety (or VarMax) Portfolio

The Maximum Variety (or VarMax) process aims at maximizing the Variety Ratio ($\mathscr{VR}$) of the final portfolio, that is one of the measures allowing to quantify the degree of diversification of a portfolio invested in m assets with proportions $\mathbf{w} = [w_1, \ldots, w_m]^T$. The Variety Ratio ($\mathscr{VR}$) of the portfolio is defined as follows:

$$\mathscr{VR}(\mathbf{w}, \Sigma) = \frac{\mathbf{w}^T \mathbf{s}}{\left(\mathbf{w}^T \Sigma \mathbf{w}\right)^{1/2}}, \tag{1}$$

where $\mathbf{w}$ is the m-vector of weights, w_i representing the allocation in asset i, Σ is the $m \times m$ covariance matrix of the m assets returns and $\mathbf{s}$ is the m-vector of the square roots of the diagonal element of Σ, i.e. $s_i = \sqrt{\Sigma_{ii}}$, representing the standard deviation of the returns of the m assets.

One way to allocate among the assets would be to maximize the above diversification ratio with respect to the weight vector $\mathbf{w}$ to obtain the solution $\mathbf{w}^*_{vr}$, also called the Maximum Diversified Portfolio in [8]:

$$\mathbf{w}^*_{vr} = \underset{\mathbf{w}}{\operatorname{argmax}} \mathscr{V}\mathscr{R}\left(\mathbf{w}, \Sigma\right), \tag{2}$$

under some conditions and constraints on the individual values of $\mathbf{w}$. In the following, we will impose $0 \leq w_i \leq 1\ \forall i \in [1, m]$ and $\sum_{i=1}^{m} w_i = 1$.

The VarMax Portfolio verifies some interesting properties, as described in [9]:

- VarMax is invariant by duplication: if an asset is duplicated in the universe, then VarMax will be unchanged giving half the weight to each duplicated asset,
- VarMax stays unchanged if a positive linear combination of the assets of the universe is added as a new asset,
- any asset of the universe not held in VarMax is more correlated to the portfolio than to any asset of the portfolio. Furthermore, the more diversified a long-only portfolio is, the greater its correlation with VarMax.

VarMax portfolios are often considered as interesting diversifying investments with respect to the other investments. The above last property would therefore suggest that the other portfolios might then be weakly diversified portfolios.

2.2 Minimum Variance (or MinVar) Portfolio

The Global Minimum Variance Portfolio (or GMVP) is obtained by computing the portfolio whose m-vector of weights $\mathbf{w}_{gmvp}$ minimizes the variance of the final portfolio. It can be formulated as a quadratic optimization problem including the linear constraint that the sum of the weights is equal to 1:

$$\min_{\mathbf{w}} \sigma^2(\mathbf{w}, \Sigma) = \min_{\mathbf{w}} \mathbf{w}^T \Sigma \mathbf{w}, \quad \text{s.t. } \mathbf{w}^T \mathbf{1}_m = 1 \tag{3}$$

with $\mathbf{1}_m$ being a m-vector of ones. The solution to (3), when there is no other constraint on the weight values, is then:
$\mathbf{w}_{gmvp} = \dfrac{\Sigma^{-1}\mathbf{1}_m}{\mathbf{1}_m^T \Sigma^{-1} \mathbf{1}_m}$, and the corresponding portfolio variance writes $\sigma^2(\mathbf{w}_{gmvp}, \Sigma) = \dfrac{1}{\mathbf{1}_m^T \Sigma^{-1} \mathbf{1}_m}$.
As for the VarMax portfolio, the covariance matrix needs to be estimated. If we denote $\widehat{\Sigma}$ an estimate of Σ, then we have:

$$\widehat{\mathbf{w}}_{gmvp} = \frac{\widehat{\Sigma}^{-1}\mathbf{1}_m}{\mathbf{1}_m^T \widehat{\Sigma}^{-1} \mathbf{1}_m}.$$

In [55], the authors derive an optimal optimization strategy in order to minimize the realized portfolio variance, under an assumption of spiked structures[1] of both Σ and Σ^{-1}.
In our case, the weights have to be positive, so that the optimal minimum variance portfolio weights cannot be obtained in a closed form expression. We will nevertheless compare several competing methods of covariance matrix estimation in order to get the GMVP.

---

[1] A spiked structure denotes a covariance model where some eigenvalues are located out of the "bulk", like outliers.

To get solutions for (2) and (3), the unknown covariance matrix Σ has to be determined or estimated. Covariance matrix estimation is a challenging problem in portfolio allocation and several methods can apply. The optimization problem is shown to be very sensitive to outliers and to the chosen method of covariance matrix estimation. Apart from the classical Sample Covariance Matrix (SCM) or Minimum Covariance Determinant (MCD, [46]) that is a method robust to outliers, reside subspace methods that aim at separating the signal space from the noise space, using the eigen-decomposition of the SCM. Usually, the main questions are the choice of the covariance matrix estimate as well as the choice of the eigenvalues threshold value. To overcome these drawbacks and to answer these two questions, we propose a robust and quite simple technique based both on the class of the robust M-estimators and the RMT.

3 Model and assumptions

Suppose that our investment universe is composed of m assets characterized at each time t by their returns. Let us denote by $\mathbf{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_N]$ the $m \times N$-matrix containing N observations of the m-vector $\{\mathbf{r}_t\}_{t \in [1,N]}$. We assume next that the returns of the m assets can conjointly be expressed as a multi-factor model where an unknown number $K < m$ of factors may be characteristic of this universe (i.e. among the m assets, there exists K principal factors that are driving the universe and expressed as a linear combination of the m assets). We assume the additive noise to be a multivariate Elliptical Symmetric noise [24, 39] generalizing a correlated multivariate Gaussian noise.
We then have, for all $t \in [1,N]$:

$$\mathbf{r}_t = \mathbf{B}_t \, \mathbf{f}_t + \sqrt{\tau_t} \, \mathbf{C}^{1/2} \mathbf{x}_t, \tag{4}$$

where

- $\mathbf{r}_t$ is the m-vector of returns at time t,
- $\mathbf{B}_t$ is the $m \times K$-matrix of coefficients that define the assets sensitivities to each factor at time t,
- $\mathbf{f}_t$ is the K-vector of random factor values at t, supposed to be common to all the assets,
- $\mathbf{x}_t$ is a m-vector of independent Gaussian white noise with unit variance and non-correlated with the factors, i.e. $\mathbb{E}[\mathbf{x}_t \, \mathbf{f}_t^T] = \mathbf{0}_{m \times K}$,
- $\mathbf{C}$ is called the $m \times m$ scatter matrix that is supposed to be Toeplitz[2] structured [19] and time invariant over the period of observation,
- τ_t is a family of i.i.d positive random variables with expectation τ that is independent of the noise and the factors and drives the variance of the noise. These random variables are time-dependent and generate the Elliptical distribution [6] of the noise.

The Toeplitz assumption made on $\mathbf{C}$ is a required assumption for the proposed methodology described in section 4.1. This hypothesis imposes a particular structure for the covariance matrix of the additive noise, and is generally used to describe stationary processes [19]. In the case of model (4) this hypothesis is plausible as it states that the additional white noise admits a Toeplitz-structured covariance matrix. In practice, and especially in the case of financial series, where we observe only one sample at each time, the cross-sectional stationarity is quite unobservable and even unprovable. This motivates the extension we propose in this paper, described in section 4.4, to splitting the assets universe into groups composed of assets having similar distributions, and being most probingly sampled from a stationary process representing a unique distribution for each group.

---

[2] A Toeplitz matrix is a diagonal-constant matrix.

Given equation (4) the covariance matrix writes for a fixed period of time t:

$$\Sigma_t = \mathbf{B}_t \Sigma_t^f \mathbf{B}_t^T + \tau \mathbf{C}, \tag{5}$$

that is a $m \times m$-matrix composed of two terms: the factor-related term with $\Sigma_t^f = \mathbb{E}[\mathbf{f}_t \mathbf{f}_t^T]$ being of rank K, and the noise-related term being of rank m. Subspace methods aims at identifying the K highest eigenvalues of Σ_t supposed to represent the K-factors especially when the power of the factors is higher than the noise power.

The efficient estimation of the number of factors K is really a challenging problem for many applications including financial applications:

- identifiability of the main K factors to build new portfolios. This problem is for example closely related to linear unmixing problem in Hyperspectral Imaging [2],
- identifiability of the main K factors to separate signal and noise subspaces in order to build projectors, to filter noisy part of the data through jointly robust and efficient covariance matrix estimation. This is for example useful for portfolio allocation or in risk management [15, 16, 22, 37].

The identified theoretical problems to solve are clearly the estimation of the order K of the model and the efficiency of covariance matrix estimation under correlated non-Gaussian noise hypothesis.

4 Proposed Methodology

4.1 General framework

Under general non-Gaussian noise hypothesis proposed in Section 3, Tyler M-estimator [40, 51] is shown to be the most robust covariance matrix estimate. Given N observations of the m-vector $\mathbf{r}_t$ with $m < N$, the Tyler-M estimate $\widehat{\mathbf{C}}_{tyl}$ is defined as the solution of the following "fixed-point" equation:

$$\mathbf{X} = \frac{m}{N} \sum_{t=1}^{N} \frac{\mathbf{r}_t \mathbf{r}_t^T}{\mathbf{r}_t^T \mathbf{X}^{-1} \mathbf{r}_t}, \tag{6}$$

with $Tr(\widehat{\mathbf{C}}_{tyl}) = m$. The scatter matrix, solution of (6) has some remarkable properties [31, 42] like being a robust estimator of the true scatter matrix and being also "variance-free": it really reflects the true structure of the underlying process without noise pollution.

When the noise is assumed to be white, several methods, based on the RMT, have been proposed [12] to extract information of interest from the received signals. One can cite for instance the number of embedded sources estimation [25], the problem of radar detection [13], signal subspace estimation [21]. However, when the additive noise is correlated, some RMT methods require the estimation of a specific threshold which has no explicit expression and can be very difficult to obtain [11, 52] while the others assume that the covariance matrix is known and use it, through some source-free secondary data, to whiten the signal. According to the following *consistency theorem* found and proved in [48–50], recent works have proposed to solve the problem through a biased Toeplitz estimate of $\widehat{\mathbf{C}}_{tyl}$, let us say $\widetilde{\mathbf{C}}_{tyl} = \mathscr{T}\left(\widehat{\mathbf{C}}_{tyl}\right)$:

Consistency Theorem [48–50]
Under the RMT regime assumption, i.e. that $N, m \to \infty$*, and the ratio* $m/N \to \xi > 0$*, we have the following spectral convergence:*

$$\left\| \mathscr{T}\left(\widehat{\mathbf{C}}_{tyl}\right) - \mathbf{C} \right\| \xrightarrow{a.s.} 0. \tag{7}$$

This powerful theorem says that it is possible to estimate the covariance matrix of the correlated noise even if the observations contain the sources or information to be retrieved.

According to this result, the first step is then to whiten the observations using $\widetilde{\mathbf{C}}_{tyl}$. The whitened observations are then defined as $\mathbf{r}_{w,t} = \widetilde{\mathbf{C}}_{tyl}^{-1/2}\, \mathbf{r}_t$.

Given the set of N whitened observations $\{\mathbf{r}_{w,t}\}$ and their Tyler's covariance matrix $\widehat{\Sigma}_w$, it has been shown in [50] that the eigenvalues distribution of $\widehat{\Sigma}_w$ fit the predicted bounded distribution of Marčenko-Pastur [33]. However, if one or several sources are contained in the observations, being powerful enough to be detected, then there will be as many eigenvalues as there are sources standing outside the upper bound of the Marčenko-Pastur distribution, given in that case by $\bar{\lambda} = \sigma^2 \left(1+\sqrt{c}\right)^2$ where $c = m/N$ and $\sigma^2 = 1$ (due to the preceding whitening process σ^2 is equal to one). Once the K largest eigenvalues larger than $\bar{\lambda}$ are detected, we process as for the Eigenvalue clipping in [27] to set the values of the remaining $m-K$ lowest eigenvalues to a unique value equal to $\left(Tr\left(\widehat{\Sigma}_w\right) - \sum_{k=K+1}^{m} \lambda_k\right)/(m-K)$. Using also the corresponding eigenvectors, we then build back the de-noised assets covariance matrix to be used in (2) and (3) or in any other objective function. The whitening procedure is detailed more precisely in the next subsection.

4.2 Detailed whitening procedure

Given $\mathbf{R}$ the $m \times N$-matrix of observations, the de-noised covariance matrix estimate $\widehat{\Sigma}_w$ is obtained through the following procedure steps:

S1 Set $\widehat{\mathbf{C}}_{tyl}$ as the Tyler-M estimate of $\mathbf{R}$, solution of (6),
S2 Set $\widetilde{\mathbf{C}}_{tyl} = \mathscr{T}\left(\widehat{\mathbf{C}}_{tyl}\right)$, the Toeplitz rectification matrix built from $\widehat{\mathbf{C}}_{tyl}$ for the Toeplitz operator $\mathscr{T}()$,
S3 Set $\mathbf{R}_w = \widetilde{\mathbf{C}}_{tyl}^{-1/2}\,\mathbf{R}$, the $m \times N$ matrix of the whitened observations,
S4 Set $\widehat{\Sigma}_{tyl}$ as the Tyler-M estimate of $\mathbf{R}_w$, solution of (6),
S5 Set $\widehat{\Sigma}_{tyl}^{clip} = \mathbf{U}\Lambda^{clip}\mathbf{U}^T$ where $\mathbf{U}$ is the $m \times m$ eigenvectors matrix and Λ^{clip} is the $m \times m$ diagonal matrix of the eigenvalues $(\lambda_k^{clip})_{k\in[1,m]}$ corrected using the Eigenvalue clipping method [27] as described in (10),
S6 Finally, $\widehat{\Sigma}_w = \left(\widetilde{\mathbf{C}}_{tyl}^{1/2}\right)\widehat{\Sigma}_{tyl}^{clip}\left(\widetilde{\mathbf{C}}_{tyl}^{1/2}\right)^T$.

4.3 Simulation example

As an illustration, Figure 1 compares the eigenvalues distribution of the SCM $\widehat{\mathbf{C}}_{scm} = \mathbf{R}\,\mathbf{R}^T/N$, $\widehat{\mathbf{C}}_{tyl}$ and $\widehat{\Sigma}_w$ for $K = 3$ sources of information embedded in highly non-Gaussian and highly correlated noise, a K-distributed noise [39] with shape parameter $\nu = 0.5$ and a Toeplitz covariance matrix (let us say $\mathbf{A}$), whose element $a_{i,j} = \rho^{|i-j|}, i, j = 1, \cdots, m$ and where the Toeplitz coefficient $\rho = 0.8$. If no whitening operation is made before applying the Marčenko-Pastur boundary properties of the eigenvalues (like on the left and the middle sides of the figure), then there is no chance to detect any of the sources. When the whitening process is firstly applied, then the only detected sources above the Marčenko-Pastur threshold correspond to the K sources. As a matter of fact, there is no need anymore to adapt the value of the threshold value regarding the distribution of τ_t and the estimated value of $\mathbb{E}[\tau]$ [50]. The robust Tyler M-estimator is "τ-free", so does not depend anymore of the distribution of τ_t.

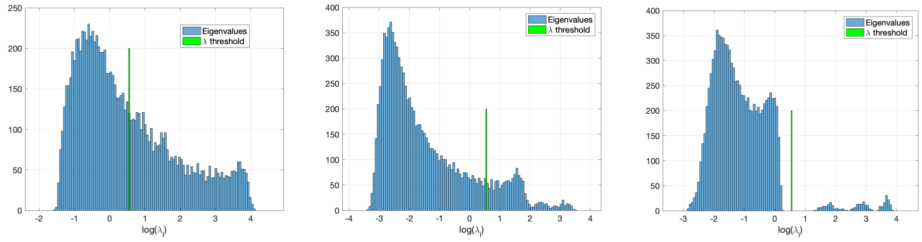

Fig. 1 Distributions of the logarithm of the eigenvalues of three covariance matrix estimates. Left side: Eigenvalues (log) of the SCM of the observations; Middle: Eigenvalues (log) of the Tyler covariance matrix of the observations; Right side: Eigenvalues (log) of the Tyler covariance matrix of the whitened observations. Observations contain $K = 3$ sources embedded in a multivariate K-distributed noise with shape parameter $\nu = 0.5$, and a Toeplitz coefficient $\rho = 0.8$. $m = 100$, $N = 1000$ ($c = 0.1$), and the (log) Marčenko-Pastur upper bound is here: $\log(\bar{\lambda}) = \log(1.7325)$.

4.4 The case of non-homogeneous assets returns

The whitening process proposed above is made under the implicit assumption that the assets returns are drawn from a unique multivariate law and are therefore homogeneous in law. As described hereafter this assumption is unrealistic for financial time series of returns. We therefore propose to split the m assets into $p < m$ groups, each composed of $\{m_q\}_{q=1}^{p}$ assets (with $\sum_{q=1}^{p} m_q = m$), and formed to be composed of assets having similar distributions. We set a fixed number of groups, and group the assets regarding their returns distributions. Under this new assumption, model (4) applies for each group q as follows:

$$\mathbf{r}_t^{(q)} = \mathbf{B}_t^{(q)}\,\mathbf{f}_t + \sqrt{\tau_t}\,\mathbf{C}_{(q)}^{1/2}\,\mathbf{x}_t, \tag{8}$$

Then, the full model (4) rewrites:

$$\begin{bmatrix} \mathbf{r}_t^{(1)} \\ \vdots \\ \mathbf{r}_t^{(p)} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_t^{(1)} \\ \vdots \\ \mathbf{B}_t^{(p)} \end{bmatrix} \mathbf{f}_t + \sqrt{\tau_t} \begin{pmatrix} \mathbf{C}_{(1)} & \mathbf{0}_{1,2} & \cdots & \mathbf{0}_{1,p} \\ \mathbf{0}_{2,1} & \mathbf{C}_{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{p-1,p} \\ \mathbf{0}_{p,1} & \cdots & \mathbf{0}_{p,p-1} & \mathbf{C}_{(p)} \end{pmatrix}^{1/2} \mathbf{x}_t, \tag{9}$$

where $\mathbf{0}_{i,j}$ denotes the null matrix of size $m_i \times m_j$, $i, j = 1, \cdots, p$, corresponding to the additional hypothesis that the groups are uncorrelated each others. The complete scatter matrix $\mathbf{C}$ is therefore block-constructed, and block-Toeplitz.

To form the groups of assets at each date t and given a past period of N observations, we proceed as follows:

- for each asset i, we compute the sample mean μ_i and the sample standard deviation σ_i using its N returns $\mathbf{r}_i$,
- we compute the "standardized" returns $\tilde{\mathbf{r}}_i = (\mathbf{r}_i - \mu_i)/\sigma_i$,
- we compute several quantiles from $\tilde{\mathbf{r}}_i$, and append μ_i and σ_i to the vector of the computed quantiles to get our variables on which to group the assets, and finally,
- we use the classical Ascending Hierarchical Classification (AHC) using the Euclidean distance and the Ward measure [53] to form the p groups.

AHC is a very classical classification method but ensures to get homogeneous groups for which the intra-group variances are smaller than the inter-group variances.

The three first steps **S1**, **S2** and **S3** of the whitening process described in 4.2 are therefore repeated for each group (q), $q = 1, \cdots, p$: given $\mathbf{R}^{(q)}$ the $m_q \times N$ matrix of observations for assets in group (q), going through **S1** to **S3** leads to $\mathbf{R}_w^{(q)}$ the whitened matrix of observations for group (q). Once $\mathbf{R}_w$ has been completed, then steps **S4** to **S6** are applied and lead to the block-constructed covariance matrix estimate $\widehat{\Sigma}_w$. This is a mixed version between a global whitening process and a diagonal whitening process (applied when the series are only standardized). Our process can be viewed as a block-diagonal whitening process and ensures that the whitened groups are more homogeneous than the overall group of assets.

5 Application

This section is devoted to show the benefits of using our proposed methodology when applied to the Maximum Variety and Minimum Variance portfolios. The investment universe[3] consists of $m = 43$ baskets of European equity stocks representing twenty-four industry subsectors (e.g. transportation, materials, energy...), thirteen countries (e.g. Sweden, France, Netherlands,...) and six factor-based indexes (e.g. momentum, quality, growth, ...). Using baskets instead of single stocks allows to reduce the idiosyncratic risks and the number of assets to be considered. We observe the prices of these assets on a daily basis from the 27th of July 2000 to the 20th of May 2019. The daily prices are close prices, i.e. the price being fixed before the financial marketplaces close at the end of each weekday.

The portfolios weights are computed as follows: every four weeks, we estimate the covariance matrix of the assets using the past one year of daily returns (so $N = 260$ weekdays) and we run the optimization procedure in order to get the vector of weights that maximizes the variety ratio (1) or minimizes the variance of the final portfolio (3) given this past period. The weights, computed say at time t, are then kept constant for the next four-week period. We apply our methodology in two manners: the first one, named "RMT-Tyler-Wh", contains the whitening process applied on the universe as a whole, whereas the one denoted by "RMT-Tyler-Wh-by-Gr" refers to the whitening process applied on each group of assets[4].

---

[3] Data are available upon request.

We compare the results with those obtained using the "SCM" and also with three other competing methods: the first one, denoted as "RMT-SCM" uses the Eigenvalue clipping of [27], the second one, that we denote as "LW", is the method that uses the Ledoit & Wolf shrinkage of [29], and finally the method using the Rotational Invariant Estimator of [4, 5], denoted as "RIE". These methods are briefly described in appendix.

We report several portfolios statistics computed over the whole period in order to quantify the benefits of the proposed methodology: the annualized return, the annualized volatility, the ratio between the annualized return and the annualized volatility, the value of the maximum drawdown (that is the return between the highest and the lowest portfolio levels observed during the whole period), and the average of the Variety Ratios computed at each rebalancing date. Higher is the return/volatility ratio, lower is the maximum drawdown and higher is the variety ratio, and better performing is the portfolio. Performances are also compared to the performance of the MSCI® Europe Index [38] (composed of large and mid cap equity stocks across 15 countries of the European regions), and to the performance of the equi-weighted portfolio, composed of all the assets that are equally weighted.

5.1 Variety Maximum (or VarMax) portfolios results

Figure 2 shows the evolution of the VarMax portfolios wealth, starting at 100 at the beginning of the first period.
The "SCM", "RMT-SCM", "LW", "RIE", "RMT-Tyler-Wh" and "RMT-Tyler-Wh-by-Gr" VarMax portfolios are respectively in red, dashed red, dash-dotted blue, blue, purple, and green lines. The naive equi-weighted portfolio is reported as the dotted black line, and the price of the benchmark, also rebased at 100 at the beginning of the period, is the black line.

The proposed "RMT-Tyler-Wh"-based techniques clearly outperform the conventional ones. Moreover, whitening homogeneous groups of data instead of the whole data set improves even more the results. Regarding the other methods, "RMT-SCM" is the only one that outperforms significantly "SCM", but shows weaker performances than our proposed method does; "LW" and "RIE" are quite similar to "SCM".

On the figure we have reported the "net of transaction fees" portfolios wealth, considering 0.07% of fees (or 7 basis points denoted as "bp") applied to any weight change from one time to the next one. Measuring the total weights changes is referred as the turnover of the portfolio. We assume that the turnover between two consecutive periods t and $t+1$ is measured by $\sum_{i=1}^{m} |w_{i,t+1} - w_{i,t}|$. If, for example, the turnover is equal to 0.15 for changing weights from t to $t+1$, then the portfolio performance computed between t and $t+1$ will be decreased by 0.15×7 bp $= 0.0105\%$. Turnover is an important number in portfolio allocation. If you ever find an apparently well performing strategy that indicates you to change the overall portfolio at each time, then the cost of changing the overall portfolio will surely be equivalent or larger than would be the performance of the strategy itself. Here, the proposed technique leads to increase the cumulated turnover, but reasonably enough to let the improvement be a significant improvement that do not cost all the benefits of the technique. Limiting the turnover is often added as an additional non linear constraint to any optimization process like (2) or (3).

---

[4] The number of group is $p = 6$ and the quantiles used are q_θ and $q_{1-\theta}$ with $\theta \in [1\%, 2.5\%, 5\%, 10\%, 15\%, 25\%, 50\%]$.

We finally report on Table 1 some statistics on the overall portfolios performance: we compare, for the whole period, the annualized return, the annualized volatility, the ratio between the return and the volatility, the maximum drawdown and the average value of the diversification ratio, for the portfolios and the benchmark. All the indicators related to the proposed technique show a significant improvement with respect to the other methods: a higher annualized return, a lower volatility (so a higher return/volatility ratio), a lower maximum drawdown and a higher diversification ratio.

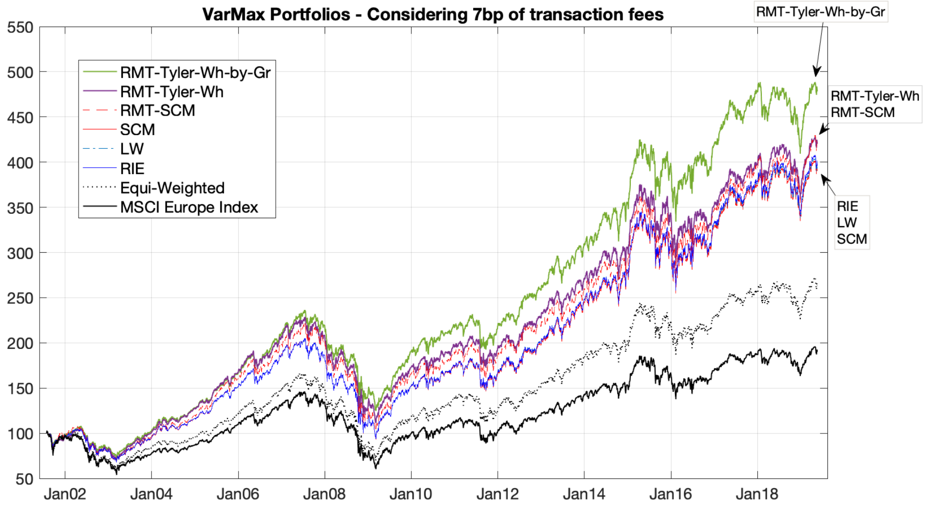

Fig. 2 VarMax portfolios wealth from July 2001 to May 2019. The proposed "RMT-Tyler-Wh-by-Gr" (green line) leads to improved performances vs the "RMT-Tyler-Wh" (purple), the "RMT-SCM" (dashed red), the "LW" (dash-dotted blue), the "RIE" (blue) and the "SCM" (red), as shown in Table 1: higher annualized return, lower annualized volatility, lower maximum drawdown and higher Diversification Ratio. But it results in a twice higher turnover: we then have taken into account 7bp (or 0.07%) of transactions fees to compare the portfolios wealth.

VarMax Portfolios	Annualized Return	Annualized Volatility	Ratio (Return / Volatility)	Maximum Drawdown	Diversification Ratio (avg)
RMT-Tyler-Wh-by-Gr	9.65%	12.03%	0.80	46.84%	1.57
RMT-Tyler-Wh	8.90%	13.16%	0.68	51.18%	1.44
RMT-SCM	8.94%	13.79%	0.65	54.15%	1.27
RIE	8.65%	13.65%	0.63	54.44%	1.38
LW	8.59%	13.57%	0.63	54.28%	1.40
SCM	8.56%	13.68%	0.63	54.45%	1.38
Equi-Weighted	*6.60%*	*15.37%*	*0.43*	*57.82%*	*1.19*
Benchmark	*4.71%*	*14.87%*	*0.32*	*58.54%*	

Table 1 Some performance numbers for VarMax portfolios with 0.07% of fees from July 2001 to May 2019. The results are ranked in descending order according to the ratio (Return / Volatility).

5.2 Minimum Variance (or MinVar) portfolios results

Results obtained for the MinVar portfolios also show some improvements but less important than for the VarMax portfolios. Figure 3 shows that whitening by groups ("RMT-Tyler-Wh-by-Gr") improves the performance whereas whitening the whole assets ("RMT-Tyler-Wh") do not bring improvement with respect to all the other approaches, even if the variety ratio is higher. "RMT-SCM", "LW" and "RIE" provide lower or similar performances if compared to "SCM". Minimizing the portfolio variance leads to choosing the assets having the lowest volatilities. Then, using a robust approach does flatten the volatility differences between assets and then the ex-post portfolio volatility, computed classically, will be higher than the ex-post portfolio volatility computed using the robust matrix. Nevertheless, our process leads to higher performance that the classical SCM exhibiting a higher diversification ratio, and also a lower maximum drawdown.

To illustrate this purpose, Figure 4 plots the standard deviations of the invested assets versus the resulting weights obtained for MinVar/SCM weights (on the top graph) the VarMax/SCM (on the bottom graph). The same conclusion arises for the "RMT-Tyler-Wh-by-Gr". It shows explicitly which assets are preferred and when, according to their volatility level. On a similar way, Figure 5 shows that VarMax assigns non-zeros weights to the less correlated assets if compared to the non-zeros MinVar weights.

As for the VarMax portfolios, Table 2 reports the MinVar portfolios statistics. Again, the indicators related to the proposed technique show an improvement if compared to the classical techniques.

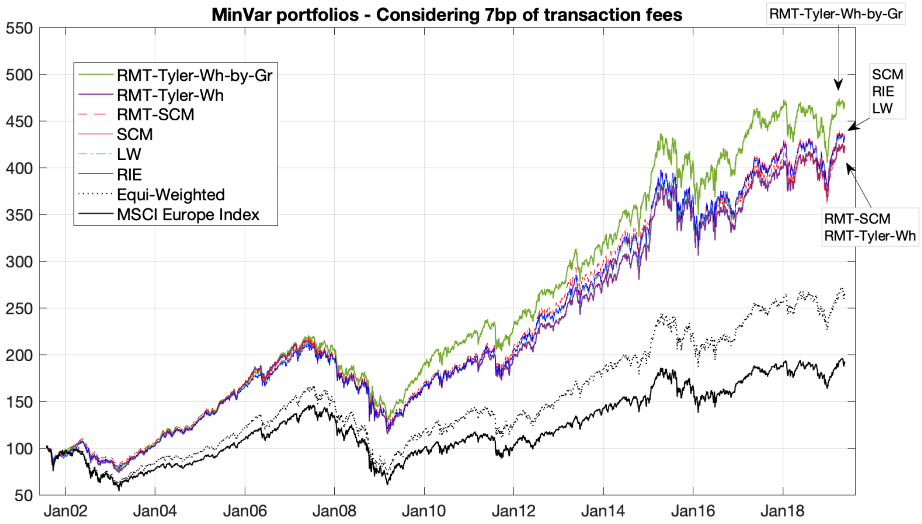

Fig. 3 MinVar portfolios wealth from July 2001 to May 2019. The proposed "RMT-Tyler-Wh-by-Gr" (green line) leads to improved performances vs the "RMT-Tyler-Wh" (purple), the "RMT-SCM" (dashed red), the "LW" (dash-dotted blue), the "RIE" (blue) and the "SCM" (red), as shown in Table 2. MinVar portfolios are known to result in poorly diversified portfolios and to invest in the lowest volatile assets. But surprisingly, the low-volatility anomaly applies in such cases.

MinVar Portfolios	Annualized Return	Annualized Volatility	Ratio (Return / Volatility)	Maximum Drawdown	Diversification Ratio (avg)
RMT-Tyler-Wh-by-Gr	9.35%	11.08%	0.84	41.07%	1.52
LW	8.75%	10.75%	0.81	43.69%	1.21
RIE	8.76%	10.78%	0.81	43.24%	1.19
SCM	8.74%	10.92%	0.80	43.78%	1.19
RMT-SCM	8.62%	10.80%	0.80	43.95%	1.14
RMT-Tyler-Wh	8.72%	11.58%	0.75	46.50%	1.36
Equi-Weighted	*6.60%*	*15.37%*	*0.43*	*57.82%*	*1.19*
Benchmark	*4.71%*	*14.87%*	*0.32*	*58.54%*	

Table 2 Some performance numbers for MinVar portfolios with 0.07% of fees from July 2001 to May 2019. The results are ranked in descending order according to the ratio (Return / Volatility).

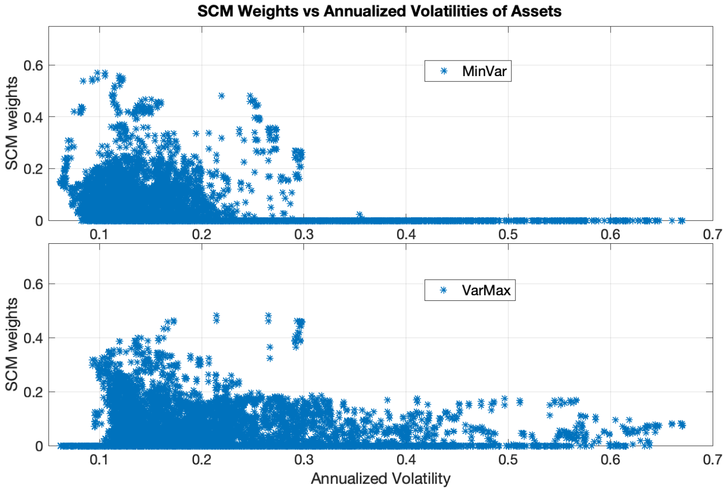

Fig. 4 VarMax and MinVar SCM weights versus the assets volatilities. As expected, MinVar weights are mostly non-zeros for the assets having the lowest volatilities. VarMax weights are more indifferent to the volatility levels.

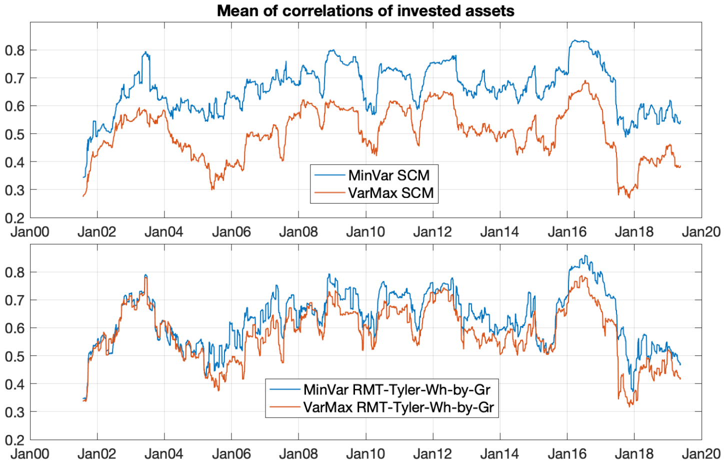

Fig. 5 Average correlation of the invested assets for the VarMax and MinVar portfolios combined with either SCM or RMT-Tyler-Wh-by-Gr method. VarMax SCM weights are assigned to the less correlated assets if compared to the SCM MinVar weights and the difference is reduced in the RMT-Tyler-Wh-by-Gr case.

6 Conclusion

In this paper, we have shown that when the covariance matrix is estimated with the Tyler M-estimator and the RMT, the Maximum Variety and the Minimum Variance Portfolio allocation processes lead to improved performances with respect to several classical estimators. The improvements come especially from the robust and de-noised version of the covariance matrix estimate. Indeed, we have modelled the assets returns as a multi-factor model embedded in a correlated elliptical and symmetric noise, allowing to account for non-Gaussian and correlated noise. Given this model setup, then we show how to separate the signal from the noise subspace using a "toeplitzified" robust and consistent Tyler-M estimator and the Random Matrix theory applied on the whitened covariance matrix estimate. Moreover, we show that if the assets are grouped within homogeneously distributed classes before processing, then the results show much improvements. This paper has focused on both the Maximum Variety and Minimum Variance portfolios but can be applied on other allocation framework involving covariance matrix estimation (and/or model order selection). Moreover this can also be exploited to define the main directions of information and to construct pure factor driven models. These methods have also shown their importance in the radar and hyperspectral fields and are very promising techniques for many applications.

Acknowledgements We would like to thank Fideas Capital for supporting this research and providing the data. We thank particularly Pierre Filippi and Alexis Merville for their constant interaction with the research team at Fideas Capital. Moreover, this research was conducted within the "Construction of factorial indexes and allocation" under the aegis of the Europlace Institut of Finance, a joint initiative with Fideas Capital.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix: Brief description of alternative covariance matrix estimators

Here, we briefly introduce some well-known covariance matrix estimators. In the following, $c = m/N$ and $\widehat{\mathbf{E}} = \widetilde{\mathbf{R}}\widetilde{\mathbf{R}}^T/N$ is the standardized SCM where $\widetilde{\mathbf{R}} = (\widetilde{\mathbf{r}}_i)_{i\in[1,m]}$ as defined in section 4.4.

A.1 Eigenvalue clipping (or RMT-SCM)

Laloux et al. [27] proposed Eigenvalue clipping in order to separate signal and noise subspaces using Marčenko-Pastur [36] boundary properties of the eigenvalues. The Eigenvalue clipping estimator of $\widehat{\mathbf{E}}$ is as follows:

$$\widehat{\mathbf{E}}_{clip} = \sum_{k=1}^{m} \lambda_k^{clip} \mathbf{u}_k \mathbf{u}_k^T$$

with $\mathbf{u}_k$ the eigenvector associated to the eigenvalue λ_k of $\widehat{\mathbf{E}}$, and λ_k^{clip} defined as follows:

$$\lambda_k^{clip} = \begin{cases} \lambda_k, \text{ if } \lambda_k \geq (1+\sqrt{c})^2 \\ \widetilde{\lambda}, \text{ otherwise} \end{cases} \tag{10}$$

where $\widetilde{\lambda}$ is chosen such that $Tr(\widehat{\mathbf{E}}_{clip}) = Tr(\widehat{\mathbf{E}})$.

A.2 Ledoit & Wolf shrinkage (or LW)

Ledoit & Wolf [29] introduced some shrinkage estimators particularly adapted to financial asset returns and based on the single factor model of Sharpe [47], where the factor is a market index. LW is a linear combination of the SCM and the covariance matrix containing the market information. This model can be written as follows:

$$r_{i,t} = \alpha_i + \beta_i F_t + \varepsilon_{i,t}, \forall i \in [1,m] \text{ and } \forall t \in [1,N] \tag{11}$$

where $r_{i,t}$ is the return of stock i at time t , α_i is the active return of the asset i, F_t is the market index return at time t, β_i is the asset sensitivity to the market index return, and $\varepsilon_{i,t}$ is the idiosyncratic return for asset i at t. This latter term is assumed to be uncorrelated to the market index. Then the covariance matrix writes:

$$\mathbf{M}_r = \sigma_F^2 \beta \beta^T + \Omega_\varepsilon$$

with $\beta = [\beta_1, \cdots, \beta_m]^T$, σ_F^2 is the variance of the market returns and Ω_ε the covariance matrix of the idiosyncratic error.
An estimator for $\mathbf{M}_r$ can be determined:

$$\widehat{\mathbf{M}}_r = \hat{\sigma}_F^2 \widehat{\beta} \widehat{\beta}^T + \widehat{\Omega}_\varepsilon$$

where each $\hat{\beta}_i$ is estimated individually using the OLS estimator based on equation (11) and the $\widehat{\Omega}_\varepsilon$ is a diagonal matrix composed of the OLS residual variances. Finally, $\hat{\sigma}_F^2$ is the sample variance of the market returns.

The Shrinkage-to-Market estimator from Ledoit & Wolf is therefore equal to:

$$\widehat{\Sigma}(\gamma) = \gamma \widehat{\mathbf{M}}_r + (1-\gamma)\mathbf{S}$$

where $\gamma \in [0,1]$ is the shrinkage parameter estimated as in [29], and $\mathbf{S}$ is the SCM of asset returns.

A.3 Rotational invariant estimator (or RIE)

Bun et al. [4, 5] proposed an optimal rotational invariant estimator for general covariance matrices by computing the overlap between the true and sample eigenvectors introduced first by Ledoit & Péché [28]. For large m, the optimal rotational invariant estimator (RIE) of $\widehat{\mathbf{E}}$ is as follows:

$$\widehat{\mathbf{E}}_{RIE} = \sum_{k=1}^{m} \lambda_k^{RIE} \mathbf{u}_k \mathbf{u}_k^T$$

with $\mathbf{u}_k$ the eigenvector associated to the eigenvalue λ_k of $\widehat{\mathbf{E}}$, and λ_k^{RIE} defined as follows:

$$\lambda_k^{RIE} = \frac{\lambda_k}{|1 - c + c z_k s(z_k)|^2}$$

where $z_k = \lambda_k - iN^{-1/2}$ is a complex number and $s(z)$ denotes the discrete form of the limiting Stieltjes transform

$$s(z) = \frac{1}{m} \sum_{j=1}^{m} \frac{1}{z - \lambda_j}$$

We also ensure that $Tr(\widehat{\mathbf{E}}_{RIE}) = Tr(\widehat{\mathbf{E}})$. For this purpose, we multiply each λ_k by ν with $\nu = \sum_{k=1}^{m} \lambda_k / \sum_{k=1}^{m} \lambda_k^{RIE}$.

References

1. Abramovich, Y., Spencer, N.K.: Diagonally loaded normalised sample matrix inversion (LNSMI) for outlier-resistant adaptive filtering. In: IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), vol. 3 (2007)

2. Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J.: Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5**(2), 354–379 (2012)
3. Bouchaud, J.P., Potters, M.: Financial applications of random matrix theory: a short review. The Oxford handbook of Random Matrix Theory, Oxford University Press (2011)
4. Bun, J., Allez, R., Bouchaud, J.P., Potters, M.: Rotational invariant estimator for general noisy matrices. IEEE Transactions on Information Theory **62**(12), 7475–7490 (2016)
5. Bun, J., Bouchaud, J.P., Potters, M.: Cleaning large correlation matrices: tools from random matrix theory. Physics Reports **666**, 1–165 (2017)
6. Cambanis, S., Huang, S., Simons, G.: On the theory of elliptically contoured distributions. Journal of Multivariate Analysis **11**(3), 368–385 (1981)
7. Chen, Y., Wiesel, A., Hero, A.O.: Robust shrinkage estimation of high-dimensional covariance matrices. IEEE Transactions on Signal Processing **59**(9) (2011)
8. Choueifaty, Y., Coignard, Y.: Toward maximum diversification. Journal of Portfolio Management **35**(1), 40–51 (2008)
9. Choueifaty, Y., Froidure, T., Reynier, J.: Properties of the most diversified portfolio. Journal of investment strategies **2**(2), 49–70 (2013)
10. Clarke, R., Silva, H.D., Thorley, S.: Minimum variance, maximum diversification, and risk parity: an analytic perspective. Journal of Portfolio Management (2012)
11. Couillet, R.: Robust spiked random matrices and a robust G-MUSIC estimator. Journal of Mult. Analysis **140**, 139–161 (2015)
12. Couillet, R., Debbah, M.: Random matrix methods for wireless communications. Cambridge University Press (2011)
13. Couillet, R., Greco, M.S., Ovarlez, J.P., Pascal, F.: RMT for whitening space correlation and applications to radar detection. In: IEEE CAMSAP, pp. 149–152 (2015)
14. Couillet, R., Pascal, F., Silverstein, J.W.: Robust estimates of covariance matrices in the large dimensional regime. IEEE Transactions on Information Theory **60**(11) (2014)
15. Darolles, S., Duvaut, P., Jay, E.: Multi-factor models and signal processing techniques: Application to quantitative finance. John Wiley & Sons (2013)
16. Darolles, S., Gouriéroux, C., Jay, E.: Robust portfolio allocation with risk contribution restrictions. In: Forum GI - Paris (2013)
17. Fama, E.F., French, K.R.: Common risk factors in the returns on stocks and bonds. Journal of Financial Economics **33**(1), 3–56 (1993)
18. Fama, E.F., French, K.R.: A five-factor asset pricing model. Journal of Financial Economics **116**(1), 1–22 (2015)
19. Gray, R.M.: Toeplitz and circulant matrices: A review. Foundations and Trends® in Communications and Information Theory **2**(3), 155–239 (2006)
20. Grinold, R., Rudd, A., Stefek, D.: Global factors: Fact or fiction? Journal of Portfolio Management pp. 79–88 (1989)
21. Hachem, W., Loubaton, P., Mestre, X., Najim, J., Vallet, P.: A subspace estimator for fixed rank perturbations of large random matrices. Journal of Multivariate Analysis **114**, 427–447 (2013)
22. Jay, E., Duvaut, P., Darolles, S., Chrétien, A.: Multi-factor models: examining the potential of signal processing techniques. IEEE Signal Processing Magazine **28**(5) (2011)
23. Jay, E., Terreaux, E., Ovarlez, J.P., Pascal, F.: Improving portfolios global performance with robust covariance matrix estimation: Application to the maximum variety portfolio. In: 26th European Signal Processing Conference (EUSIPCO) (2018)
24. Kelker, D.: Distribution theory of spherical distributions and a location-scale parameter generalization. Sankhyā: The Indian Journal of Statistics, Series A **32**(4), 419–430 (1970)
25. Kritchman, S., Nadler, B.: Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory. IEEE Transactions on Signal Processing **57**(10), 3930–3941 (2009)
26. Laloux, L., Cizeau, P., Bouchaud, J.P., Potters, M.: Noise dressing of financial correlation matrices. Physycal Review Letters **83**(1468) (1999)
27. Laloux, L., Cizeau, P., Potters, M., Bouchaud, J.P.: Random Matrix Theory and financial correlations. International Journal of Theoretical and Applied Finance **3**(03), 391–397 (2000)
28. Ledoit, O., Péché, S.: Eigenvectors of some large sample covariance matrix ensembles. Probability Theory and Related Fields **151**(1), 233–264 (2011)
29. Ledoit, O., Wolf, M.: Improved estimation of covariance matrix of stock returns with an application to portfolio selection. Journal of Empirical Finance **10**, 603–621 (2003)
30. Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis **88**, 365–411 (2004)

31. Mahot, M., Pascal, F., Forster, P., Ovarlez, J.P.: Asymptotic properties of robust complex covariance matrix estimates. IEEE Transactions on Signal Processing **61**(13), 3348–3356 (2013)
32. Maillard, S., Roncalli, T., Teiletche, J.: The properties of equally weighted risk contributions portfolios. Journal of Portfolio Management **36**, 60–70 (2010)
33. Marchenko, V.A., Pastur, L.A.: Distribution of eigenvalues for some sets of random matrices. Matematicheskii Sbornik (1967)
34. Markowitz, H.M.: Portfolio selection. Journal of Finance **7**(1), 77–91 (1952)
35. Maronna, R.A.: Robust *M*-estimators of multivariate location and scatter. Annals of Statistics **4**(1), 51–67 (1976)
36. Marčenko, V.A., Pastur, L.A.: Distribution of eigenvalues for some sets of random matrices. Math. USSR-Sbornik **1**(4), 457–483 (1967)
37. Melas, D., Suryanarayanan, R., Cavaglia, S.: Efficient replication of factor returns (2009). MSCI Barra Research Paper No. 2009-23
38. MSCI: Msci europe index. https://www.msci.com/europe
39. Ollila, E., Tyler, D.E., Koivunen, V., Poor, H.V.: Complex Elliptically Symmetric distributions: Survey, new results and applications. IEEE Transactions on Signal Processing **60**(11), 5597–5625 (2012)
40. Pascal, F., Chitour, Y., Ovarlez, J.P., Forster, P., Larzabal, P.: Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis. IEEE Transactions on Signal Processing **56**(1), 34–48 (2008)
41. Pascal, F., Chitour, Y., Quek, Y.: Generalized robust shrinkage estimator and its application to STAP detection problem. IEEE Transactions on Signal Processing **62**(21) (2014)
42. Pascal, F., Forster, P., Ovarlez, J.P., Larzabal, P.: Performance analysis of covariance matrix estimates in impulsive noise. IEEE Transactions on Signal Processing **56**(6), 2206–2217 (2008)
43. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., Stanley, H.E.: Collective behavior of stock price movements: A Random Matrix Theory approach. Physica A **299**, 175–180 (2001)
44. Potters, M., Bouchaud, J.P., Laloux, L.: Financial applications of Random Matrix Theory: old laces and new pieces. Acta Physica Polonica B **36**(9) (2005)
45. Rosenberg, B.: Extra-Market components of covariance in security markets. Journal of Financial and Quantitative Analysis pp. 263–274 (1974)
46. Rousseeuw, P., Driessen, K.V.: A fast algorithm for the minimum covariance determinant estimator. Technometrics **41**, 212–223 (1999)
47. Sharpe, W.F.: Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance **19**(3), 425–442 (1964)
48. Terreaux, E., Ovarlez, J.P., Pascal, F.: New model order selection in large dimension regime for Complex Elliptically Symmetric noise. In: 25th European Signal Processing Conference (EUSIPCO), pp. 1090–1094 (2017)
49. Terreaux, E., Ovarlez, J.P., Pascal, F.: Robust model order selection in large dimensional Elliptically Symmetric noise. arXiv preprint, https://arxiv.org/abs/1710.06735 (2017)
50. Terreaux, E., Ovarlez, J.P., Pascal, F.: A Toeplitz-Tyler estimation of the model order in large dimensional regime. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2018)
51. Tyler, D.E.: A distribution-free *M*-estimator of multivariate scatter. The annals of Statistics **15**(1), 234–251 (1987)
52. Vinogradova, J., Couillet, R., Hachem, W.: Statistical inference in large antenna arrays under unknown noise pattern. IEEE Transactions on Signal Processing **61**(22), 5633–5645 (2013)
53. Ward, J.H.J.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association **58**, 236–244 (1963)
54. Yang, L., Couillet, R., McKay, M.R.: A robust statistics approach to minimum variance portfolio optimization. IEEE Transactions on Signal Processing **63**(24), 6684–6697 (2015)
55. Yang, L., Couillet, R., McKay, M.R.: Minimum variance portfolio optimization in the spiked covariance model. In: Proc. IEEE 7th Computational Advances in Multi-Sensor Adaptive Processing Workshop (CAMSAP)) (2017)