A TOEPLITZ-TYLER ESTIMATION OF THE MODEL ORDER IN LARGE DIMENSIONAL REGIME.

Eugénie Terreaux [*1], *Jean-Philippe Ovarlez*[1,2], *Frédéric Pascal*[3]

[1] CentraleSupélec/SONDRA, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France
[2] ONERA, DEMR/TSI, Chemin de la Hunière, F-91120 Palaiseau, France
[3] L2S/CentraleSupélec-Université Paris-Sud, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France

ABSTRACT

This paper presents a new algorithm to estimate the number of sources embedded in a correlated Complex Elliptically Distributed (CES) noise in the context of large dimensional regime. The proposed method is a two-steps ones: first the data covariance matrix is estimated with a robust and consistent estimator exploiting the Toeplitz structure assumption of the true scatter matrix. Then, after whitening the signal thanks to the first estimator, the distribution of its Tyler estimator eigenvalues is studied, as in [1]. This allows to derive a threshold, estimated thanks to the Marchenko-Pastur law, to separate the eigenvalues corresponding to the noise and those corresponding to the sources. The number of sources can therefore be deduced. The proposed method is compared to classical ones as the Akaike Information Criterion (AIC) or other algorithms recently developed.

1. INTRODUCTION

The problem of estimating the number of sources, or equivalently the model order, is a recurrent and challenging issue in a lot of signal processing fields (see e.g., [2], [3] or [4]). To that end, statistical methods often rely on the estimation of the covariance matrix eigenvalues. This is the case for instance for the Akaike Information Criterion (AIC) [5] or others methods of model order selection [6]. Thereby, with such methods, the better the covariance matrix estimation is, the better the number of sources estimation is. In this context of number of sources estimation, this article introduces new results when the observations turn to be non-Gaussian and of large dimension.

Large dimensional regime corresponds to the following regime: the number of received signals N and the dimension m tend to infinity at the same rate (see e.g., [7]). For white or whitened noise, several methods, based on the Random Matrix Theory (RMT) have been proposed to extract information of interest from the received signals. One can cite for instance the number of embedded sources estimation [8], the problem of radar detection [9], signal subspace estimation [10]. 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 [6, 11] while the others assume that the covariance matrix is known and use it to whiten the signal.

Concerning non-Gaussian signals, robust methods have been recently developed based on Complex Elliptically Symmetric (CES) distribution. CES distributions belong to a wide-class of distributions including the Gaussian distribution, the K-distribution and others, as detailed in [12]. Its flexibility enables to model a large panel of random signals and justifies its wide use in signal processing applications for example in communication applications. To deal with non-Gaussian data, robust methods adapted to large dimensional observations [13] have been developed in the case of white (or perfectly whitened) noise.

To fill this gap, [14] proposed a method that relaxes this strong assumption thanks to a Toeplitz-based correlation model. Moreover, this methodology is based on Maronna's assumptions [15], that is to say, for well-behaved weighting functions $u(.)$ appearing in the definition of the covariance M-estimators. For model order selection, the threshold proposed in [14] depends on the selected function $u(.)$ and remains quite difficult to derive. This is the case for instance for the Huber [16] M-estimators.

This article expands former results of [14] to the case of Tyler [17] M-estimators for correlated CES noise. Note that Tyler M-estimators do not respect Maronna's conditions but provide more flexibility since it is a distribution-free estimator [18]. Moreover, recent works in [1] proved that the eigenvalues distribution of the Tyler M-estimator follows, in the white CES case, the Marchenko-Pastur distribution [19]. The main contribution of this article is first to propose a consistent estimator of the covariance matrix for a correlated CES noise. Then, it aims at providing an estimation of the number of sources when the signal is constituted of a correlated CES noise and some additive sources.

This estimation process is shown to be easier than those proposed in [14]. The proposed method splits into two steps

as in [14]: the first one is the whitening of the signal with a good estimated covariance matrix. This estimated whitening covariance matrix is obtained thanks to a Tyler estimator and a Toeplitz rectification. The second step consists in estimating the eigenvalues of the Tyler covariance matrix of the whitened signal and then, it exploits the support of the Marchenko-Pastur law to find the number of sources in the signal. These steps are described in the two first parts of this article. The third part shows a comparison in terms of performance between the proposed method, the method proposed in [14] and AIC through some simulations.

Notations: Matrices are in bold and capital, vectors in bold. Let $\mathbf{X}$ be a square matrix of size $s \times s$, $(\lambda)_i(\mathbf{X})$, $i \in]\!]1, ..., s[\![$, are the eigenvalues of $\mathbf{X}$. $Tr(\mathbf{X})$ is the trace of the matrix $\mathbf{X}$. $\|\mathbf{X}\|$ stands for the spectral norm. Let $\mathbf{A}$ be a matrix, $\mathbf{A}^T$ is the transpose of $\mathbf{A}$ and $\mathbf{A}^H$ the Hermitian transpose of $\mathbf{A}$. $\mathbf{I}_n$ is the $n \times n$ identity matrix. For any $m-$vector $\mathbf{x}$, $\mathcal{L} : \mathbf{x} \mapsto \mathcal{L}(\mathbf{x})$ is the $m \times m$ matrix defined as the Toeplitz operator: $([\mathcal{L}(\mathbf{x})]_{i,j})_{i \leq j} = x_{i-j}$ and $([\mathcal{L}(\mathbf{x})]_{i,j})_{i>j} = x^*_{i-j}$. 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 $\check{a}_{i,\, 0<i<m-1}$ contains the sum of the $i-$th diagonal of $\mathbf{A}$ divided by m. For $x \in \mathbb{R}$, δ_x is the Dirac measure at x. supp is the support of a set. The notation $\xrightarrow{a.s.}$ means "tends to almost surely". The notation *dist* stands for the distance associated to the L_1 norm.

1.1. Model and Assumptions

The N observations $\mathbf{Y} = [\mathbf{y}_0, ..., \mathbf{y}_{N-1}]$ are made of a correlated CES noise added to p mixed sources. The equation (1) sets down the model:

$$\mathbf{y}_i = \sum_{j=1}^{p} a_{i,j}\, \mathbf{m}_j + \sqrt{\tau_i}\, \mathbf{C}^{1/2}\, \mathbf{x}_i, \quad i \in [\![0, N-1]\!], \quad (1)$$

where the $\{\tau_i\}_{i \in [\![0,N-1]\!]}$ are positive random variables and where $\mathbf{X} = [\mathbf{x}_0, ..., \mathbf{x}_{N-1}]$ is the matrix of independant multivariate m-vectors uniformly distributed on the sphere of dimension m. Moreover, the $m \times p$ matrix $\mathbf{M}$ with elements $M_{i,j} = (\mathbf{M})_{i,j} = (\mathbf{m}_j)_i$ is referred to as the mixing matrix and contains the p vectors of the sources. Each element $a_{i,j}$ of the $p \times N$ matrix $\mathbf{A}$ corresponds to the amplitude variation of each source in the received vector.

The matrix $\mathbf{C}$ is assumed to be a nonnegative definite Toeplitz matrix:

$$\mathbf{C} = \begin{pmatrix} c_0 & c_1 & ... & c_{m-1} \\ c_1^* & c_0 & ... & c_{m-2} \\ ... & & & \\ c_{m-1}^* & c_{m-2}^* & ... & c_0 \end{pmatrix}.$$

The signal can thus be written as:

$$\mathbf{Y} = \mathbf{M}\,\mathbf{A} + \mathbf{C}^{1/2}\,\mathbf{X}\,\mathbf{T}^{1/2}, \quad (2)$$

where $\mathbf{T}$ is the $N \times N$-diagonal matrix containing the $\{\tau_i\}_{i \in [\![0,N-1]\!]}$. In the sequel, we will consider the following assumptions: first, $N \to \infty$, $m \to \infty$ and $c_N = \frac{m}{N} \to c > 0$.

Then $\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\tau_i}$ satisfies $\int \tau \mu_N(d\tau) \to 1$ almost surely and $\frac{1}{N} \sum \delta_{\lambda_i(\mathbf{C})}$ converges almost surely toward the true measure ν. We will also assume that $\max_i dist(\lambda_i(\mathbf{C}), \text{supp}(\nu)) \to 0$ and that $\{c_k\}_{k \in [\![0,m-1]\!]}$ are absolutely summable coefficients, such that $c_0 \neq 0$.

The problem considered here is the estimation of p, the number of sources. Next section presents the main contribution of this article.

2. RESULTS

2.1. Signal Whitening

The first step of the proposed algorithm is to whiten the signal with a consistent estimation of $\mathbf{C}$, the scatter matrix of the CES noise. The proposed estimator is a Tyler M-estimator of the scatter matrix $\hat{\mathbf{C}}$ enforced to be Toeplitz-structured with the operator $\mathcal{T}(\hat{\mathbf{C}})$. The Tyler M-estimator $\hat{\mathbf{C}}$ is defined as the unique solution of [20]:

$$\boldsymbol{\Sigma} = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{y}_i\mathbf{y}_i^H}{\mathbf{y}_i^H\boldsymbol{\Sigma}^{-1}\mathbf{y}_i}. \quad (3)$$

This equation requires an iterative approach to compute $\hat{\mathbf{C}}$ as the final value of $\boldsymbol{\Sigma}$. Note that the $\mathbf{y}_i$ contain also the sources. The consistency of $\mathcal{T}(\hat{\mathbf{C}})$ is proven thanks to the following theorem:

Theorem 1 (A consistent estimation of $\mathbf{C}$). *Under the prior assumptions, we have the following convergence:*

$$\left\| \mathcal{T}(\hat{\mathbf{C}}) - \mathbf{C} \right\| \xrightarrow{a.s.} 0. \quad (4)$$

Proof. Let $\hat{\mathbf{S}}$ be the Sample Covariance Matrix (SCM) that is $\hat{\mathbf{S}} = \frac{1}{N} \mathbf{X}\mathbf{X}^H$. Then:

$$\left\| \mathcal{T}(\hat{\mathbf{C}}) - \mathbf{C} \right\| \leq \left\| \mathcal{T}(\hat{\mathbf{C}}) - \mathcal{T}(\hat{\mathbf{S}}) \right\| + \left\| \mathcal{T}(\hat{\mathbf{S}}) - \mathbf{C} \right\| \quad (5)$$

Moreover, it is proven in [14] that $\left\| \mathcal{T}(\hat{\mathbf{S}}) - \mathbf{C} \right\| \xrightarrow{a.s.} 0$. Since [1] shows that $\left\| \hat{\mathbf{C}} - \hat{\mathbf{S}} \right\| \xrightarrow{a.s.} 0$ and as $\mathcal{T}$ is a linear operator, it appears that $\left\| \mathcal{T}(\hat{\mathbf{C}}) - \mathcal{T}(\hat{\mathbf{S}}) \right\|$ and that $\left\| \mathcal{T}(\hat{\mathbf{C}}) - \mathbf{C} \right\| \xrightarrow{a.s.} 0$. $\square$

The consistency of this estimator is shown on Figure 1 where the $\{\tau\}_{i \in [\![1, N]\!]}$ are inverse gamma distributed. The

rate of the convergence is relatively slow but it is possible to enhance it with banding or tapering (see for example [21]).
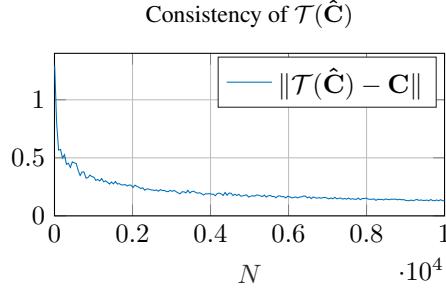

Fig. 1. Consistency of the proposed estimator of $\mathbf{C}$, $N \in [\![5, 10000]\!]$, $c = 0.7$, $\{\tau\}_{i \in [\![1, N]\!]}$ following an inverse gamma distribution

Let us note $\check{\mathbf{C}}$ the estimator $\mathcal{T}\left(\hat{\mathbf{C}}\right)$. The signal can be whitened as following:

$$\hat{\mathbf{Y}}_w = \check{\mathbf{C}}^{-1/2}\mathbf{Y} = \check{\mathbf{C}}^{-1/2}\mathbf{M}\mathbf{A} + \check{\mathbf{C}}^{-1/2}\mathbf{C}^{1/2}\mathbf{X}\mathbf{T}^{1/2}, \quad (6)$$

with $\hat{\mathbf{Y}}_w = [\hat{\mathbf{y}}_{w1}, ..., \hat{\mathbf{y}}_{wN}]$.

2.2. Estimation

The signal being whitened, it is now possible to apply a Tyler M-estimator in order to threshold its eigenvalues. Let $\hat{\mathbf{W}}$ denote a Tyler estimation of the "white" scatter matrix, that is, $\hat{\mathbf{W}}$ is the unique solution if it exists of:

$$\mathbf{\Sigma}_w = \frac{m}{N} \sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{wi} \hat{\mathbf{y}}_{wi}^H}{\hat{\mathbf{y}}_{wi}^H \mathbf{\Sigma}_w^{-1} \hat{\mathbf{y}}_{wi}}. \quad (7)$$

Let us define $\hat{\mathbf{S}}_w = \frac{1}{N}\mathbf{X}\mathbf{X}^H$. In order to threshold its eigenvalues and estimate the number of sources, Theorem 2 enables to choose the threshold as the right edge of the Marchenko-Pastur law support [19]:

Theorem 2 (Convergence of $\hat{\mathbf{W}}$). *Under same assumptions as previously,*

$$\left\|\hat{\mathbf{W}} - \hat{\mathbf{S}}_w\right\| \xrightarrow{a.s.} 0. \quad (8)$$

Proof. It follows the one of Theorem 2 in [14]. Let $\check{\mathbf{W}}$ denote the Tyler M-estimator of $\mathbf{Y}_w = \mathbf{C}^{-1/2}\mathbf{M}\mathbf{A} + \mathbf{X}\mathbf{T}^{1/2}$ with $\mathbf{Y}_w = [\mathbf{y}_{w1}, ..., \mathbf{y}_{wN}]$, i.e., $\check{\mathbf{W}} = \frac{m}{N}\sum_{i=1}^{N} \frac{\mathbf{y}_{wi}\mathbf{y}_{wi}^H}{\mathbf{y}_{wi}^H \check{\mathbf{W}}^{-1} \mathbf{y}_{wi}}$. Since $\hat{\mathbf{W}} = \frac{m}{N}\sum_{i=1}^{N} \frac{\hat{\mathbf{y}}_{wi}\hat{\mathbf{y}}_{wi}^H}{\hat{\mathbf{y}}_{wi}^H \hat{\mathbf{W}}^{-1} \hat{\mathbf{y}}_{wi}}$, if we replace $\hat{\mathbf{y}}_{wi}$ by $\check{\mathbf{C}}^{-1/2}\mathbf{C}^{1/2}\mathbf{y}_{wi}$, $0 \le i \le N$, one obtains:

$$\mathbf{C}^{-1/2}\check{\mathbf{C}}^{1/2}\hat{\mathbf{W}}\check{\mathbf{C}}^{1/2}\mathbf{C}^{-1/2} = \quad (9)$$
$$\frac{m}{N}\sum_{i=1}^{N} \frac{\mathbf{y}_{wi}\mathbf{y}_{wi}^H}{\mathbf{y}_{wi}^H \left(\mathbf{C}^{-1/2}\check{\mathbf{C}}^{1/2}\hat{\mathbf{W}}\check{\mathbf{C}}^{1/2}\mathbf{C}^{-1/2}\right)^{-1} \mathbf{y}_{wi}},$$

which leads to:

$$\hat{\mathbf{W}} = \check{\mathbf{C}}^{-1/2}\mathbf{C}^{1/2}\check{\mathbf{W}}\mathbf{C}^{1/2}\check{\mathbf{C}}^{-1/2}. \quad (10)$$

The equation (8) can be rewritten as:

$$\left\|\hat{\mathbf{W}} - \hat{\mathbf{S}}_w\right\| \le \left\|\hat{\mathbf{W}} - \check{\mathbf{W}}\right\| + \left\|\check{\mathbf{W}} - \hat{\mathbf{S}}_w\right\|. \quad (11)$$

Moreover, it is proven in [1] that $\left\|\check{\mathbf{W}} - \hat{\mathbf{S}}_w\right\| \xrightarrow{a.s.} 0$. Concerning the first term, it can be rewritten as:

$$\left\|\hat{\mathbf{W}} - \check{\mathbf{W}}\right\| \le \left\|\check{\mathbf{C}}^{-1/2}\mathbf{C}^{1/2}\check{\mathbf{W}}\mathbf{C}^{1/2}\check{\mathbf{C}}^{-1/2} - \check{\mathbf{W}}\mathbf{C}^{1/2}\check{\mathbf{C}}^{-1/2}\right\|$$
$$+ \left\|\check{\mathbf{W}}\mathbf{C}^{1/2}\check{\mathbf{C}}^{-1/2} - \check{\mathbf{W}}\right\|. \quad (12)$$

After left and right factorization, one obtains:

$$\left\|\hat{\mathbf{W}} - \check{\mathbf{W}}\right\| \le \left\|\check{\mathbf{C}}^{-1/2}\mathbf{C}^{1/2} - \mathbf{I}_m\right\| \|\check{\mathbf{W}}\| \left(\left\|\mathbf{C}^{1/2}\check{\mathbf{C}}^{-1/2}\right\| + 1\right).$$

As $\|\mathbf{C}\|$ has a bounded support, $\|\check{\mathbf{C}}\|$ has a bounded support too thanks to Theorem 1. Moreover, this theorem proves the consistency as $\left\|\mathbf{C} - \check{\mathbf{C}}\right\| \xrightarrow{a.s.} 0$ and this ensures the proof. $\square$

Thereby, as the distribution of the eigenvalues of $\hat{\mathbf{S}}_w$ converges in law toward the Marchenko-Pastur distribution [19], the scatter matrix eigenvalues distribution converges in law to the Marchenko-Pastur law, if some sources have sufficiently high SNR, it is possible to detect them. The right edge of the Marchenko-Pastur law, that is $(1 + \sqrt{c})^2$, can be chosen to threshold the eigenvalue distribution of $\hat{\mathbf{W}}$. The number of eigenvalues upon the threshold corresponds to the number of sources in the signal, and is denoted by $\hat{p}$.

2.3. Simulations

In order to enlighten these theoretical results, Figure 2 shows that the eigenvalues distribution of the estimated scatter matrix $\hat{\mathbf{W}}$ follows the Marchenko-Pastur distribution. On this figure, the $\{\tau\}_{i \in [\![1, N]\!]}$ are inverse gamma distributed. The 4 sources with SNR=10 dB present in the signal can be detected since 4 eigenvalues are found upon the Marchenko-Pastur distribution support, even if the SNR is low.

Then, Figure 3 shows the eigenvalues distribution with the same parameters as in Figure 2 but for a Fixed Point estimation of $\mathbf{C}$ noted FP, directly defined as the unique solution of $\mathbf{\Sigma} = \frac{1}{N}\sum_{i=1}^{N} u\left(\mathbf{y}_i^H \mathbf{\Sigma}^{-1} \mathbf{y}_i\right) \mathbf{y}_i \mathbf{y}_i^H$ build with the function $u(x) = \frac{1+t}{t+x}$, $t = 0.1$ and compared to the Marchenko-pastur law. As expected, this distribution does not follow the Marchenko-Pastur law and the chosen threshold is no more relevant. To find a correct threshold for this estimator, it is necessary to whiten the signal first and then calculate a more complicated threshold, without explicit expression, as deteiled in [14].
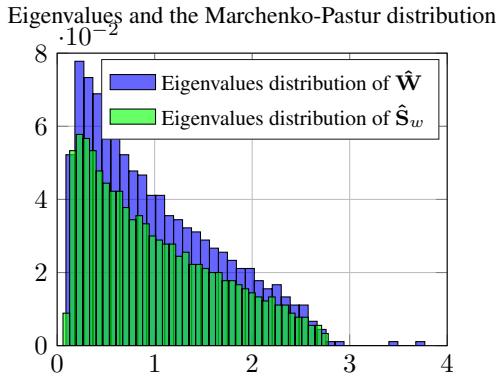
Fig. 2. The Marchenko-Pastur distribution and the eigenvalues distribution of $\hat{\mathbf{W}}$, 4 sources, SNR=10 dB, $N = 2000$, $m = 900$, $c = 0.7$, $\{\tau\}_{i\in[\![1,\,N]\!]}$ are inverse gamma distributed.
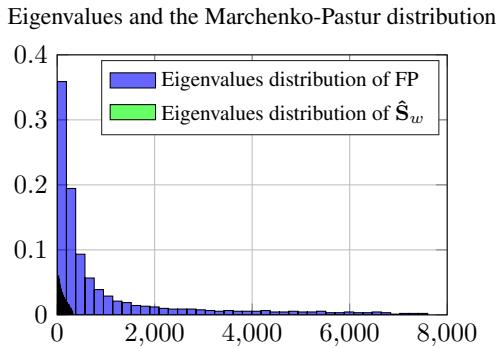

Fig. 3. 4 sources, $N = 2000$, $m = 900$, $\{\tau\}_{i\in[\![1,\,N]\!]}$ are inverse-gamma distributed.
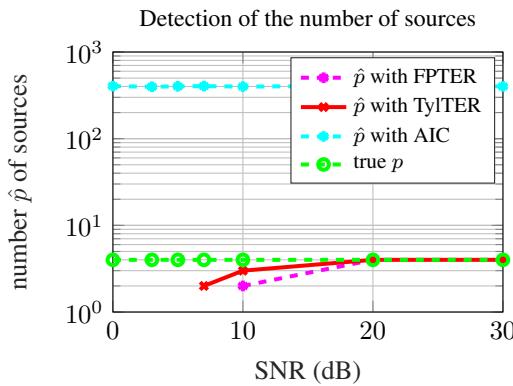

Fig. 4. 4 sources, $N = 2000$, $m = 900$, $\{\tau\}_{i\in[\![1,\,N]\!]}$ are inverse-gamma distributed.
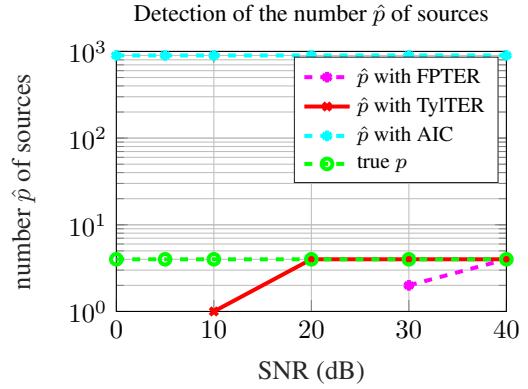

Fig. 5. 4 sources, $N = 2000$, $m = 900$, $\{\tau\}_{i\in[\![1,\,N]\!]}$ are Student-t distributed.

The number $\hat{p}$ of sources detected in the signal is plotted on Figure 4, for different SNR. Our method, denoted TylTER, is compared to the method of [14] noted FPTER, and to the AIC method. The true number of sources is also plotted on the Figure. The scale is in logarithm. The AIC method overestimates the number of sources as our method finds the true number of sources from a SNR= 20dB. The method FPTER begins to find sources for higher SNR than TylTER. Other criterions adapted to colored noise give results similar to the AIC method, as for exemple the method of [22].

On Figure 5 are drawn the same results but for the $\{\tau\}_{i\in[\![1,\,N]\!]}$ Student-t distributed. As expected, the proposed method still detects the sources for correct SNR. The results are almost the same except that the SNR required for detecting sources is higher for the methods TylTER and FPTER. Moreover, the FPTER method detects sources for a higher SNR than for the distribution of the τ_i in Figure 4. Indeed, in Figure 4, the chosen function u for FPTER is the optimal one for the τ_i. As the τ_i are slightly different but the function u is the same, the method FPTER is still robust but need higher SNR than TylTER to detect all the sources. These results are very promising and encourage the application of the proposed method to a large panel of signals.

3. CONCLUSION

A new robust algorithm has been proposed in order to estimate the number of sources in a signal. The algorithm consists first in whitening the signals and then in estimating the scatter matrix eigenvalues distribution. Simulations show the efficiency of the method, even if the noise is different from what expected. Two major results of this paper are first that this algorithm is derived for correlated CES noise in large dimensional regime and secondly that it proves it is no longer necessary to determine a complicated threshold as in [14]: the right edge of the Marchenko-Pastur law support can simply be used. Thereby, independently of the τ_i distribution, the proposed threshold does not require any a priori on the noise.

4. REFERENCES

[1] T. Zhang, C. Xiuyuan, and A. Singer, "Marchenko-pastur law for tyler's and maronna's m-estimators," *arXiv preprint*, 2014.

[2] J. Vinogradova, R. Couillet, and W. Hachem, "A new method for source detection, power estimation, and localization in large sensor networks under noise with unknown statistics," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 3943–3946.

[3] N. Arkind and B. Nadler, "Parametric joint detection-estimation of the number of sources in array processing," in *2010 IEEE Sensor Array and Multichannel Signal Processing Workshop*, Oct 2010, pp. 269–272.

[4] B. Ottersten, M. Viberg, and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Transactions on Signal Processing*, vol. 40, no. 3, pp. 590–600, March 1992.

[5] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, Dec 1974.

[6] J. Vinogradova, R. Couillet, and W. Hachem, "Statistical inference in large antenna arrays under unknown noise pattern," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5633–5645, Nov 2013.

[7] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*, Cambridge University Press, 2011.

[8] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3930–3941, Oct 2009.

[9] F. Pascal and A. Kammoun, "Théorie des matrices aléatoires robustes et applications à la détection radar," *Traitement du Signal*, 2016.

[10] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet, "A subspace estimator for fixed rank perturbations of large random matrices," *Journal of Multivariate Analysis*, vol. 114, pp. 427–447, 2013.

[11] Romain Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *Journal of Multivariate Analysis*, vol. 140, pp. 139 – 161, 2015.

[12] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, Nov 2012.

[13] R. Couillet, F. Pascal, and J. W. Silverstein, "The random matrix regime of maronna's M-estimator with elliptically distributed samples," *Journal of Multivariate Analysis*, 2015.

[14] E. Terreaux, J. P. Ovarlez, and F. Pascal, "Robust model order selection in large dimensional elliptically symmetric noise," *arXiv preprint*, 2017.

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

[16] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[17] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, March 1987.

[18] E. Ollila and D. Tyler, "Distribution-free detection under complex elliptically symmetric clutter distribution," in *Proc. IEEE 7th Sensor Array Multichannel Signal Process. Workshop (SAM)*, June 2012, pp. 413–416.

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

[20] 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 Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, Jan 2008.

[21] T. T. Cai, Z. Ren, and H. H. Zhou, "Optimal rates of convergence for estimating Toeplitz covariance matrices," *Probability Theory and Related Fields*, 2013.

[22] Emanuel Radoi and André Quinquis, "A new method for estimating the number of harmonic components in noise with application in high resolution radar," *EURASIP J. Appl. Signal Process.*, vol. 2004, pp. 1177–1188, Jan. 2004.