ROBUST COVARIANCE MATRIX ESTIMATION AND PORTFOLIO ALLOCATION: THE CASE OF NON-HOMOGENEOUS ASSETS

E. Jay[◇⋆] *T. Soler*[⋆♯] *J.-P. Ovarlez*[§] *P. De Peretti*[♯] *C. Chorro*[♯]

[◇] Quanted & EIF, Palais Brongniart, 28 place de la Bourse, 75002 Paris, France
[⋆] Fideas Capital, 21 avenue de l'Opéra, 75001 Paris, France
[§] DEMR, ONERA, Université Paris-Saclay, F-91123, Palaiseau, France
[♯] Centre d'Economie de la Sorbonne, Université Paris 1, 106-112 bd de l'Hôpital, 75647 Paris Cedex 13, France

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. Our study is based on the case of the Maximum Variety Portfolio and may be obviously extended to other classical frameworks with analogous results. We focus on the fact that the assets should preferably be classified in homogeneous groups before applying the proposed methodology which is to whiten the data before estimating the covariance matrix using the robust Tyler M-estimator and the Random Matrix Theory (RMT). The proposed procedure is applied and compared to standard techniques on real market data showing promising improvements.

Index Terms— Robust Covariance Matrix Estimation, Model Order Selection, Random Matrix Theory, Portfolio Optimization, Elliptical Symmetric Noise.

1. INTRODUCTION

Portfolio allocation is often associated with the mean-variance framework fathered by Markowitz in the 50's [1]. 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, the Equally Risk Contribution Portfolio, [2], or the one developed in this paper, the Variety Maximum Portfolio, also known as the Maximum Diversified Portfolio [3, 4]. This process is designed to find weights that maximize a diversification indicator such as the variety (or diversification) ratio only involving the covariance matrix of the assets returns.

It is now well-documented that, under Normal assumptions, the Sample Covariance Matrix (SCM) is an optimal estimator of the covariance. Nevertheless, most of financial time series of returns exhibit fat tails and asymmetry hardly compatible with the Gaussian hypothesis. The field of robust covariance estimation under non-Gaussian distributions [5, 6] intends to deal with this problem especially when N, the number of samples, is larger than m, the size of the observations vector. Recent works [7, 8, 9, 10] based on Random Matrix Theory (RMT) have therefore considered robust estimation in the m, N regime in order to detect the number of targets embedded in compound or elliptical noise. In this setup, the Tyler M-estimator [5] has particularly shown its interest as being the most robust covariance matrix estimate [11].

When financial assets returns are modelled through a multi-factor model based on statistical factors, it aims at capturing the effects of the systematic risks borne by the common factors, that have to be determined from the assets universe using statistical tools. The number of factors being unknown, subspace methods help in separating the factor subspace from the noisy one based on the relative magnitudes of their eigenvalues. The problem of factor identification is quite similar as determining the number of sources embedded in compound or elliptical noise, and RMT is a powerful tool. Under the assumption of a white multivariate noise and random matrices, V. A. Marčenko and L. A. Pastur [12] defined the distribution of the eigenvalues as a bounded law, the so-called Marčenko-Pastur (MP) distribution. Then any signal or factor whose power is higher that the noise power will be detected thanks to its eigenvalue greater that the MP upper bound. In practice, additional noise is most likely non Gaussian and correlated, so that the RMT theoretical results should not be applied in their original form, but some methods have nevertheless shown their interest in financial applications [13, 14, 15, 16]. In [17] the authors found how to efficiently whiten the data before applying the original bound.

This procedure has been adapted to estimate the covariance matrix before allocating a portfolio [18], and shows improvement with respect to other classical procedures. Given that the elliptical noise assumption is made on the whole universe, the authors found that considering sub-groups of homoge-

neous assets might bring some better performance [19].
This paper focuses on assets classification and compares two different methods: the Ascending Hierarchical Clustering (AHC) method that requires the number of classes to be fixed a priori or determined using a predefined criterion (we choose here the Caliński-Harabasz (CH) criterion [20]), and the Affinity Propagation (AP) method [21] that self-determines the number of classes. Empirical tests are conducted on two different assets universes: a set of European assets and a set of American assets, both allocated with the Maximum Variety process. These tests extend our preliminary results and show that the way the assets are grouped might improve again the portfolio performance.
The article is constructed as follows: section 2 describes the Maximum Variety process, the model under consideration and its extension with groups of assets. Section 3 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 4 is dedicated to the empirical results highlighting the efficiency of the proposed method with regards to the conventional ones. Conclusion closes this paper.

Notations Matrices are in bold and capital, vectors in bold. $\mathrm{Tr}(\mathbf{X})$ is the trace of the matrix $\mathbf{X}$. For any matrix $\mathbf{A}$, $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. For any $m-$vector $\mathbf{x}$, $\mathcal{L}:\ \mathbf{x}\mapsto\mathcal{L}(\mathbf{x})$ is defined as the $m\times m$ symmetric and real-valued matrix obtained through the Toeplitz operator: $([\mathcal{L}(\mathbf{x})]_{i,j})=x_{|i-j|+1}$. For any matrix $\mathbf{A}$ of size $m\times m$, $\mathcal{T}(\mathbf{A})$ represents the matrix $\mathcal{L}(\check{\mathbf{a}})$ where $\check{\mathbf{a}}$ is a vector for which each component $\check{\mathbf{a}}_{i,\,1\le i\le m}$ 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 $\mathbf{A}$.

2. PROBLEM FRAMEWORK

Portfolio allocation is a widely studied problem. In this section we describe the Maximum Variety process that is one of the allocation process depending on the sole covariance matrix of the asset returns.

2.1. Maximum Variety (VarMax) Portfolio

The Maximum Variety (VarMax) process aims at maximizing the Variety Ratio ($\mathcal{VR}$) 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 ($\mathcal{VR}$) of the portfolio is defined as follows:

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

where $\mathbf{w}$ is the m-vector of weights, w_i representing the allocation in asset i, $\mathbf{\Sigma}$ 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 $\mathbf{\Sigma}$, i.e. $s_i=\sqrt{\mathbf{\Sigma}_{ii}}$, representing the standard deviation of the returns of the m assets.

The objective is to maximize (1) with respect to $\mathbf{w}$ in order to obtain the solution $\mathbf{w}^*_{vr}$, also called the Maximum Diversified Portfolio in [3]:

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

where we impose $0\le w_i\le 1\ \forall i\in[1,m]$ and $\displaystyle\sum_{i=1}^{m}w_i=1$.
The VarMax Portfolio verifies some interesting properties described in [4].

2.2. General 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 also 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. The additive noise is assumed to be a multivariate Elliptical Symmetric noise [22, 23] generalizing a correlated multivariate Gaussian noise. The general model writes therefore as follows, for all $t\in[1,N]$:

$$\mathbf{r}_t=\mathbf{B}_t\,\mathbf{f}_t+\sqrt{\tau_t}\,\mathbf{C}^{1/2}\,\mathbf{x}_t, \tag{3}$$

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 structured [24] and time invariant over the period of observation, and 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. Here, we do not focus on the K common factors and their sensitivities ($\mathbf{B}_t$), but they are indirectly linked to the K eigenvectors related to the K largest eigenvalues as detailed in section 3.

2.3. The case of non-homogeneous assets returns

We propose in this section to re-write model (3) for the m assets splitted into $p<m$ groups. Each group is 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. It follows that:

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

The complete scatter matrix $\mathbf{C}$ is therefore block-constructed, block-Toeplitz, and the groups are assumed to be uncorrelated each others.

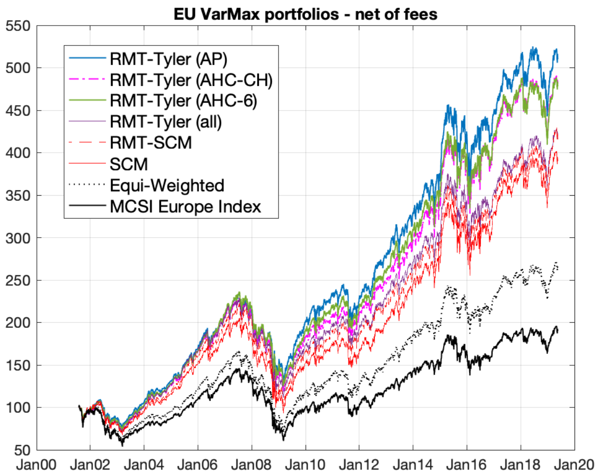

Fig. 1. EU VarMax portfolios' wealth with 0.07% of fees from July 2001 to May 2019.

3. PROPOSED METHODOLOGY

In this section our methodology is presented: we describe the methods used to group the assets, the Tyler M-estimator and finally we detail the whole whitening procedure.

3.1. Assets classification

Under the assumption of non-homogeneous assets returns, we propose to form groups of assets before applying the whitening process. We propose here to use the Affinity Propagation algorithm (AP) [21] that does not require to specify the number of clusters, and the classical Ascending Hierarchical Classification (AHC).

3.1.1. Affinity Propagation algorithm (AP)

The Affinity Propagation algorithm (AP) [21] is an iterative partitioning method similar to the K-means, but instead of regrouping individuals around central values, AP algorithm regroups them around exemplar values. The algorithm is based on a similarity matrix $\mathbf{S}$, where $s_{i,j} = -\|\mathbf{v}_i - \mathbf{v}_j\|^2$ for $i \neq j$, and with $\mathbf{v}_i$ and $\mathbf{v}_j$ the input variables vectors of the asset i and j. The number of groups is influenced by the main diagonal of $\mathbf{S}$ ($s_{i,i}\, \forall i \in [1, m]$) also called "preferences" parameters. In order to moderate the number of groups p, the parameters are set to a common value using the median of pairwise similarities as in [21].

3.1.2. Ascending Hierarchical Classification (AHC)

The classical Ascending Hierarchical Classification (AHC) is an iterative and unsupervised method. The algorithm is based on the distances between the variables $(\mathbf{v}_i)_{i\in[1,m]}$ used to represent individuals to be grouped and seeks at each step to build the groups by aggregation. AHC ensures to get homogeneous groups for which the intra-group variances are smaller than the inter-group variances. We use AHC with the Euclidean distance and the Ward measure [25] to form the p groups. The number of groups p is determined arbitrary or with Caliński-Harabasz (CH) criterion [20].

3.2. The robust Tyler's M-estimator

Under general non-Gaussian noise hypothesis proposed in Section 2.2, Tyler M-estimator [5, 11] 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:

$$\mathbf{X} = \frac{m}{N} \sum_{t=1}^{N} \frac{\mathbf{r}_t\, \mathbf{r}_t^T}{\mathbf{r}_t^T\, \mathbf{X}^{-1}\, \mathbf{r}_t}, \tag{5}$$

with $Tr(\widehat{\mathbf{C}}_{tyl}) = m$. The scatter matrix, solution of (5) has some remarkable properties [26, 27] 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 sources are present in the observations $\{\mathbf{r}_t\}$, the direct use of this estimator (contrary to the SCM estimate) may lead to whiten the observations and to slightly destroy the main information concentrated in the K factors. According to the *consistency theorem* found and proved in [17, 28], the problem can be solved through a biased Toeplitz estimate of $\widehat{\mathbf{C}}_{tyl}$, let us say $\widetilde{\mathbf{C}}_{tyl} = \mathcal{T}\left(\widehat{\mathbf{C}}_{tyl}\right)$. This 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.

3.3. Detailed whitening procedure

Given $\mathbf{R}$ the $m \times N$-matrix of observations, and $\mathbf{R}^{(q)}$ the $m_q \times N$-matrix of observations for group (q), the de-noised covariance matrix estimate $\widehat{\mathbf{\Sigma}}_w$ is obtained through the following procedure steps:

- Compute the p groups using the methods described in 3.1 with $(\mathbf{v}_i)_{i\in[1,m]}$ composed of the mean μ_i, the standard deviation σ_i and of several quantiles computed from $\tilde{\mathbf{r}}_i = (\mathbf{r}_i - \mu_i\, \mathbf{1}_N)\,/\sigma_i$ the "standardized" returns, where $\mathbf{1}_N$ is the N-vector of ones,
- Set $\widehat{\mathbf{C}}_{tyl}^{(q)}$ the Tyler-M estimate of $\mathbf{R}^{(q)}$, solution of (5),
- Set $\widetilde{\mathbf{C}}_{tyl}^{(q)} = \mathcal{T}\left(\widehat{\mathbf{C}}_{tyl}^{(q)}\right)$, the Toeplitz rectification matrix built from $\widehat{\mathbf{C}}_{tyl}^{(q)}$ for the Toeplitz operator $\mathcal{T}$,
- Set $\mathbf{R}_w^{(q)} = \left(\widetilde{\mathbf{C}}_{tyl}^{(q)}\right)^{-1/2} \mathbf{R}^{(q)}$, the $m_q \times N$ matrix of the whitened observations of group q,
- Set $\widehat{\mathbf{\Sigma}}_{tyl}$ as the Tyler-M estimate of $\mathbf{R}_w$, solution of (5), where $\mathbf{R}_w = [\mathbf{R}_w^{(1)T} \ldots \mathbf{R}_w^{(p)T}]^T$ of size $m \times N$,

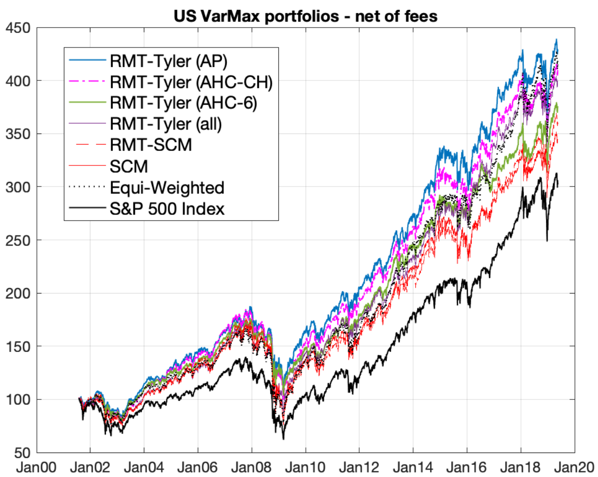

Fig. 2. US VarMax portfolios' wealth with 0.07% of fees from July 2001 to May 2019.

- Set $\widehat{\mathbf{\Sigma}}_{tyl}^{clip} = \mathbf{U}\,\mathbf{\Lambda}^{clip}\,\mathbf{U}^T$ where $\mathbf{U}$ is the $m \times m$ eigenvectors matrix and $\mathbf{\Lambda}^{clip}$ is the $m \times m$ diagonal matrix of the eigenvalues $(\lambda_k^{clip})_{k \in [1,m]}$ of $\widehat{\mathbf{\Sigma}}_{tyl}$ corrected using the Eigenvalue clipping method [14],
- Finally, $\widehat{\mathbf{\Sigma}}_w = \left(\widetilde{\mathbf{C}}_{tyl}^{1/2}\right) \widehat{\mathbf{\Sigma}}_{tyl}^{clip} \left(\widetilde{\mathbf{C}}_{tyl}^{1/2}\right)^T$.

4. APPLICATION

This section is devoted to showing the benefits of using AP algorithm to form asset groups before applying the whitening process and estimating the whole covariance matrix in order to allocate a portfolio through the Maximum Variety process. Two investment universes are tested: the first one consists of European equity indices ($m = 43$) and the second one to US equity indices ($m = 30$). These indices represent industry sub-sectors (e.g. transportation or materials), factor-based indices (e.g. momentum or growth), and also countries (e.g. Sweden or France) for the European universe. We observe daily closing prices from July 27th, 2000 to May 20th, 2019, and 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 ($N = 260$) and we maximize the variety ratio (1) to obtain the vector of weights. The weights are kept constant for the next four-weeks period. When applicable, assets are classified either by the AP algorithm ("RMT-Tyler (AP)") or by AHC where the number of groups is set to $p = 6$ ("RMT-Tyler (AHC-6)") or set according to the CH criterion ("RMT-Tyler (AHC-CH)"). The quantiles used for the clustering algorithms are q_θ and $q_{1-\theta}$ with $\theta \in \{1\%, 2.5\%, 5\%, 10\%, 15\%, 25\%, 50\%\}$. These methods are compared to the whitening process applied on the whole universe ("RMT-Tyler (all)"), the Eigenvalue clipping method [14] ("RMT-SCM"), and the classical SCM ("SCM"). We also add for comparison the equally weighted portfolio and the respective benchmark for each universe (MSCI® Europe Index or S&P® 500 Index). The portfolio performances are net of transaction fees considering 0.07% of fees applied between two rebalancing dates in order to take into account the costs associated with buying or selling positions (turnover) as in [19]. Whitening returns within each group formed using either AHC or AP leads to outperforming conventional methods. The AP algorithm finds 7 groups on average and adapts the number of groups more dynamically than AHC-CH does, finding only 2 groups. Moreover, RMT-Tyler (AP) improves even more the results for the two universes, as shown in Figures 1 and 2. In Table 1 and Table 2, we report some portfolio statistics comparing: the annualized return (Ann. Ret.), the annualized volatility (Ann.Vol.), the ratio between the return and the volatility, the maximum drawdown and the average value of the diversification ratio. All the indicators related to the "RMT-Tyler (AP)" show a significant improvement with respect to the other methods: a higher annualized return, a lower volatility (higher return/volatility ratio), and a lower maximum drawdown for the European universe.

EU VarMax Portfolios	Ann. Ret.	Ann. Vol.	Ratio Ret/Vol	Max Drawdown	$\mathcal{VR}$ (avg)
RMT-Tyler (AP)	9.87%	12.14%	0.81	45.37%	1.46
RMT-Tyler (AHC-6)	9.65%	12.03%	0.80	46.84%	1.57
RMT-Tyler (AHC-CH)	9.58%	12.45%	0.77	48.16%	1.51
RMT-Tyler (all)	8.90%	13.16%	0.68	51.18%	1.44
RMT-SCM	8.94%	13.79%	0.65	54.15%	1.27
SCM	8.56%	13.68%	0.63	54.45%	1.38
Equi-Weighted	6.60%	15.37%	0.43	57.82%	1.19
MSCI Europe	*4.71%*	*14.87%*	*0.32*	*58.54%*	

Table 1. Performance numbers for the Europe (EU) VarMax portfolios with 0.07% of fees from July 2001 to May 2019.

US VarMax Portfolios	Ann. Ret.	Ann. Vol.	Ratio Ret/Vol	Max Drawdown	$\mathcal{VR}$ (avg)
RMT-Tyler (AP)	8.76%	11.11%	0.79	42.82%	1.51
RMT-Tyler (AHC-CH)	8.57%	11.53%	0.74	46.57%	1.55
RMT-Tyler (AHC-6)	7.98%	10.79%	0.74	41.50%	1.52
RMT-Tyler (all)	8.49%	12.09%	0.70	49.27%	1.53
Equi-Weighted	8.92%	13.83%	0.65	53.70%	1.25
RMT-SCM	8.03%	13.13%	0.61	56.53%	1.34
SCM	7.80%	13.27%	0.59	55.47%	1.46
S&P 500	*7.21%*	*14.18%*	*0.51*	*55.71%*	

Table 2. Performance numbers for the US VarMax portfolios with 0.07% of fees from July 2001 to May 2019.

5. CONCLUSION

This paper questioned the ability of classification methods (AP algorithm and AHC) to improve the estimation of the covariance matrix of financial assets using the Tyler M-estimator and the RMT. We test our methodology on the Maximum Variety portfolio optimization problem and prove the superiority of the AP algorithm to produce higher performances for both EU and US universes. The same improvements can be observed for the Minimum Variance Portfolio and are available upon request.

6. REFERENCES

[1] H. M. Markowitz, “Portfolio selection,” *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[2] S. Maillard, T. Roncalli, and J. Teiletche, “The properties of equally weighted risk contributions portfolios,” *Journal of Portfolio Management*, vol. 36, pp. 60–70, 2010.

[3] Y. Choueifaty and Y. Coignard, “Toward maximum diversification,” *Journal of Portfolio Management*, vol. 35, no. 1, pp. 40–51, 2008.

[4] Y. Choueifaty, T. Froidure, and J. Reynier, “Properties of the most diversified portfolio,” *Journal of investment strategies*, vol. 2, no. 2, pp. 49–70, 2013.

[5] D. E. Tyler, “A distribution-free M-estimator of multivariate scatter,” *The annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[6] R. A. Maronna, “Robust M-estimators of multivariate location and scatter,” *Annals of Statistics*, vol. 4, no. 1, pp. 51–67, January 1976.

[7] Y. Chen, A. Wiesel, and A. O. Hero, “Robust shrinkage estimation of high-dimensional covariance matrices,” *IEEE Trans. on Signal Processing*, vol. 59, no. 9, September 2011.

[8] F. Pascal, Y. Chitour, and Y. Quek, “Generalized robust shrinkage estimator and its application to STAP detection problem,” *IEEE Trans. on Signal Processing*, vol. 62, no. 21, November 2014.

[9] R. Couillet, F. Pascal, and J. W. Silverstein, “Robust estimates of covariance matrices in the large dimensional regime,” *IEEE Trans. on Information Theory*, vol. 60, no. 11, September 2014.

[10] L. Yang, R. Couillet, and M. R. McKay, “A robust statistics approach to minimum variance portfolio optimization,” *IEEE Trans. on Signal Processing*, vol. 63, no. 24, pp. 6684–6697, Aug 2015.

[11] F. Pascal, Y. Chitour, J. P. Ovarlez, P. Forster, and P. Larzabal, “Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis,” *IEEE Trans. on Signal Processing*, vol. 56, no. 1, pp. 34–48, Jan 2008.

[12] V. A. Marchenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices,” *Matematicheskii Sbornik*, 1967.

[13] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, “Noise dressing of financial correlation matrices,” *Physycal Review Letters*, vol. 83, no. 1468, 1999.

[14] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud, “Random Matrix Theory and financial correlations,” *International Journal of Theoretical and Applied Finance*, vol. 3, no. 03, pp. 391–397, 2000.

[15] M. Potters, J. P. Bouchaud, and L. Laloux, “Financial applications of Random Matrix Theory: old laces and new pieces,” *Acta Physica Polonica B*, vol. 36, no. 9, 2005.

[16] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, “Collective behavior of stock price movements: A Random Matrix Theory approach,” *Physica A*, vol. 299, pp. 175–180, 2001.

[17] E. Terreaux, J. P. Ovarlez, and F. Pascal, “A Toeplitz-Tyler estimation of the model order in large dimensional regime,” in *ICASSP*, Apr 2018.

[18] E. Jay, E. Terreaux, J. P. Ovarlez, and F. Pascal, “Improving portfolios global performance with robust covariance matrix estimation: Application to the maximum variety portfolio,” in *26th EUSIPCO*, Sept 2018.

[19] E. Jay, T. Soler, E. Terreaux, J. P. Ovarlez, F. Pascal, P. De Peretti, and C. Chorro, “Improving portfolios global performance using a cleaned and robust covariance matrix estimate,” DOI:10.13140/RG.2.2.35800.44804, May 2019.

[20] T. Caliński and J. Harabasz, “A dendrite method for cluster analysis,” *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974.

[21] B. J. Frey and D. Dueck, “Clustering by passing messages between data points,” *Science*, vol. 315, no. 5814, pp. 972–976, 2007.

[22] D. Kelker, “Distribution theory of spherical distributions and a location-scale parameter generalization,” *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 32, no. 4, pp. 419–430, 1970.

[23] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, “Complex Elliptically Symmetric distributions: Survey, new results and applications,” *IEEE Trans. on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, Nov 2012.

[24] R. M. Gray, “Toeplitz and circulant matrices: A review,” *Foundations and Trends® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[25] J. H. Jr. Ward, “Hierarchical grouping to optimize an objective function,” *Journal of the American Statistical Association*, vol. 58, pp. 236–244, 1963.

[26] F. Pascal, P. Forster, J. P. Ovarlez, and P. Larzabal, “Performance analysis of covariance matrix estimates in impulsive noise,” *IEEE Trans. on Signal Processing*, vol. 56, no. 6, pp. 2206–2217, June 2008.

[27] M. Mahot, F. Pascal, P. Forster, and J. P. Ovarlez, “Asymptotic properties of robust complex covariance matrix estimates,” *IEEE Trans. on Signal Processing*, vol. 61, no. 13, pp. 3348–3356, July 2013.

[28] E. Terreaux, J. P. Ovarlez, and F. Pascal, “New model order selection in large dimension regime for Complex Elliptically Symmetric noise,” in *25th EUSIPCO*, Aug 2017, pp. 1090–1094.