

# New model order selection in large dimension regime for complex elliptically symmetric noise

Eugénie Terreaux, Jean-Philippe Ovarlez, Frédéric Pascal

► **To cite this version:**

Eugénie Terreaux, Jean-Philippe Ovarlez, Frédéric Pascal. New model order selection in large dimension regime for complex elliptically symmetric noise. 25th European Signal Processing Conference (EUSIPCO 2017), Aug 2017, Kos Island, Greece. 10.23919/EUSIPCO.2017.8081376 . hal-01692642

**HAL Id: hal-01692642**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-01692642>**

Submitted on 27 Mar 2020

**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.

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

Eugénie Terreaux\*, Jean-Philippe Ovarlez\*<sup>†</sup>, and Frédéric Pascal<sup>‡</sup>

\*CentraleSupélec-SONDRA, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France  
e-mail: eugenie.terreaux@centralesupelec.fr

<sup>†</sup> ONERA, DEMR/TSI, Chemin de la Hunière, 91120 Palaiseau, France and SONDRA/CentraleSupélec  
e-mail: jean-philippe.ovarlez@onera.fr

<sup>‡</sup>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 modelled by Complex Elliptically Symmetric (CES) distribution, 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 (AIC) and the Minimum Description Length (MDL).

## 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<sup>1</sup> is also degraded if the dimensions  $N$  and  $m$  go to infinity at the same [2]. And this situation arises frequently since sensors are developed to be more accurate and numerous: the improvement 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 signal subspace estimation, the covariance matrix is also used in detection [3], or filtering [4]. Thus, developing more efficient tools and techniques is essential.

All these related techniques are commonly based on information theoretic criteria for model order selection such as

the Akaike Information Criterion (AIC) [5] or the Minimum Description Length (MDL) [6], [7]. Since many techniques are based on the eigenvalues decomposition or the singular values decomposition of the collected data covariance matrix, Wax and Kailath [8] have derived these two criteria based on the eigenvalues which can be conveniently applied 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 [9].

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 [10]. The Random Matrix Theory (RMT), recently developed in signal processing (see [11] for one of the first use of RMT in a signal processing field), provides some useful properties or attributes to handle with these kinds of matrices [12] without using dimensionality reduction. This theory proposes, among others, methods to estimate the eigenvalues distribution for large matrices [13] or for mixed-model with a signal composed of few sources and additive noise [14], [15].

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 context of Complex Elliptical Symmetric (CES) noise [16], which provides a better noise characterisation in many applications (e.g., in hyperspectral imaging [17]). For non-RMT processes, Toeplitz matrices have been widely studied (see for instance [18] or [19]). But herein they are exploited in a large dimension regime. The proposed method is decomposed into two steps. A first part sets the chosen model while the second one presents the beginning of the method, that is to “toeplitzify” the empirical estimation of the covariance matrix (the Sample Covariance Matrix (SCM)) and to prove the consistency of this “toeplitzified” estimator compared to the true covariance matrix. This estimator is required to whiten the data. Then, we propose to use a Maronna’s  $M$ -estimator [20] to estimate the so-called scatter

<sup>1</sup>More precisely, its eigenvalues distribution is not correct anymore

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 Vinogradova [21] to the one of the MDL and AIC methods. Finally, we focus on the model order selection. Indeed a threshold can be applied on the covariance matrix eigenvalues. All eigenvalues greater than this threshold can be proved to be relative to sources due to RMT properties. In conclusion, the relevance of this method on simulated signals and on real hyperspectral images is presented.

*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  $a$ ,  $a^*$  denotes its complex conjugate.  $d_1(\cdot)$  means the distance associated with the  $l_1$ -norm. The distribution  $\delta$  denotes the Dirac measure,  $\text{supp}(\cdot)$  the support of any measure and  $\|\cdot\|$  the spectral matrix norm. The Toeplitz matrix operator is acting on any vector  $\mathbf{x}$  as  $\mathcal{T} : \mathbf{x} \rightarrow \mathcal{T}(\mathbf{x})$  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  $\{\mathbf{y}_i\}_{i \in \{1,\dots,N\}}$  where each  $\mathbf{y}_i$  is a multidimensional  $m$ -vector. In this article, we suppose here the usual random matrix regime, *i.e.*  $N \rightarrow \infty$ ,  $m \rightarrow \infty$  with the constant regime  $c_N = \frac{m}{N} \rightarrow c$ ,  $c > 0$ . The general model characterizing the presence of  $p$  sources corrupted with an additive CES noise can be stated as the following binary hypothesis test:

$$\begin{cases} H_0 : \mathbf{y}_i = \sqrt{\tau_i} \mathbf{C}^{1/2} \mathbf{x}_i, & i \in \{1, \dots, N\}, \\ H_1 : \mathbf{y}_i = \sum_{j=1}^p s_{i,j} \mathbf{m}_j + \sqrt{\tau_i} \mathbf{C}^{1/2} \mathbf{x}_i, & i \in \{1, \dots, N\}, \end{cases} \quad (1)$$

where, for each observation  $i$ ,  $\mathbf{m}_j$  stands for the unknown  $m$ -steering vector of the  $j$ -th deterministic source with power  $s_{i,j}$ ,  $\mathbf{x}_i$  is a multivariate zero-mean white noise of independent entries identically and uniformly distributed on the  $m$ -unit-sphere,  $\tau_1, \dots, \tau_N$  are positive scalar random texture variables (independent from the  $\mathbf{x}_i$  and with a non-specified distribution) and where  $\mathbf{C}$  is a Hermitian Toeplitz covariance matrix defined as  $\mathbf{C} = \mathcal{T}((c_0, \dots, c_{m-1})^T)$ . For large random matrix regime, *i.e.* when  $N \rightarrow \infty$ , we suppose that  $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{\tau_i}$  satisfies  $\int \tau \mu_N(d\tau) \rightarrow 1$  almost surely, that  $\frac{1}{N} \sum \delta_{\lambda_i(\mathbf{C})}$  converges almost surely toward the true measure  $\nu$  and moreover,  $\max_i d_1(\lambda_i(\mathbf{C}), \text{supp}(\nu)) \rightarrow 0$  ( $d_1$  is not specified!!!!) and that  $\{c_k\}_{k \in \{0,\dots,m-1\}}$  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  $\{\tau_i\}_{i \in \{1,\dots,N\}}$  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  $(\mathbf{S}^T)_{i,j} = s_{i,j}$  the  $N \times p$ -matrix representing all the source power, the model can be rewritten as:

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

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 source is contained in the measurements.

Let us define  $\check{c} = (\check{c}_0, \dots, \check{c}_{m-1})^T$  the vector built with elements given by averaging along each diagonal the signal-free SCM defined as  $\mathbf{Y} \mathbf{Y}^H / N = \mathbf{R} \mathbf{R}^H / N$ :

$$\check{c}_k = \frac{1}{mN} \sum_{i=1}^m \sum_{j=1}^N y_{i,j} y_{i+k,j}^* \mathbb{1}_{1 \leq i+k \leq m}, \quad (2)$$

where  $k \in \{0, \dots, m-1\}$  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 assumptions stated above, one has the following result:*

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

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

*Proof.* The proof follows the one in [21] and will not be detailed due to the lack of space. The proof relies on the Lemma 4.1 from [22], 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 decomposed into two parts. Then, we prove that each term converges towards zero. Finally, in the Fourier space, the steps of the proof are similar than the ones in [21], except the additional term  $\|\mathbf{T}\|$  found on the denominator of the upper bound of the inequality.  $\square$

The consistency of the proposed estimator is illustrated in Figure 1. The chosen signal is a white Gaussian noise  $\mathbf{X}$  correlated with a Toeplitz matrix  $\mathbf{C} = \mathcal{T}((\rho^0, \rho^1, \dots, \rho^{m-1})^T)$  where  $\rho = 0.7$ . The texture  $\{\tau_i\}_{i \in [1,N]}$  is randomly extracted from an inverse gamma distribution with unit mean. The resulting noise is therefore Student-t-distributed. This figure presents the  $l_2$ -norm (log scale) of the difference between the real covariance matrix  $\mathbf{C}$  and the proposed estimator (green curve) or the usual SCM estimator equal to  $\frac{1}{N} \mathbf{R} \mathbf{R}^H$  (red curve) when  $N$  varies from 20 to 2000 and  $c = 0.2$  (20 Monte Carlo trials). As expected, the SCM is not consistent while the proposed estimator of Theorem 1 has a slow convergence toward the true covariance.

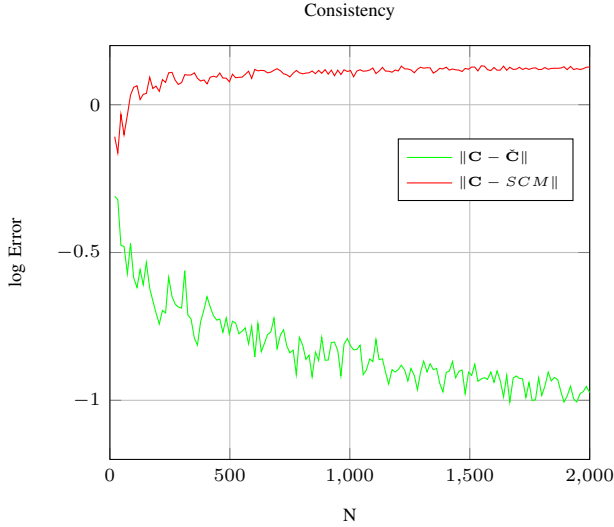


Fig. 1. Consistency of  $\tilde{\mathbf{C}}$  estimator when  $\mathbf{C} = \mathcal{T}((\rho^0, \dots, \rho^{m-1})^T)$  with  $\rho = 0.7$  and  $c = 0.2$  for Student-t distributed  $m$ -vectors.

### C. Signal Subspace Rank Estimation

In this section, the observations  $\mathbf{Y}$  are whitened by the estimator  $\tilde{\mathbf{C}}$  introduced before. Let us note  $\mathbf{Y}_w$  the whitened signal:

$$\mathbf{Y}_w = \tilde{\mathbf{C}}^{-1/2} \mathbf{Y} = \tilde{\mathbf{C}}^{-1/2} \mathbf{M} \mathbf{S} + \mathbf{R}_w, \quad (4)$$

where  $\mathbf{R}_w = \tilde{\mathbf{C}}^{-1/2} \mathbf{C}^{1/2} \mathbf{X} \mathbf{T}^{1/2}$ .

Due to the fact that the observations are polluted by the textures  $\{\tau_i\}_{i \in \{1, \dots, N\}}$ , 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  $\tilde{\Sigma}$  is defined as the unique solution of the following equation:

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

where  $\mathbf{R}_w = [\mathbf{r}_{w_1}, \dots, \mathbf{r}_{w_N}]$  and  $u : [0, +\infty) \rightarrow [0, +\infty)$  nonnegative, continue and non-increasing (see [23] for details). In order to evaluate the rank of the signal subspace, it is possible to set a threshold on the eigenvalues of  $\tilde{\Sigma}$  in a non-RMT regime. But it is not easy to analyze the behavior of  $\tilde{\Sigma}$  eigenvalues when  $N, m \rightarrow \infty$  using RMT classical tools since the term  $u \left( \frac{1}{m} \mathbf{r}_{w_i}^H \tilde{\Sigma}^{-1} \mathbf{r}_{w_i} \right)$  depends on  $\mathbf{r}_{w_i}$ . For that purpose, a possible solution is to find and deal with another useful mathematical object having similar properties. Before establishing the next theorem, here are some definitions:

- $\phi : x \mapsto x u(x)$ , is increasing and bounded with  $\lim_{x \rightarrow \infty} \phi(x) = \phi_\infty > 1$  where  $\phi_\infty < \lim_{N \rightarrow \infty} c_N$   
**PROBLEM!!!! Car c est juste positif!!!! C'est très réducteur cette hypothèse et d'ailleurs ça ne correspond pas aux simulations où  $c = 0.2$ ,**

- $g : x \mapsto \frac{x}{1 - c_N \phi(x)}$ ,  $v : x \mapsto u \circ g^{-1}(x)$ ,
- $\psi : x \mapsto x v(x)$  and  $\gamma_m$  is the unique solution (if it exists) of  $\sum_{i=1}^N \frac{\psi(\tau_i \gamma)}{1 + c_N \psi(\tau_i \gamma)} = 1$ .

**Theorem 2** (Convergence of  $\tilde{\Sigma}$ ). *Using previous definitions, one has the following almost sure convergence*

$$\left\| \tilde{\Sigma} - \hat{\Sigma} \right\| \rightarrow 0 \quad (6)$$

where the matrix  $\hat{\Sigma}$  is defined by:

$$\hat{\Sigma} \triangleq \frac{1}{N} \sum_{i=1}^N \tau_i v(\tau_i \gamma_m) \mathbf{x}_i \mathbf{x}_i^H. \quad (7)$$

*Proof.* Define  $\hat{\Sigma}$  as the unique solution of

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

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

$$\tilde{\Sigma} = \tilde{\mathbf{C}}^{-1/2} \mathbf{C}^{1/2} \hat{\Sigma} \mathbf{C}^{1/2} \tilde{\mathbf{C}}^{-1/2}, \quad (9)$$

and

$$\left\| \tilde{\Sigma} - \tilde{\mathbf{C}}^{-1/2} \mathbf{C}^{1/2} \hat{\Sigma} \mathbf{C}^{1/2} \tilde{\mathbf{C}}^{-1/2} \right\| \leq \left\| \hat{\Sigma} - \hat{\mathbf{S}} \right\| \|\tilde{\mathbf{C}}\| \|\mathbf{C}\|. \quad (10)$$

It is proven in [14] that the matrix  $\hat{\Sigma}$  given by (7) is such that

$$\left\| \hat{\Sigma} - \hat{\mathbf{S}} \right\| \rightarrow 0 \text{ a.s.} \quad (11)$$

As  $\|\mathbf{C}\|$  has a bounded support,  $\|\tilde{\mathbf{C}}\|$  is bounded too since its eigenvalues support converges almost surely toward the true distribution. Moreover, it can also be easily shown that the consistency implies that  $\tilde{\mathbf{C}}^{-1/2} \mathbf{C}^{1/2} \rightarrow \mathbf{I}$ ,  $\mathbf{I}$  being the identity matrix, term to term. The result given in (11) ensures the proof.  $\square$

In [14], the threshold  $t = \frac{\phi_\infty (1 + \sqrt{c})^2}{\gamma_m (1 - c \phi_\infty)}$  has been set to ensure that all the eigenvalues of the matrix  $\hat{\Sigma}$  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 [14]. Thanks to Theorem 2, the threshold  $t$  can also be applied on the eigenvalues of scatter matrix  $\tilde{\Sigma}$  built on  $\mathbf{Y}_w = [\mathbf{y}_{w_1}, \dots, \mathbf{y}_{w_N}]$  in order to test both hypothesis  $H_0$  and  $H_1$ :

$$\tilde{\Sigma} = \frac{1}{N} \sum_{i=1}^N u \left( \frac{1}{m} \mathbf{y}_{w_i}^H \tilde{\Sigma}^{-1} \mathbf{y}_{w_i} \right) \mathbf{y}_{w_i} \mathbf{y}_{w_i}^H. \quad (12)$$

Let  $\{\lambda_i(\tilde{\Sigma})\}_{i \in \{1, \dots, N\}}$  be the eigenvalues of  $\tilde{\Sigma}$  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  $\hat{p} = \min_k (\lambda_k > t)$ .

### III. RESULTS AND SIMULATIONS

In this section, the proposed method is used to estimate the rank of the signal subspace, on simulated data.

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

Figure 2 represents the Marchenko-Pastur law, the threshold  $t$  and the distribution of the eigenvalues of  $\check{\Sigma}$  for hypothesis  $H_0$ . The chosen signal ( $m = 900$ ,  $N = 2000$ ) is a correlated Student-t noise with a Toeplitz matrix ( $\rho = 0.7$ ). In addition, we set the function  $u$  with the inverse gamma distribution for  $\tau$ . As a reminder,  $u$  is computed in order to obtain the Maximum Likelihood for the Fixed-Point estimator in the non-RMT regime even if the RMT regime does not ensure to maintain the performances:  $u(\cdot) = -h'(\cdot)/h(\cdot)$  where  $h(\cdot)$  is the CES generator function and where  $'$  denotes the derivative. J'ENLEVERAI TOUTE LA PHRASE QUI N'EST PAS CLAIRE et n'apporte pas grand chose For the corresponding Student-t distributed noise with parameter  $\nu$ , one has  $u : x \mapsto \frac{1 + \nu}{\nu + x}$ . Hence, the threshold  $t$  is equal to  $\frac{(1 + \nu)(1 + \sqrt{c})^2}{\gamma_m(1 - c(1 + \nu))}$ . For the following figures,  $\nu$  is set to 0.1 (peut-être ici une footnote disant que en régime classique  $\nu$  doit être  $> 2$  mais que en RMT régime  $\nu$  doit juste être positif... car c'est  $\nu/m$  en fait...). In Figure 2, the eigenvalues distribution is almost the same than the Marchenko-Pastur law. The fixed-point  $M$ -estimator cancels the influence of the texture  $\tau$  on the  $\mathbf{X}$  observations:  $\check{\Sigma}$  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, denoted  $\check{\Sigma}_{nw}$ , the threshold is clearly not greater than the largest eigenvalue of  $\check{\Sigma}_{nw}$ . This result is illustrated in Figure 3: this shows that the proposed whitening process is very important when applying the threshold technique.

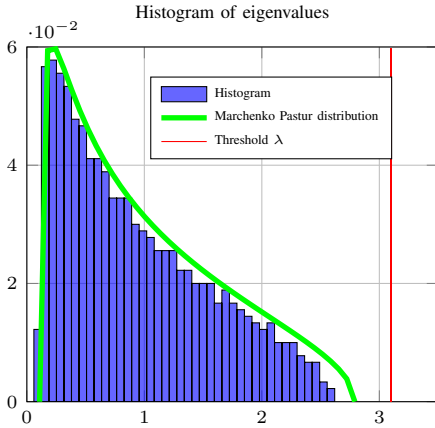


Fig. 2. Eigenvalues distribution of  $\check{\Sigma}$  for  $m = 900$ ,  $N = 2000$ .

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

For simulated and correlated ( $\rho = 0.7$ ) CES noise, the  $\{\tau_i\}_{i \in \{1, \dots, N\}}$  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 with a Signal-to-Noise Ratio (SNR) varying from  $-50$  to  $50$ dB. In this figure, the

Histogram of Eigenvalues

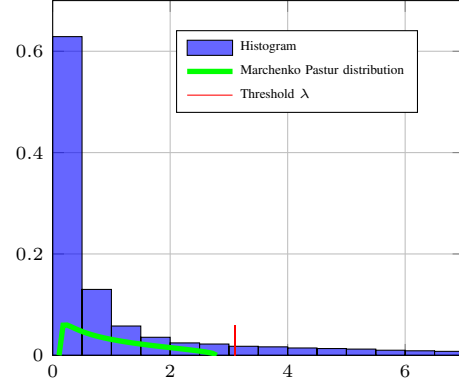


Fig. 3. Eigenvalues distribution of  $\check{\Sigma}_{nw}$  for  $m = 900$ ,  $N = 2000$ .

number of sources  $\hat{p}$  (mean of 4 trials) is estimated thanks to three different 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. Figure 5 presents the same simulation for  $p = 16$  sources. Indeed, similar remarks occurs: the proposed estimator still presents better performance than the others.

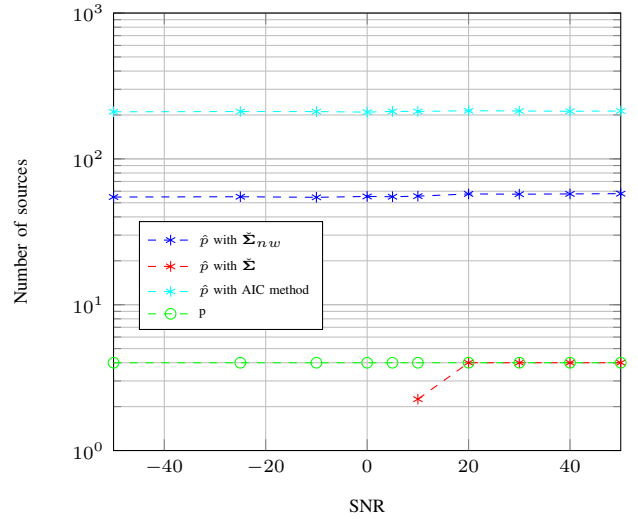


Fig. 4. Estimation of the number  $p$  of sources embedded in CES correlated noise ( $m = 400$ ,  $c = 0.2$ ,  $p = 4$  sources,  $\rho = 0.7$ ) versus SNR.

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  $u(\cdot)$  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

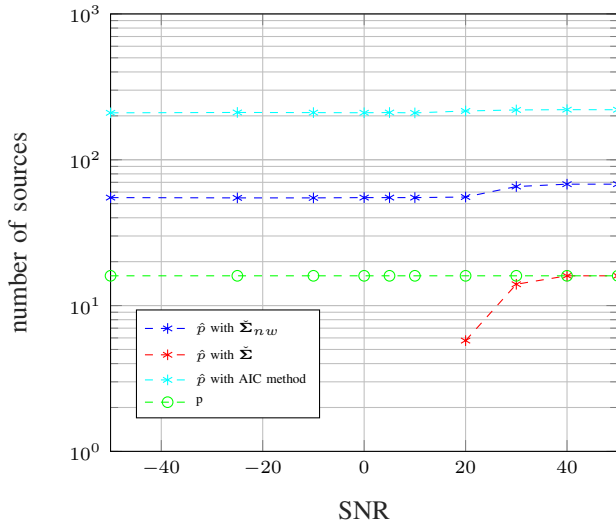


Fig. 5. Estimation of the number  $p$  of sources embedded in CES correlated noise ( $m = 400$ ,  $c = 0.2$ ,  $p = 16$  sources,  $\rho = 0.7$ ) versus SNR.

TABLE I  
ESTIMATED  $p$  FOR DIFFERENT HYPERSPECTRAL IMAGES.

Images	Indian Pines	SalinasA	PaviaU	Cars
$\hat{p}$	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

empirical estimator of  $\gamma$  is used,  $\hat{\gamma} = \frac{1}{N} \sum_{i=1}^N \frac{1}{m} \mathbf{y}_i^H \tilde{\Sigma}_{(i)}^{-1} \mathbf{y}_i$ ,

where  $\tilde{\Sigma}_{(i)} = \tilde{\Sigma} - \frac{1}{N} u \left( \frac{1}{m} \mathbf{y}_i^H \tilde{\Sigma}^{-1} \mathbf{y}_i \right) \mathbf{y}_i \mathbf{y}_i^H$ . Then [14] shows that  $\gamma - \hat{\gamma} \rightarrow 0$  a.s.. The results are summarised in table I. On each image, the result tends to be better than those of classical methods.

#### IV. CONCLUSION AND PERSPECTIVES

This article is devoted to the model order selection of sources embedded in correlated CES noise. First, a Toeplitz-based covariance matrix estimator for the correlated noise has been proposed and has been proven to be consistent. 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 has been therefore used on the corresponding scatter matrix eigenvalues to correctly separate sources from the noise. We have applied successfully this general technique on both simulated correlated CES noise and real hyperspectral data. This method provides interesting and encouraging results on several hyperspectral images containing known sources. Note that the proposed method is general and can be 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] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.
- [13] 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.
- [14] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *Journal of Multivariate Analysis*, vol. 140, pp. 139–161, 2015.
- [15] 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.
- [16] D. Kelker, "Distribution theory of spherical distributions and a location-scale parameter generalization," *The Indian Journal of Statistics*, pp. 419–430, 1970.
- [17] 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.
- [18] A. Semeniaka, D. Likhovitskiy, 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.
- [19] 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.
- [20] R. Maronna and V. Yohai, *Robust Estimation of Multivariate Location and Scatter*. John Wiley & Sons, Inc., 2004.
- [21] 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.

- [22] 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.
- [23] 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.