New Model Order Selection in Large Dimension Regime for Complex Elliptically Symmetric Noise

Eugénie Terreaux*, Jean-Philippe Ovarlez*†, and Frédéric Pascal‡

*CentraleSupélec-SONDRA, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France
e-mail: eugenie.terreaux@centralesupelec.fr
† ONERA, DEMR/TSI, Chemin de la Hunière, 91120 Palaiseau, France and SONDRA/CentraleSupélec
e-mail: jean-philippe.ovarlez@onera.fr
‡L2S/CentraleSupélec-CNRS-Université Paris-Sud, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France
e-mail: frederic.pascal@lss.centralesupelec.fr

***Abstract*—This paper presents a new model order selection technique for signal processing applications related to source localization or subspace orthogonal projection techniques in large dimensional regime (Random Matrix Theory) when the noise environment is Complex Elliptically Symmetric (CES) distributed, with unknown scatter matrix. The proposed method consists first in estimating the Toeplitz structure of the background covariance matrix. In a second step, after a whitening process, the eigenvalues distribution of any Maronna's M-estimators is exploited, leading to the order selection. Simulations made on different kinds of CES noise as well as analysis of real hyperspectral images demonstrate the superiority of the proposed technique compared to those of Akaike Information Criterion and the Minimum Description Length.**

I. INTRODUCTION

In signal processing, covariance matrices often contain information related to signals of interest. For example in source localization [1], the estimation of the direction of arrival, *i.e.* the estimation of the signal subspace, can be done thanks to the estimation and the exploitation of the covariance matrix. When sources are polluted with noise coming from the sensors, the channel, or other disturbances, the estimation of the covariance matrix is degraded. In the example of N sensors detecting signals of size m, the estimation of the covariance matrix is also degraded if the dimensions N and m go to infinity with a fixed ratio [2]. And this situation arises frequently since sensors are developed to be more accurate and numerous: this search of accuracy leads to an accumulation of large data that need to be processed. Model order selection for detecting the number of sources in a multichannel time-series or for estimating the rank of a subspace remains therefore an important problem in statistical signal processing as in security, medical screening, astronomy, finance, communication and other scientific fields. Beyond the problem of estimation of the signal subspace, the covariance matrix is also used in detection [3], or filtering [4]. Thus, developing more efficient tools and techniques is relevant.

All these related techniques are commonly based on the application of the information theoretic criteria for model order selection such as the Akaike Information Criterion (AIC) [5] or the Minimum Description Length (MDL) [6], [7] and [8] for gaussian processes. Since many techniques are based on the eigenvalue-decomposition or the singular value decomposition of the collected data covariance matrix, Wax and Kailath [9] have derived the eigenvalue forms of these two criteria which can be applied conveniently in array signal processing problems. In their seminal work as well as in most applications, the additive noise process is assumed to be spatially and temporally white Gaussian random process. These methods are shown to fail when the noise is not white Gaussian distributed or even non-Gaussian. When the dimension of the observation is growing, these methods also give very bad performances [10].

In a lot of fields such as hyperspectral imaging which deals with multivariate data of large dimension, data are generally compressed or projected before being processed [11]. The Random Matrix Theory (RMT), recently developed in signal processing (see [12] for one of the first use of RMT in a signal processing field), provides some useful properties or attributes to handle with these kind of matrices [13] without using dimensionality reduction. This theory proposes, among others, methods to estimate the distribution of eigenvalues for large matrices [14] or for mixed-model with a signal composed of few sources and additive noise [15] or [16].

This article first proposes a new estimator for the covariance matrix when the noise is correlated and non-Gaussian. This estimator is developed here in the Complex Elliptical Symmetric noise (CES) context [17], which is a better characterization for the noise in a lot of applications (for example in hyperspectral imaging [18]). For non-RMT processes, Toeplitz matrices have been widely studied for example in [19], or [20]. But herein they are exploited in a large dimension regime. The proposed method is decomposed in two parts and presented in two different sections. A first section sets the chosen model, a second presents the beginning of the method, that is to "toeplitzify" the empirical estimation of the covariance matrix (the Sample Covariance Matrix (SCM)) and prove the consistency of this "toeplitzified" estimator compared to the true covariance matrix. This estimator is needed to white the data. In a last section of the second part, we propose to use a Maronna's M-estimator [21] to estimate the so-called scatter

ISBN 978-0-9928626-7-1 © EURASIP 2017

matrix (covariance matrix up to a factor) of the uncorrelated CES noise. This proposed robust estimator extends the field of application of the article of J. Vinogradova [22] to the one of the MDL and AIC methods. The end of this third section is devoted to the model order selection. Indeed a threshold can be applied on the eigenvalues of the covariance matrix. All eigenvalues greater than this threshold can be proved to be relative to sources due to RMT properties. A final part presents the relevance of this method on simulated signals and on real hyperspectral images.

Notation: Vectors are in bold and matrices in bold and capitals letters. Let $\mathbf{A}$ be a matrix, $\mathbf{A}^T$ and $\mathbf{A}^H$ are respectively the transpose and the Hermitian transpose of $\mathbf{A}$, $(\mathbf{A})_{i,j}$ is the (i,j)-th element of the matrix. If $\mathbf{A}$ is a square matrix of size $m \times m$ then $\{\lambda_i(\mathbf{A})\}_{i=1,\dots,m}$ are the eigenvalues of $\mathbf{A}$. $\mathbb{E}[x]$ is the statistical mean of the random variable x. $a.s.$ stands for the almost sure convergence. For any complex scalar a, $a^\star$ denotes its complex conjugate. $d_1(.)$ means the distance associated with the l_1-norm. The distribution δ denotes the Dirac measure, $\mathrm{supp}(.)$ the support of any measure and $\|.\|$ the spectral matrix norm. The Toeplitz matrix operator is acting on any vector $\mathbf{x}$ as $\mathcal{T} : \mathbf{x} \to \mathcal{T}(\mathbf{x})$ where $([\mathcal{T}(\mathbf{x})]_{i,j})_{i \leq j} = x_{i-j}$ and $([\mathcal{T}(\mathbf{x})]_{i,j})_{i>j} = x^*_{i-j}$.

II. Theoretical Aspects

A. Model and Assumption

Let us consider a set of N observations $\{\mathbf{y}_i\}_{i\in[1,N]}$ where each $\mathbf{y}_i$ is a multidimensional m-vector. In this article, we suppose here the usual random matrix regime, *i.e.* $N \to \infty$, $m \to \infty$ with the constant regime $c_N = \frac{m}{N} \to c$, $c > 0$.
The general model characterizing the presence of p sources corrupted with an additive Complex Elliptically Symmetric (CES) noise can be stated as the following set of binary hypothesis test:

$$\begin{cases} H_0 & \mathbf{y}_i = \sqrt{\tau_i}\,\mathbf{C}^{1/2}\,\mathbf{x}_i\,, \quad i \in [\![1,N]\!]\,, \\ H_1 & \mathbf{y}_i = \displaystyle\sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \sqrt{\tau_i}\,\mathbf{C}^{1/2}\,\mathbf{x}_i\,, \quad i \in [\![1,N]\!]\,, \end{cases} \tag{1}$$

where $\mathbf{m}_j$ are, for each observation i, the unknown m-steering vector of the j-th deterministic source with power $s_{i,j}$, where $\mathbf{x}_i$ is a multivariate zero-mean white noise of independent entries identically and uniformly distributed on the m-unit-sphere, where $\tau_1,\dots,\tau_N$ are positive scalar random texture variables and where $\mathbf{C}$ is a Hermitian Toeplitz covariance matrix defined as $\mathbf{C} = \mathcal{T}\left((c_0,\dots,c_{m-1})^T\right)$. For large random matrix regime, *i.e.* when $N \to \infty$, we suppose that $\mu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\tau_i}$ satisfies $\int \tau \mu_N(d\tau) \to 1$ almost surely, that $\frac{1}{N}\sum \delta_{\lambda_i(\mathbf{C})}$ converges almost surely toward the true measure ν and moreover, $\max_i d_1(\lambda_i(\mathbf{C}), \mathrm{supp}(\nu)) \to 0$ and that $\{c_k\}_{k\in[0,m-1]}$ are absolutely summable coefficients, such that $c_0 \neq 0$.

By denoting $\mathbf{Y} = [\mathbf{y}_1,\dots,\mathbf{y}_N]$ the $m \times N$-matrix containing all the observations, $\mathbf{X} = [\mathbf{x}_1,\dots,\mathbf{x}_N]$ the $m \times N$-matrix containing white noise, $\mathbf{T}$ the $N \times N$-matrix containing the $\{\tau_i\}_{i\in[1,N]}$ on its diagonal and zero elsewhere, $\mathbf{M} = [\mathbf{m}_1,\dots,\mathbf{m}_p]$ the $m \times p$-mixing matrix containing the p steering vectors corresponding to the sources and $\left(\mathbf{S}^T\right)_{i,j} = s_{i,j}$ the $N \times p$-matrix representing all the source power, we can write:

$$\mathbf{Y} = \mathbf{M}\,\mathbf{S} + \mathbf{R}\,,$$

where $\mathbf{R} = \mathbf{C}^{1/2}\,\mathbf{X}\,\mathbf{T}^{1/2}$ is the additive correlated CES noise.

B. Signal Whitening

The noise being correlated, we propose in this section a consistent estimator of the covariance matrix $\mathbf{C}$ built on the measurements $\mathbf{Y}$. We first analyze the hypothesis H_0, *i.e.* no-sources are contained in the measurements.
Let us define $\check{\mathbf{c}} = (\check{c}_0,\dots,\check{c}_{m-1})^T$ the vector built with elements given by averaging along each diagonal the signal-free Sample Covariance Matrix (SCM) $\mathbf{Y}\,\mathbf{Y}^H/N = \mathbf{R}\,\mathbf{R}^H/N$:

$$\check{c}_k = \frac{1}{m\,N}\sum_{i=1}^{m}\sum_{j=1}^{N} y_{i,j}\, y^\star_{i+k,j}\, \mathbb{1}_{1\leq i+k\leq m}\,, \tag{2}$$

where $k \in [\![0, m-1]\!]$ and where $\mathbb{1}_A$ is the indicator function on the set A. It can be noted that this averaging process consists in dividing each sum on each diagonal by m, leading to a so-called biased Toeplitz estimate.

Theorem 1 (Consistent estimator of $\mathbf{C}$)**.** *Under the same assumptions defined above, we have the following convergence:*

$$\|\mathcal{T}(\check{\mathbf{c}}) - \mathbb{E}[\tau]\,\mathbf{C}\| \to 0\,. \tag{3}$$

Up to an unknown scale factor $\mathbb{E}[\tau]$, *a consistent estimator of* $\mathbf{C}$ *is therefore given by* $\check{\mathbf{C}} = \mathcal{T}(\check{\mathbf{c}})$.

Proof. The proof follows the one in [22] and will not be done due to the lack of space. The proof relies on the Lemma 4.1 from [23], which sets an inequality between the l_2-norm of a matrix and the Fourier series of the correlation coefficients of the matrix (power spectral density). The left term of the theorem is cut in two parts, and we prove than each converges to zero. Once in the Fourier space, the steps of the proof are overall the same than in [22], with an additional term $\|\mathbf{T}\|$ found on the denominator of the upper bound of the inequality. □

The consistency of the proposed estimator can be shown in Figure 1. The chosen signal is a white Gaussian noise $\mathbf{X}$ correlated with a Toeplitz matrix $\mathbf{C} = \mathcal{T}\left((\rho^0, \rho^1, \dots, \rho^{m-1})^T\right)$ where $\rho = 0.7$. The texture $\{\tau_i\}_{i\in[1,N]}$ is randomly extracted from an inverse gamma distribution with mean equal to one. The corresponding noise is therefore Student-t-distributed. This figure presents the spectral norm (log scale) of the difference between the real covariance matrix $\mathbf{C}$ and respectively the proposed estimator in green, and the usual SCM estimator equal to $\frac{1}{N}\mathbf{R}\,\mathbf{R}^H$ in red when N varies from 20 to 2000 and $c = 0.2$ (20 Monte Carlo trials). The SCM is shown here not to be consistent. The estimator proposed in **Theorem 1** has

ISBN 978-0-9928626-7-1

a slow convergence toward the true covariance matrix but it converges toward it.

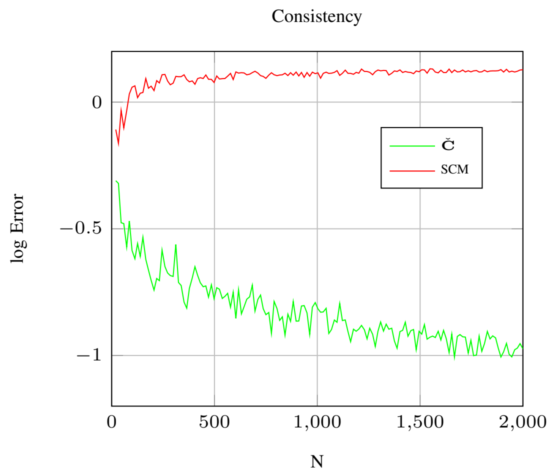

Fig. 1. Consistency of $\check{\mathbf{C}}$ estimator when $\mathbf{C} = \mathcal{T}\left(\left(\rho^0, \ldots, \rho^{m-1}\right)^T\right)$ with $\rho = 0.7$ and $c = 0.2$ for Student-t distributed m-vectors.

C. Signal Subspace Rank Estimation

In this section, the observations $\mathbf{Y}$ are whitened by the estimator $\check{\mathbf{C}}$ presented above. We note $\mathbf{Y}_w$ the whitened signal:

$$\mathbf{Y}_w = \check{\mathbf{C}}^{-1/2}\,\mathbf{Y} = \check{\mathbf{C}}^{-1/2}\,\mathbf{M}\,\mathbf{S} + \mathbf{R}_w\,, \tag{4}$$

where $\mathbf{R}_w = \check{\mathbf{C}}^{-1/2}\,\mathbf{C}^{1/2}\,\mathbf{X}\,\mathbf{T}^{1/2}$. Due to the fact that the observations are polluted by the textures $\{\tau_i\}_{i\in[1,N]}$, the number of sources can be estimated thanks to the distribution of the eigenvalues of any Maronna's M-estimator of the scatter matrix of the observations $\mathbf{Y}_w$. Under the H_0 hypothesis, the chosen robust estimator $\check{\mathbf{\Sigma}}$ is then defined as the unique solution, if it exists, of the equation:

$$\mathbf{\Sigma} = \frac{1}{N}\sum_{i=1}^{N} u\left(\frac{1}{m}\,\mathbf{r}_{wi}^{\;H}\,\mathbf{\Sigma}^{-1}\,\mathbf{r}_{wi}\right)\,\mathbf{r}_{wi}\,\mathbf{r}_{wi}^{\;H}\,, \tag{5}$$

where $\mathbf{R}_w = [\mathbf{r}_{w1}, \ldots, \mathbf{r}_{wN}]$ and $u : [0, +\infty) \to [0, +\infty)$ nonnegative, continue and non-increasing (see [24] for details). In order to evaluate the rank of the signal subspace, it is possible to set a threshold on the eigenvalues of $\check{\mathbf{\Sigma}}$ in a non-RMT regime. But it is not so easy to analyze the behavior of $\check{\mathbf{\Sigma}}$ eigenvalues when $N, m \to \infty$ using RMT classical tools since the term $u\left(\frac{1}{m}\,\mathbf{r}_{wi}^{\;H}\,\check{\mathbf{\Sigma}}^{-1}\,\mathbf{r}_{wi}\right)$ depends on $\mathbf{r}_{wi}$. So, we have to find and deal with another useful mathematical object having similar properties and behavior. Before proposing the following theorem, here are some definitions:

- $\phi : x \mapsto x\,u(x)$, increasing and bounded with $\lim_{x\to\infty}\phi(x) = \phi_\infty > 1$ where $\lim_{N\to\infty} c_N < \phi_\infty^{-1}$,
- $g : x \mapsto \dfrac{x}{1 - c_N\,\phi(x)}$, $v : x \mapsto u \circ g^{-1}(x)$,
- $\psi : x \mapsto x\,v(x)$ and γ_m is the unique solution (if it exists) of $\displaystyle\sum_{i=1}^{N} \frac{\psi(\tau_i\,\gamma)}{1 + c_N\,\psi(\tau_i\,\gamma)} = 1$.

Theorem 2 (Convergence of $\check{\mathbf{\Sigma}}$). *With the definitions given above, we have the following convergence*

$$\left\|\check{\mathbf{\Sigma}} - \hat{\mathbf{S}}\right\| \to 0 \;\; a.s. \tag{6}$$

where the matrix $\hat{\mathbf{S}}$ *is defined by:*

$$\hat{\mathbf{S}} \triangleq \frac{1}{N}\sum_{i=1}^{N} \tau_i v\left(\tau_i\,\gamma_m\right)\,\mathbf{x}_i\,\mathbf{x}_i^H\,. \tag{7}$$

Proof. Let us define $\hat{\mathbf{\Sigma}}$ as the unique solution of

$$\hat{\mathbf{\Sigma}} = \frac{1}{N}\sum_{i=0}^{N-1} \tau_i\,u\left(\frac{1}{m}\,\tau_i\,\mathbf{x}_i^H\,\hat{\mathbf{\Sigma}}^{-1}\,\mathbf{x}_i\right)\,\mathbf{x}_i\,\mathbf{x}_i^H\,. \tag{8}$$

As $\mathbf{r_w}_i = \check{\mathbf{C}}^{-1/2}\,\mathbf{C}^{1/2}\,\sqrt{\tau_i}\,\mathbf{x}_i$, it can be easily shown that

$$\check{\mathbf{\Sigma}} = \check{\mathbf{C}}^{-1/2}\,\mathbf{C}^{1/2}\,\hat{\mathbf{\Sigma}}\,\mathbf{C}^{1/2}\,\check{\mathbf{C}}^{-1/2}\,. \tag{9}$$

Equation (6) can be rewritten as

$$\left\|\check{\mathbf{\Sigma}} - \hat{\mathbf{S}}\right\| \le \left\|\check{\mathbf{\Sigma}} - \hat{\mathbf{\Sigma}}\right\| + \left\|\hat{\mathbf{\Sigma}} - \hat{\mathbf{S}}\right\|\,. \tag{10}$$

Concerning the second term of the right hand side of (10), it is proven in [15] that the matrix $\hat{\mathbf{S}}$ given by (7) is such that

$$\left\|\hat{\mathbf{\Sigma}} - \hat{\mathbf{S}}\right\| \to 0 \;\; a.s. \tag{11}$$

With (9), the first term of right hand side of (10) can be rewritten as:

$$\begin{aligned}\left\|\check{\mathbf{\Sigma}} - \hat{\mathbf{\Sigma}}\right\| \le &\left\|\check{\mathbf{C}}^{-1/2}\,\mathbf{C}^{1/2}\,\hat{\mathbf{\Sigma}}\,\mathbf{C}^{1/2}\,\check{\mathbf{C}}^{-1/2} - \hat{\mathbf{\Sigma}}\,\mathbf{C}^{1/2}\,\check{\mathbf{C}}^{-1/2}\right\| \\ &+ \left\|\hat{\mathbf{\Sigma}}\,\mathbf{C}^{1/2}\,\check{\mathbf{C}}^{-1/2} - \hat{\mathbf{\Sigma}}\right\|\,.\end{aligned} \tag{12}$$

After left and right factorizations, we obtain:

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

As $\|\mathbf{C}\|$ has a bounded support, $\|\check{\mathbf{C}}\|$ is bounded too since its eigenvalues support converges almost surely toward the true distribution. Moreover, **Theorem 1** has proved the consistency $\left\|\mathbf{C} - \check{\mathbf{C}}\right\| \to 0 \;\; a.s.$ This ensures the proof. □

In the paper [15], the threshold $t = \dfrac{\phi_\infty\,(1+\sqrt{c})^2}{\gamma_m\,(1 - c\,\phi_\infty)}$ has been set to ensure that all the eigenvalues of the matrix $\hat{\mathbf{S}}$ beyond t correspond to sources. This threshold comes from the upper bound of the support of the Marchenko-Pastur law and details can be found in [15]. Thanks to **Theorem 2**, the threshold t can be applied on eigenvalues of scatter matrix $\check{\mathbf{\Sigma}}$ built on observations $\mathbf{Y}_w = [\mathbf{y}_{w1}, \ldots, \mathbf{y}_{wN}]$ in order to test both hypothesis H_0 and H_1:

$$\check{\mathbf{\Sigma}} = \frac{1}{N}\sum_{i=1}^{N} u\left(\frac{1}{m}\,\mathbf{y}_{wi}^{\;H}\,\check{\mathbf{\Sigma}}^{-1}\,\mathbf{y}_{wi}\right)\,\mathbf{y}_{wi}\,\mathbf{y}_{wi}^{\;H}\,. \tag{13}$$

Let $\left\{\lambda_i(\check{\mathbf{\Sigma}})\right\}_{i=[1,N]}$ be the eigenvalues of $\check{\mathbf{\Sigma}}$ sorted in descending order. As all sources are assumed to be independent, the estimated number $\hat{p}$ of sources that is the rank of the signal subspace is given by $\hat{p} = \min_k(\lambda_k > t)$.

ISBN 978-0-9928626-7-1

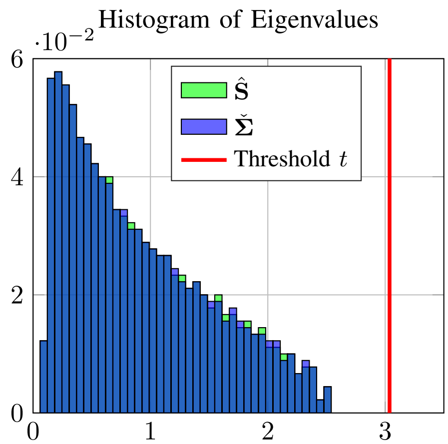

Fig. 2. Eigenvalues of the covariance matrices $\check{\mathbf{\Sigma}}$ and $\hat{\mathbf{S}}$ and the corresponding threshold t ($\rho = 0.7$, $m = 900$, $N = 2000$, $\tau =$ inverse gamma, $\nu = 0.1$).

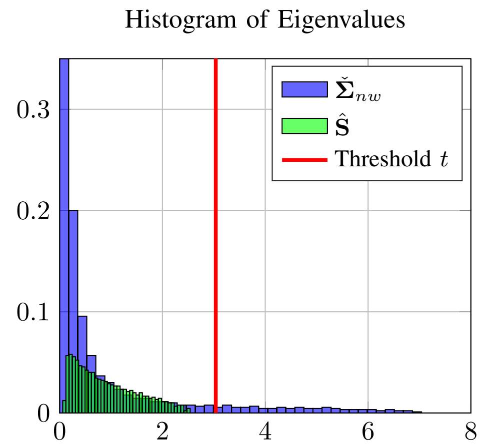

Fig. 3. Eigenvalues of the covariance matrices $\check{\mathbf{\Sigma}}_{nw}$ and $\hat{\mathbf{S}}$ when the signal $\mathbf{Y}$ has not been whitened and the corresponding threshold t ($\rho = 0.7$, $m = 900$, $N = 2000$, $\tau =$ inverse gamma, $\nu = 0.1$).

III. Results and Simulations

This section exposes some results of the proposed method to estimate the rank of the signal subspace, on simulated data.

A. *Behavior of the eigenvalues on simulated CES noise*

In Figure 2 are drawn the distribution of the eigenvalues of $\mathbf{S}$, the threshold t and the distribution of the eigenvalues of $\check{\mathbf{\Sigma}}$ for hypothesis H_0. The chosen signal ($m = 900$, $N = 2000$) is a CES Student-t noise correlated with a Toeplitz matrix with parameter $\rho = 0.7$. In addition, we set the function u with the inverse gamma distribution for τ. For the corresponding Student-t distributed CES noise with parameter ν, we have $u : x \mapsto \dfrac{1+\nu}{\nu + x}$. Hence, the threshold t is equal to $\dfrac{(1+\nu)\,(1+\sqrt{c})^2}{\gamma_m\,(1 - c\,(1+\nu))}$. For the following figures, ν is set to 0.1. In this figure, the eigenvalue distribution is almost the same than the one of $\mathbf{S}$. The fixed-point M-estimator cancels the influence of the texture τ on the $\mathbf{X}$ observations: $\check{\mathbf{\Sigma}}$ is almost equal to $\frac{1}{N}\mathbf{X}\mathbf{X}^H$ as expected. If the noise is not whitened by the proposed estimator $\check{\mathbf{C}}$ and if the scatter matrix is directly estimated with the Fixed Point estimator, noted $\check{\mathbf{\Sigma}}_{nw}$, the threshold is clearly not greater than the largest eigenvalue of $\check{\mathbf{\Sigma}}_{nw}$. This result is shown in Figure 3: this shows that the proposed whitening process is very important when applying this threshold.

B. *Estimation of the number of sources on CES simulated noise and real data*

For simulated and correlated ($\rho = 0.7$) CES noise, the $\{\tau_i\}_{i \in [1,N]}$ are inverse gamma distributed with parameter $\nu = 0.1$. In Figure 4 ($m = 400$ and $N = 2000$), $p = 4$ sources are added in the observations for a Signal to Noise Ratio (SNR) varying from -50 to 50dB. In this figure, the number of sources $\hat{p}$ (mean of 4 trials) is estimated through three methods: AIC, the non-whitened signal and the proposed method. The proposed method starts to find sources from a SNR equal to 10dB. For a greater SNR, whereas it systematically gives the correct number of sources, the other methods overestimate it. In Figure 5 is drawn the same simulation but for $p = 16$ sources. In Figure 5, the proposed estimator still presents better performance than the others.

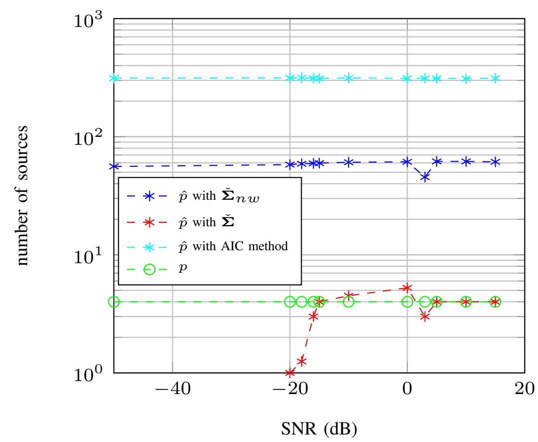

Fig. 4. Estimation of the number p of sources (4 trials) embbeded in CES correlated noise ($m = 400$, $c = 0.2$, $p = 4$ sources, $\rho = 0.7$) versus SNR.

Now we compare the results obtained with three different methods on several real hyperspectral images found in public access: *Indian Pines*, *SalinasA* from AVIRIS database and *PaviaU* from ROSIS database. Let $M1$ be the proposed method, $M2$ be the method consisting in thresholding the eigenvalues of the Fixed-Point estimator without the whitening step, and the usual AIC method. For the function $u(.)$ corresponding to Student-t distribution, we choose $\nu = 0.1$. As we do not have any access to the true distribution of the noise, an empirical estimator of γ is used, $\hat{\gamma} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \frac{1}{m}\mathbf{y}_i^H\,\check{\mathbf{\Sigma}}_{(i)}^{-1}\,\mathbf{y}_i$, where $\check{\mathbf{\Sigma}}_{(i)} = \check{\mathbf{\Sigma}} - \dfrac{1}{N}\,u\left(\dfrac{1}{m}\,\mathbf{y}_i^H\,\check{\mathbf{\Sigma}}^{-1}\,\mathbf{y}_i\right)\,\mathbf{y}_i\,\mathbf{y}_i^H$. Then [15]

ISBN 978-0-9928626-7-1

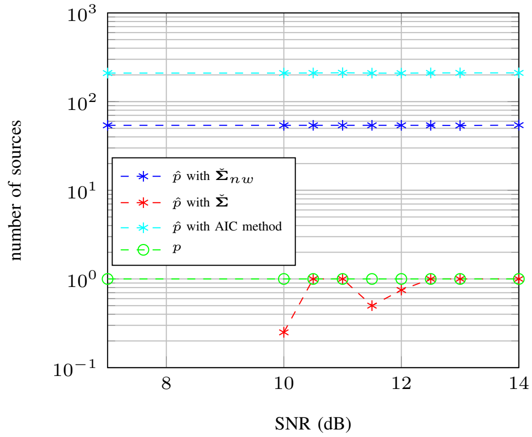

Fig. 5. Estimation of the number $\hat{p}$ of sources (4 trials) embbeded in CES correlated noise ($m = 400$, $c = 0.2$, $p = 1$ source) versus SNR.

TABLE I
ESTIMATED p FOR DIFFERENT HYPERSPECTRAL IMAGES.

Images	Indian Pines	SalinasA	PaviaU	Cars
p	16	9	9	6
$\hat{p}$ M1	11	9	1	3
$\hat{p}$ M2	220	204	103	1
$\hat{p}$ AIC	219	203	102	143

shows that $\gamma - \hat{\gamma} \to 0$ $a.s.$. The results are summarized in table I. On each image, the result tends to be better than those of classical methods.

IV. CONCLUSION AND PERSPECTIVES

In this article devoted to the model order selection of sources embedded in correlated CES noise, we have first proposed a Toeplitz-based covariance matrix estimator of the correlated noise and proved its consistency. To deal with the CES texture, any M-estimator can then be used to estimate the correct structure of the scatter matrix built on whitened observations. A Random Matrix Theory-based model order selection can therefore be applied on the corresponding scatter matrix eigenvalues to correctly separate sources from the noise. We have applied successfully this general technique to simulated correlated CES noise and we also have shown that this method provides interesting and encouraging results on several hyperspectral images containing known sources. This method can be generally applied for any model order selection problems (radar clutter rank estimation, sources localization or any hyperspectral problems such as anomaly detection or linear/non-linear unmixing techniques).

ACKNOWLEDGMENT

The authors would like to thank DGA (Ministry of Defense) for its financial support.

REFERENCES

[1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276–280, March 1986.

[2] V. Marcenko and L. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 507–536, 1967.

[3] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *Signal Processing, IEEE Transactions on*, vol. 42, no. 8, pp. 2146–2157, 1994.

[4] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 10, no. 6, pp. 853–863, Nov 1974.

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

[6] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sept 1978.

[7] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, March 1978.

[8] E. Fishler and H. Messer, "On the use of order statistics for improved detection of signals by the mdl criterion," *IEEE Transactions on Signal Processing*, pp. 2242–2247, 2000.

[9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 33, no. 2, pp. 387–392, April 1985.

[10] E. Terreaux, J. P. Ovarlez, and F. Pascal, "Anomaly detection and estimation in hyperspectral imaging using random matrix theory tools," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015 IEEE 6th International Workshop on*, Dec 2015, pp. 169–172.

[11] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 24–33, Jan 2014.

[12] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *Information Theory, IEEE Transactions on*, vol. 45, no. 2, pp. 641–657, March 1999.

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

[14] Z. Bai and J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices," *The Annals of Probability*, vol. 26, no. 1, pp. 316–345, 1998.

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

[16] R. Couillet, F. Pascal, and J. Silverstein, "Robust estimates of covariance matrices in the large dimensional regime," *Information Theory, IEEE Transactions on*, vol. 60, no. 11, pp. 7269–7278, Nov 2014.

[17] D. Kelker, "Distribution theory of spherical distributions and a location-scale parameter generalization," *The Indian Journal of Statistics*, pp. 419–430, 1970.

[18] D. Manolakis and D. Marden, "Non gaussian models for hyperspectral algorithm design and assessment," *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 3, pp. 1664–1666, 2002.

[19] A. Semeniaka, D. Lekhovitskiy, and D. Rachkov, "Comparative analysis of Toeplitz covariance matrix estimation methods for space-time adaptive signal processing," in *Proceedings of 2011 IEEE CIE International Conference on Radar*, vol. 1, Oct 2011, pp. 696–699.

[20] A. Goian, M. AlHajri, R. Shubair, L. Weruaga, A. Kulaib, R. AlMemari, and M. Darweesh, "Fast detection of coherent signals using pre-conditioned root-MUSIC based on Toeplitz matrix reconstruction," in *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2015, pp. 168–174.

[21] R. Maronna and V. Yohai, *Robust Estimation of Multivariate Location and Scatter*. John Wiley & Sons, Inc., 2004.

[22] J. Vinogradova, R. Couillet, and W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection," *Signal Processing, IEEE Transactions on*, vol. 63, no. 18, pp. 4903–4913, Sept 2015.

[23] 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.

[24] M. Mahot, F. Pascal, P. Forster, and J. P. Ovarlez, "Asymptotic properties of robust covariance matrix estimates," *Signal Processing, IEEE Transactions on*, vol. 61, no. 13, pp. 3348–3356, 2013.

ISBN 978-0-9928626-7-1 © EURASIP 2017