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

Emmanuelle Jay[a], Eugénie Terreaux[b], Thibault Soler[a,c], Jean-Philippe Ovarlez[b,d], Frédéric Pascal[e], Philippe De Peretti[c], Christophe Chorro[c]

[a] Fideas Capital, F-75001 Paris, France; [b] CentraleSupélec, Université Paris-Saclay, F91190 Gif-sur-Yvette, France; [c] Université Paris 1 Panthéon-Sorbonne, 75647 Paris Cedex 13, France
[d] ONERA, Université Paris-Saclay, F-91123, Palaiseau, France; [e] L2S, CentraleSupélec, Université Paris-Saclay, F-91190 Gif-sur-Yvette, France

Problem Setting

- Frequently used portfolio allocation processes require the estimation of the covariance matrix of the assets returns
  (e.g. Global Minimum Variance [Maillard 10, Clarke 12], Maximum Variety [Fideas Capital] or Most Diversified Portfolio [Choueifaty 08], Mean-Variance [Markovitz 52], etc.)
  - → The Sample Covariance Matrix (SCM) - optimal under the Normal assumption - is the mostly used estimator, but, financial time series of returns might exhibit outliers,
  - → The field of robust estimation intends to deal with outliers ([*Maronna 76, Tyler 87*]),
  - → Hybrid robust shrinkage covariance matrix estimates have also been proposed building estimators upon Tyler's robust M-estimator ([*Chen 11, Ollila 14, Pascal 14*]),
  - → Recent works based on Random Matrix Theory (RMT) have also considered robust estimation in the large dimensional regime ([*Yang 15*]).
- A way to mitigate covariance matrix estimation errors is to identify the most informative asset part and then to filter the noisy part of the data
  - → Standard statistical methods like the principal component analysis may fail in distinguishing informative factors from the noisy ones,
  - → RMT helps in finding a solution for filtering noise ([*Laloux 99 and 00, Potters 05*]), but needs to be adapted to non-homogeneous and correlated time series,
  - ⇒ To fill this gap, we extend our preliminary results found in [*Jay 18*] and based upon the most up-to-date RMT-based model order selection [*Vinogradova 13, Terreaux 17*] to non-homogeneous time series.

We propose in this paper...

... **a new process for estimating and denoising covariance matrix that leads to improved global portfolio performances** ***(reduced Draw-Down, increased Sharpe Ratio, etc.)***.

- → Asset returns are modelled as a multi-factor model ([*Jay 11, Darolles 13*]),
- → Assets are grouped regarding the distribution of their returns,
- → The up-to-date RMT-based Model Order Selection method [*Terreaux 17, Jay 18*] is applied on groups of homogeneous assets to denoise the assets covariance matrix.

Assets returns Model

Let $\{\mathbf{r}_t\}_{t\in[1,N]}$ be N observations of the m assets returns, modelled as a K-factor model. For each observation date t, we then have:

$$\mathbf{r}_t = \sum_{k=1}^{K} \mathbf{f}_{t,k}\, \boldsymbol{\beta}_k + \sqrt{\tau_t}\, \mathbf{C}^{1/2}\, \mathbf{n}_t\,, \quad t \in [1, N]\,,$$

or, written more compactly: $\mathbf{R} = \mathbf{B}\,\mathbf{F} + \mathbf{C}^{1/2}\,\mathbf{N}\,\mathbf{T}^{1/2}$, where

- $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N] \in \mathbb{R}^{m\times N}$ are the observations,
- $\mathbf{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K) \in \mathbb{R}^{m\times K}$ is an unknown mixing matrix of coefficients (or beta) that define the proportion of the K factors in each asset,
- $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_N) \in \mathbb{R}^{K\times N}$ is an unknown matrix of the K common returns,
- $\mathbf{T} = \mathrm{diag}(\tau_1, \ldots, \tau_N) \in \mathbb{R}^{+N\times N}$ is a diagonal matrix unknown containing random texture,
- $\mathbf{N} \in \mathbb{R}^{m\times N}$ is a white Gaussian noise ($E\left[\mathbf{n}_t^T\,\mathbf{n}_t\right] = 1$), independent of the K factors,
- $\mathbf{C} \in \mathbb{R}^{m\times m}$ is an unknown **Toeplitz** scatter matrix ($\mathrm{Tr}(\mathbf{C}) = m$).

Theoretical Results [*Terreaux 17*]

Robust Consistent Estimation for C

Let $\hat{\mathbf{M}}_{FP} = \frac{m}{N}\sum_{t=1}^{N} \frac{\mathbf{r}_t\,\mathbf{r}_t^T}{\mathbf{r}_t^T\,\hat{\mathbf{M}}_{FP}^{-1}\,\mathbf{r}_t}$ be the scatter matrix Tyler M-estimator of $\mathbf{R}$.

As $m, N \to \infty$ such that $m/N \to c\in]0,\infty[$, we have $\left\|\mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right] - \mathbf{C}\right\| \xrightarrow{a.s.} 0$,

where $\mathcal{T}[\cdot]$ is the **Toeplitz rectification** operator: $(\mathcal{T}[\mathbf{X}])_{ij} = \frac{1}{m}\sum_{k=1}^{m} \mathbf{X}_{k,k+|i-j|}$.

A consistent estimator $\hat{\mathbf{C}}$ of the background scatter matrix $\mathbf{C}$ characterizing the background noise is therefore defined through observations $\mathbf{R}$ as $\hat{\mathbf{C}} = \mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right]$.

$\Longrightarrow$ The observations $\mathbf{R}$ can now be whitened through $\hat{\mathbf{C}}^{-1/2}\,\mathbf{R}$

Behavior of whitened data

Let $\mathbf{R}_w = \left(\mathcal{T}\left[\hat{\mathbf{M}}_{FP}\right]\right)^{-1/2}\mathbf{R}$ be the whitened data and $\hat{\mathbf{W}}_{FP}$ be the Tyler M-estimator of $\mathbf{R}_w$. As $m, N \to \infty$ such that $m/N \to c\in]0,\infty[$, if $\mathbf{R}_w$ does not contain any factor, then:

$$\left\|\hat{\mathbf{W}}_{FP} - \frac{1}{N}\mathbf{N}\,\mathbf{N}^T\right\| \xrightarrow{a.s.} 0\,.$$

- Without factors, the spectral distribution of the whitened data scatter matrix of $\mathbf{R}_w$ follows a Marchenko-Pastur distribution (same spectral distribution of unobservable covariance matrix of $\mathbf{N}$) characterized by its support $\left[(1-\sqrt{c})^2, (1+\sqrt{c})^2\right]$,
- All eigenvalues greater than $\bar{\lambda} = (1+\sqrt{c})^2$ can be considered as significant factors,
- Given that the eigenvalues are 'detectable', then $\hat{\lambda}_{k,N} \overset{N\to\infty}{\longrightarrow} \frac{(\lambda_k+\sigma^2)\,(\lambda_k+\sigma^2\,c)}{\lambda_k}$.

Estimation of K the number of factors

Let $(\lambda_k)_{k\in[1,m]}$ be the sorted eigenvalues of $\hat{\mathbf{W}}_{FP}$, then: $\hat{K} = \underset{k}{\mathrm{argmax}}\left(\lambda_k > \bar{\lambda}\right)$.

Illustration: estimating the correct number of factors

Estimating K is really a challenging problem for many applications where informative signal is embedded in correlated noise. Below, we show how our process allows to detect the $K = 3$ sources embedded in non-Gaussian and strongly correlated noise.

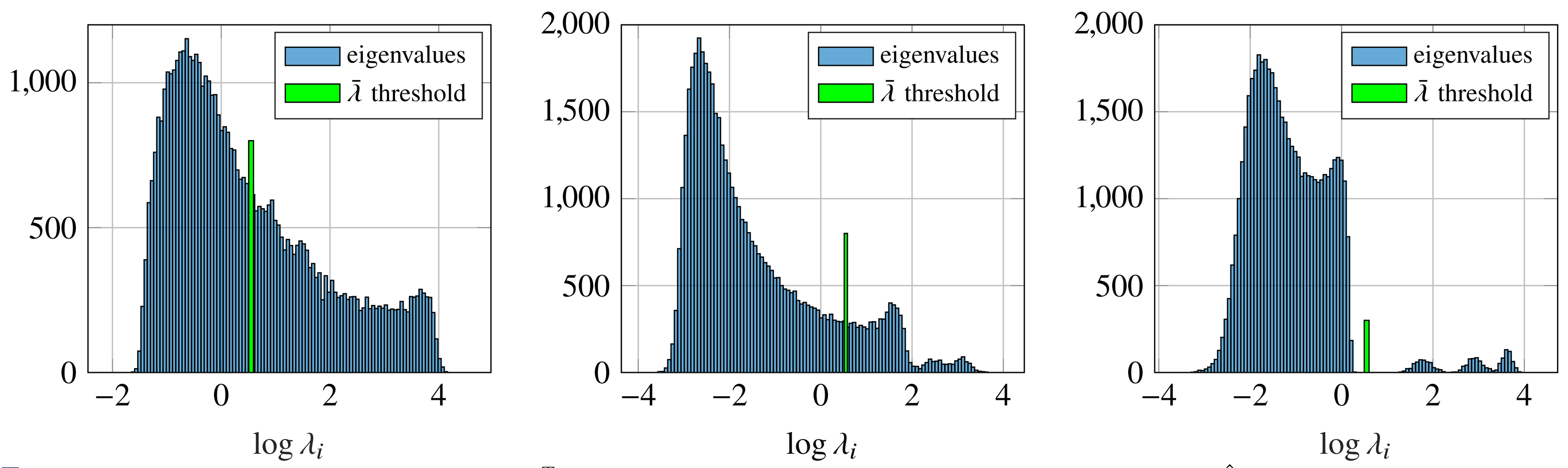

Figure 1: Eigenvalue distributions. Left: $\mathbf{R}\,\mathbf{R}^T/N$, Sample Covariance Matrix of observations. Middle: $\hat{\mathbf{M}}_{FP}$, Tyler covariance matrix of observations. Right: $\hat{\mathbf{W}}_{FP}$, Tyler covariance matrix of observations after whitening process. K-distributed case with shape parameter $\nu = 0.5$, $\rho = 0.8$, $m = 100$, $N = 1000$, $K = 3$.

Application: the Variety Maximum (VarMax) Portfolio

- **VarMax portfolio** weights are the weights that maximise the Diversification Ratio (DR): $\mathbf{w}_{vm}^* = \underset{\mathbf{w}}{\mathrm{argmax}}\,DR(\mathbf{w}) = \underset{\mathbf{w}}{\mathrm{argmax}}\frac{\mathbf{w}^T\,\mathbf{s}}{(\mathbf{w}^T\,\boldsymbol{\Sigma}\,\mathbf{w})^{1/2}}$, where $\boldsymbol{\Sigma}$ is the covariance matrix and $\mathbf{s}$ such that $s_i = \sqrt{\boldsymbol{\Sigma}_{ii}}$, $i \in [1, m]$,
- **The investment universe** consists of $m = 43$ baskets of European equity stocks representing industry subsectors, countries and factor-based indices,
- **Assets are preliminary grouped** into 5 groups regarding the closeness of their respective quantiles,
- **The whitening process** is applied in a 'block-manner' for each of the 5 groups.

We compare three different VarMax portfolios with the MSCI EU Index:

- 'RMT-Tyler-Wh-by-Gr': the proposed 'by group' whitening process,
- 'RMT-Tyler-Wh': the whitening process on the whole universe,
- 'SCM': the classical SCM estimator is used.

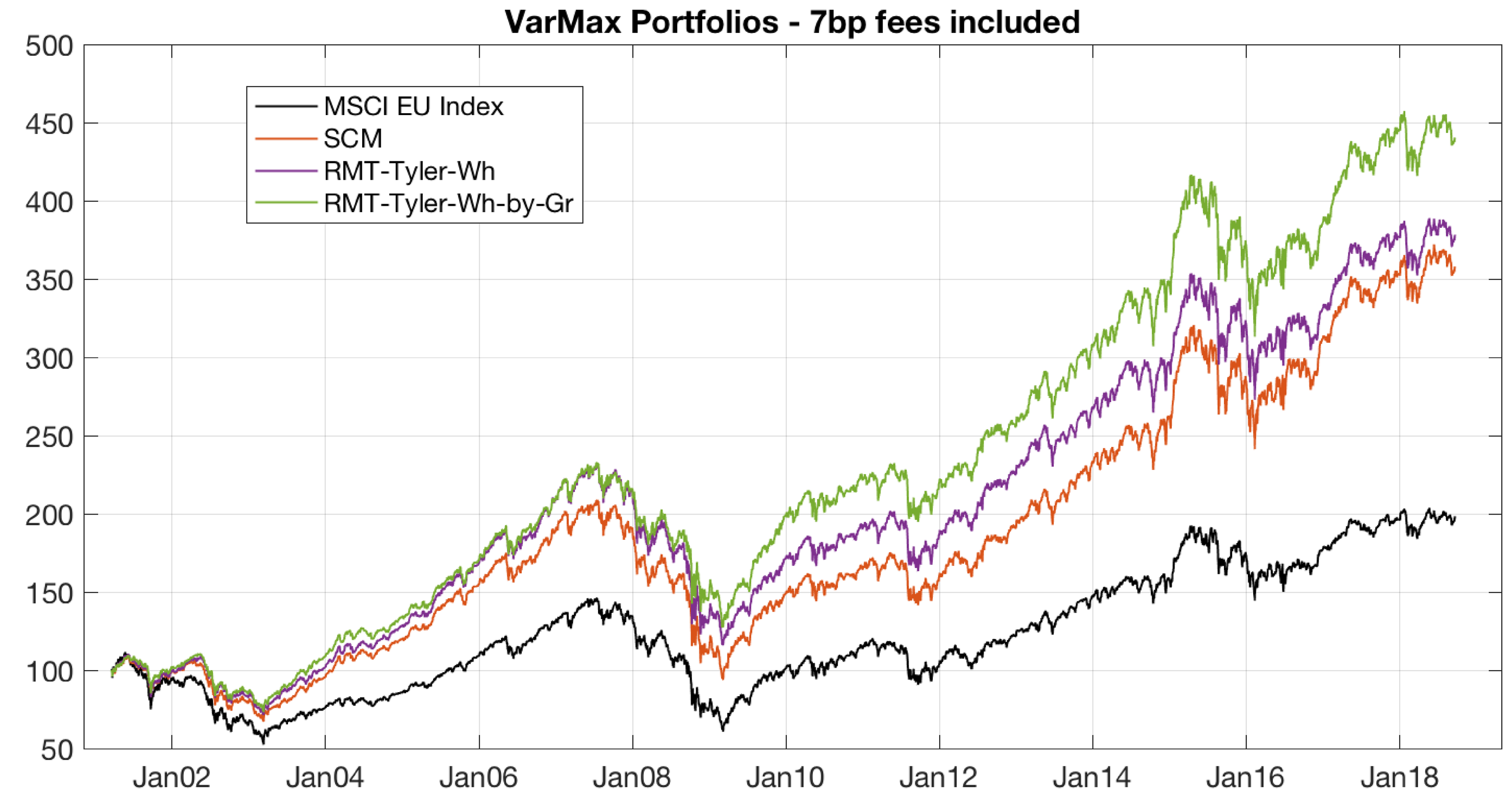

Figure 2: VarMax portfolios wealths from June 2001 to Septembre 2018. The proposed 'RMT-Tyler-Wh-by-Gr' (green line) leads to improved performances vs the 'RMT-Tyler-Wh' (purple) and the 'SCM' (red), as shown in 1: higher annualized return, lower annualized volatility, lower maximum drawdown and higher Diversification Ratio. But it results in a twice higher turnover: we then have taken into account 7bp of transactions fees to compare the portfolio wealths.

Variety Max Portfolios	Annualised Return	Annualised Volatility	Ratio (Return / Volatility)	Maximum Drawdown	Diversification Ratio (avg)
RMT-Tyler-Wh-by-Gr	8.47%	14.58%	0.58	46.27%	1.59
RMT-Tyler-Wh	7.48%	15.20%	0.49	49.22%	1.45
SCM	7.44%	16.17%	0.46	54.70%	1.38
Benchmark	3.41%	19.17%	0.18	58.54%	

Table 1: Some performance numbers for the VarMax Portfolio

Conclusion

- Asset returns have been modelled as a multi-factor model embedded in a correlated elliptical and symmetric noise, allowing to account for non-Gaussian and non correlated noise,
- Given this model setup, the most informative assets have been separated 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,
- Preliminary grouping assets leads to improved performance, as shown with the Variety Maximum portfolios.

Conf. DySES 2018 | *This research was both conducted within the "Construction of factorial indexes and allocation" under the aegis of the Europlace Institut of Finance, a joint initiative with Fideas Capital and was partly funded by the French DGA* | **Paris, 9-12 Oct. 2018**