

Research Article

Model Order Selection for Short Data: An Exponential Fitting Test (EFT)

Angela Quinlan,¹ Jean-Pierre Barbot,² Pascal Larzabal,² and Martin Haardt³

¹ Department of Electronic and Electrical Engineering, University of Dublin, Trinity College, Ireland

² SATIE Laboratory, École Normale Supérieure de Cachan, 61 avenue du Président Wilson, 94235 Cachan Cedex, France

³ Communications Research Laboratory, Ilmenau University of Technology, P.O. Box 100565, 98684 Ilmenau, Germany

Received 29 September 2005; Revised 31 May 2006; Accepted 4 June 2006

Recommended by Benoit Champagne

High-resolution methods for estimating signal processing parameters such as bearing angles in array processing or frequencies in spectral analysis may be hampered by the model order if poorly selected. As classical model order selection methods fail when the number of snapshots available is small, this paper proposes a method for noncoherent sources, which continues to work under such conditions, while maintaining low computational complexity. For white Gaussian noise and short data we show that the profile of the ordered noise eigenvalues is seen to approximately fit an exponential law. This fact is used to provide a recursive algorithm which detects a mismatch between the observed eigenvalue profile and the theoretical noise-only eigenvalue profile, as such a mismatch indicates the presence of a source. Moreover this proposed method allows the probability of false alarm to be controlled and predefined, which is a crucial point for systems such as RADARs. Results of simulations are provided in order to show the capabilities of the algorithm.

Copyright © 2007 Angela Quinlan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

In sensor array processing, it is important to determine the number of signals received by an antenna array from a finite set of observations or snapshots. A similar problem arises in line spectrum estimations. The number of sources has to be determined successfully in order to obtain good performance for high-resolution direction finding estimates. A lot of work has been published concerning the model order selection problem. Estimating the number of sources is traditionally thought of as being equivalent to the determination of the number of eigenvalues of the covariance matrix which are different from the smallest eigenvalue [1]. Such an approach leads to a rank reduction principle in order to separate the noise from the signal eigenvalues [2]. Anderson [3] gave a hypothesis testing procedure based on the confidence interval of the noise eigenvalue, in which a threshold value must be assigned subjectively. He showed [3] that the log-likelihood ratio to the number of snapshots is asymptotic to a χ^2 distribution. For a small number of snapshots, James introduced the idea of “modified statistics” [4]. In [5], Chen et al. proposed a

method based on an a priori on the observation probability density function that detects the number of sources present by setting an upper bound on the value of the eigenvalues.

For thirty years information theoretic criteria (ITC) approaches have been widely suggested for detection of multiple sources [6]. The best known of this test family are the Akaike information criterion (AIC) [7] and the minimum description length (MDL) [8–10]. Such criteria are composed of two terms. The first depends on the data and the second is a penalty term concerning the number of free parameters (parsimony). The AIC is not consistent and tends to over-estimate the number of sources present, even at high signal-to-noise ratio (SNR) values. While the MDL method is consistent, it tends to under-estimate the number of sources at low and moderate SNR. In [11] a theoretical evaluation is given of the probability of over— and under—estimation of source detection methods such as the AIC and MDL, under the assumption of asymptotical conditions.

In an effort to moderate the behavior of the AIC and MDL methods Wong et al. proposed a modified ITC

approach in [12], which uses the marginal p.d.f. of the sample eigenvalues as the log-likelihood function. In [1] a general ITC is proposed in which the first term of the criteria can be selected from a set of suitable functions. Based on this method Wu and Fuhrmann [13] then proposed a parametric technique as an alternative method of defining the first term of this criteria.

Using Bayesian methodology, Djurić then proposed an alternative to the AIC and MDL methods [14, 15] in which the penalty against over-parameterization was no longer independent of the data. Some authors have also investigated the possible use of eigenvectors for model order selection [16, 17], but they generally suffer from the necessity to introduce a priori knowledge. More recently, Wu et al. [18] proposed two ways of estimating the number of sources by drawing Gerschgorin radii.

These algorithms work correctly when the noise eigenvalues are closely clustered. However for a small sample size, where we define a sample as small when the number of snapshots is of the same order as the number of sensors, this condition is no longer valid and the noise eigenvalues can instead be seen to have an approximately exponential profile.

Recently this problem of detecting multiple sources was readdressed by means of looking directly for a gap between the noise and the signal eigenvalues [19]. In this way, and as an alternative to the traditional approaches, we recently proposed a method [20] to obtain an estimation of the number of significant targets in time reversal imaging. Motivated by experimental results reported in [21], this method exploits the exponential profile of the ordered noise eigenvalues first introduced in [22]. Assuming that the smallest eigenvalue is a noise eigenvalue, this exponential profile can then be used to find the theoretical profile of the noise-only eigenvalues. Starting with the smallest eigenvalue a recursive algorithm is then applied in order to detect a mismatch greater than a threshold value between each observed eigenvalue and the corresponding theoretical eigenvalue. The occurrence of such a mismatch indicates the presence of a source, and the eigenvalue index where this mismatch first occurs is equal to the number of sources present.

The test initially proposed in [20] uses thresholds obtained from the empirical dispersion of ordered noise eigenvalues. The proposed paper presents an alternative to determine the corresponding thresholds for a predefined false alarm probability, and through simulations we show the improvements in comparison with some of the traditional tests.

Section 2 presents the basic formulation of the problem. In Section 3, we recall the model for the eigenvalue profile and explain how the parameters of this model are calculated. Section 4 describes the detection test deduced from this model and how the corresponding thresholds are calculated in order to control the false alarm. Section 5 compares the performance of this test with that of the usual tests. Section 6 draws our conclusions concerning the method.

2. PROBLEM FORMULATION

2.1. Antenna signal model

We consider an array of M sensors located in the wavefield generated by d narrow-band point sources. Let $\mathbf{a}(\theta)$ be the steering vector representing the complex gains from one source at location θ to the M sensors. Then, if $\mathbf{x}(t)$ is the observation vector of size $M \times 1$, $\mathbf{s}(t)$ the emitted vector signal of size $d \times 1$, and $\mathbf{n}(t)$ the additive noise vector of size $M \times 1$, we obtain the following conventional model:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \mathbf{y}(t) + \mathbf{n}(t), \quad (1)$$

where \mathbf{A} is the matrix of the d steering vectors. Moreover, the vector $\mathbf{n}(t)$ denotes spatially and temporally uncorrelated circular Gaussian complex noise with distribution $N(0, \sigma^2 \mathbf{I})$ which is also uncorrelated with the signals. Thus, from (1), the observation covariance matrix \mathbf{R}_x can be expressed as

$$\mathbf{R}_x = E[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{R}_y + \mathbf{R}_n = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma^2\mathbf{I}. \quad (2)$$

2.2. Principle of statistical tests based on eigenvalue profile

According to (1), the noiseless observations $\mathbf{y}(t)$ are a linear combination of $\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_d)$. Assuming independent source amplitudes $\mathbf{s}(t)$, the random vector $\mathbf{y}(t)$ spans the whole subspace generated by the steering vectors. This is the “signal subspace.” Assuming $d < M$ and no antenna ambiguity, the signal subspace dimension is d , and consequently the number of nonzero eigenvalues of \mathbf{R}_y is equal to d , with $(M - d)$ eigenvalues being zero.

Now, in the presence of white noise, according to (2), \mathbf{R}_x has the same eigenvectors as \mathbf{R}_y , with eigenvalues $\lambda_x = \lambda_y + \sigma^2$ and the smallest $(M - d)$ eigenvalues equal to σ^2 . Then, from the spectrum of \mathbf{R}_x with eigenvalues in decreasing order, it becomes easy to discriminate between signal and noise eigenvalues and order determination would be an easy task.

In practice, \mathbf{R}_x is unknown and an estimate is made using $\hat{\mathbf{R}}_x = (1/N) \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}(t)^H$, where N is the number of snapshots available. As $\hat{\mathbf{R}}_x$ involves averaging over the number of snapshots available $\hat{\mathbf{R}}_x \rightarrow \mathbf{R}_x$, as $N \rightarrow \infty$, resulting in all the noise eigenvalues being equal to σ^2 . However, when taken over a finite number of snapshots, the sample matrix $\hat{\mathbf{R}}_x \neq \mathbf{R}_x$. In the spectrum of ordered eigenvalues, the “signal eigenvalues” are still identified as the d largest ones. But, the noise eigenvalues are no longer equal to each other, and the separation between the signal and noise eigenvalues is not clear (except in the case of high SNR, when a gap can be observed between signal and noise eigenvalues), making discrimination between signal and noise eigenvalues a difficult task.

2.3. Qualification of order estimation performance

Letting \hat{d} equal the estimated number of sources, three exclusive situations and their corresponding probabilities will

be considered:

$$\begin{aligned} \hat{d} = d &: \text{correct detection, } P_d = \text{Prob}[\hat{d} = d], \\ \hat{d} > d &: \text{false alarm, } P_{fa} = \text{Prob}[\hat{d} > d], \\ \hat{d} < d &: \text{nondetection, } 1 - P_d - P_{fa} = \text{Prob}[\hat{d} < d]. \end{aligned} \quad (3)$$

Various methods will be compared on the basis of P_d and P_{fa} values for various numbers of sources, locations, and power conditions.

Usually, a detection threshold may be adjusted to provide the best compromise between detection and false alarm. In such situations, a common practice is to set the threshold for a given value of P_{fa} (1% for instance) and to compare the corresponding values of P_d for different methods. The probabilities P_d and P_{fa} will be estimated from statistical occurrence rates by Monte Carlo simulations.

2.4. Classical tests

Several tests have been proposed for determining the number of sources in the presence of statistical fluctuations. The most common of these tests, recalled below, are the Akaike information criterion (AIC) [7], and Rissanen's minimum description length (MDL) criterion [8]. More recently, a new version of the MDL, named (MDLB), has been proposed in [10] and an information theoretic criterion, the predictive description length (PDL) has been proposed in [23], able to resolve coherent and noncoherent sources. They are based on a decomposition of the correlation matrix \mathbf{R}_x into two orthogonal components; the signal and noise subspaces. As the MDLB and PDL require a maximum likelihood (ML) estimation of the angle of arrival, their computational cost is significantly greater than for the AIC and MDL tests, but they lead to more precise model order selection.

The AIC, MDL, MDLB, and PDL tests will be used as benchmarks in this paper.

The aim of the AIC method is to determine the order of a model using information theory. Using the expression given in [9] for the AIC, the number of sources is the integer \hat{d} which, for $m \in \{0, 1, \dots, M-1\}$, minimizes the following quantity:

$$\text{AIC}(m) = -N(M-m) \log \left(\frac{g(m)}{a(m)} \right) + m(2M-m), \quad (4)$$

where $g(m)$ and $a(m)$ are, respectively, the geometric and arithmetic means of the $(M-m)$ smallest eigenvalues of the covariance matrix of the observation. The first term stands for the log-likelihood residual error, while the second is a penalty for over-fitting. This criterion does not determine the true number of sources with a probability of one, even with an infinite number of samples.

The MDL approach is also based on information theoretic arguments, and the selected model order is the one which minimizes the code length needed to describe the data.

In this paper we use the form of the MDL given in [9]:

$$\text{MDL}(m) = -N(M-m) \log \left(\frac{g(m)}{a(m)} \right) + \frac{1}{2}m(2M-m) \log N. \quad (5)$$

It appears that the MDL method is similar to AIC method except for the penalty term, leading to an asymptotic consistent test.

Concerning now the MDLB and PDL tests, ML estimates are used to find the projection of the sample correlation matrix $\hat{\mathbf{R}}_x$ onto the signal and noise subspaces. The summation of the ML estimates of these matrices is the ML estimate of the correlation matrix. The number of sources detected by the PDL and MDLB tests are, respectively, obtained by the minimization of the cost functions:

$$\begin{aligned} \hat{d}_{\text{PDL}} &= \arg \min_m \text{PDL}_m(N), \\ \hat{d}_{\text{MDLB}} &= \arg \min_m \text{MDLB}_m(N), \end{aligned} \quad (6)$$

where $m \in \{0, 1, \dots, M-1\}$, $\text{PDL}_m(N)$ and $\text{MDLB}_m(N)$ are the PDL criterion and MDLB criterion computed with N snapshots and a number of m candidate sources. Expressions of $\text{PDL}_m(N)$ and $\text{MDLB}_m(N)$ are obtained as follows.

If the estimate of $\hat{\mathbf{R}}_x$ is computed with i snapshots, $\hat{\mathbf{R}}_x(i)$, then

$$\hat{\mathbf{R}}_x(i) = \frac{1}{i} \sum_{t=1}^i \mathbf{x}(t)\mathbf{x}^H(t). \quad (7)$$

In the sequel, the sample estimates will be represented by a "hat" ($\hat{\cdot}$) placed on the top of the character and the ML estimates by a "bar" ($\bar{\cdot}$).

The estimated matrix $\hat{\mathbf{R}}_x(i-1)$ can be projected onto signal and noise subspaces. The projected correlation matrices for the m th model are given by

$$\begin{aligned} \hat{\mathbf{R}}_{xs}^m(i-1) &= \mathbf{P}_s(\boldsymbol{\theta}_m) \hat{\mathbf{R}}_x(i-1) \mathbf{P}_s(\boldsymbol{\theta}_m), \\ \hat{\mathbf{R}}_{xn}^m(i-1) &= \mathbf{P}_n(\boldsymbol{\theta}_m) \hat{\mathbf{R}}_x(i-1) \mathbf{P}_n(\boldsymbol{\theta}_m), \end{aligned} \quad (8)$$

where $\mathbf{P}_s(\boldsymbol{\theta}_m)$ and $\mathbf{P}_n(\boldsymbol{\theta}_m)$ are, respectively, the projector on the signal subspace and the projector on the noise subspace. The projectors $\mathbf{P}_s(\boldsymbol{\theta}_m)$ and $\mathbf{P}_n(\boldsymbol{\theta}_m)$ are defined by

$$\begin{aligned} \mathbf{P}_s(\boldsymbol{\theta}_m) &= \mathbf{A}(\boldsymbol{\theta}_m) (\mathbf{A}^H(\boldsymbol{\theta}_m) \mathbf{A}(\boldsymbol{\theta}_m))^{-1} \mathbf{A}^H(\boldsymbol{\theta}_m), \\ \mathbf{P}_n(\boldsymbol{\theta}_m) &= \mathbf{I} - \mathbf{P}_s(\boldsymbol{\theta}_m), \end{aligned} \quad (9)$$

where $\mathbf{A}(\boldsymbol{\theta}_m)$ is the matrix of the m steering vectors $\mathbf{a}(\boldsymbol{\theta}_j)$, $j \in \{1, 2, \dots, m\}$ and $\boldsymbol{\theta}_m$ is the direction of arrival vector.

The ML estimate of the correlation matrix for the m th model (a model with m sources) and obtained with $(i-1)$ snapshots is

$$\bar{\mathbf{R}}_x^m(i-1) = \bar{\mathbf{R}}_{xs}^m(i-1) + \bar{\mathbf{R}}_{xn}^m(i-1). \quad (10)$$

If $\hat{\boldsymbol{\theta}}_m$ is the ML estimate vector of the m directions of arrival ($\bar{\boldsymbol{\theta}}_m = \hat{\boldsymbol{\theta}}_m$), then

$$\bar{\mathbf{R}}_{xs}^m(i-1) = \hat{\mathbf{R}}_{xs}^m(i-1). \quad (11)$$

In a similar way, it is possible to show that $\bar{\mathbf{R}}_{xn}^m(i-1)$ has the same eigenvectors as $\hat{\mathbf{R}}_{xn}^m(i-1)$ and a single eigenvalue of multiplicity $(M-m)$ obtained by

$$\bar{\sigma}(\bar{\theta}_{i-1}^m) = \frac{1}{M-m} \text{tr}(\hat{\mathbf{R}}_{xn}^m(i-1)), \quad (12)$$

where $\text{tr}(\cdot)$ represents the trace of a matrix. The matrix $\bar{\mathbf{R}}_{xn}^m(i-1)$ is thus obtained while applying the linear transformation,

$$\bar{\mathbf{R}}_{xn}^m(i-1) = \mathbf{T}_{i-1}^m \hat{\mathbf{R}}_{xn}^m(i-1) \quad (13)$$

with $\lambda_j(\hat{\mathbf{R}}_{xn}^m(i-1))$, $j = 1, \dots, M-m$ the nonzero eigenvalues of $\hat{\mathbf{R}}_{xn}^m(i-1)$, $\hat{\mathbf{V}}_{n,M-m}$ the $M \times (M-m)$ matrix of the corresponding eigenvectors, $\text{diag}[\cdot]$ the diagonal matrix formed by the elements in the brackets, and

$$\mathbf{T}_{i-1}^m = \hat{\mathbf{V}}_{n,M-m} \text{diag} \left[\frac{\bar{\sigma}(\bar{\theta}_{i-1}^m)}{\lambda_j(\hat{\mathbf{R}}_{xn}^m(i-1))} \right] \hat{\mathbf{V}}_{n,M-m}^H \quad (14)$$

The PDL test for N snapshots and m candidate sources is then obtained with [23]

$$\begin{aligned} \text{PDL}_m(N) &= \sum_{i=M+1}^N \left[\log \zeta(\hat{\mathbf{R}}_{xs}^m(i-1)) \right. \\ &\quad + (M-m) \times \log \left(\frac{1}{M-m} \text{tr}(\hat{\mathbf{R}}_{xn}^m(i-1)) \right) \\ &\quad \left. + \mathbf{x}^H(i) (\hat{\mathbf{R}}_{xs}^m(i-1) + \mathbf{T}_{i-1}^m \hat{\mathbf{R}}_{xn}^m(i-1))^{-1} \mathbf{x}(i) \right] \end{aligned} \quad (15)$$

and the MDLB expression is given by [10, 23]

$$\begin{aligned} \text{MDLB}_m(N) &= N \log \zeta(\hat{\mathbf{R}}_{xn}^m(N)) \\ &\quad + N(M-m) \log \left(\frac{1}{M-m} \text{tr}(\hat{\mathbf{R}}_{xs}^m(i-1)) \right) \\ &\quad + \frac{m(m+1)}{2} \log(N), \end{aligned} \quad (16)$$

where $\zeta(\cdot)$ represents the multiplication of the nonzero eigenvalues. Note that in expression (15), the PDL test is computed for all $i = M+1, M+2, \dots, N$.

In [23], the estimate $\hat{\mathbf{R}}_x(i)$ of the true correlation matrix $\mathbf{R}_x(i)$ is obtained by the recursion $\hat{\mathbf{R}}_x(i) = \alpha \hat{\mathbf{R}}_x(i-1) + (\alpha-1) \mathbf{x}(i) \mathbf{x}^H(i)$ where $\alpha < 1$ is a real smoothing factor and the factor $1/(\alpha-1)$ is the effective length of the exponential window [24]. In this paper, $\hat{\mathbf{R}}_x(i)$ is estimated with expression (7).

The computation of the PDL and MDLB depends on the ML estimation of the angle of arrival vector $\bar{\theta}_{i-1}^m$. As suggested in [10, 23], the *alternate projection algorithm* is used to reduce the complexity [25].

These two methods (PDL, MDLB) can detect both coherent and noncoherent signals. The PDL can also be used online and then applied to time varying systems and target tracking. In this paper, as the EFT is applicable to fixed and noncoherent sources detection, only this case will be investigated.

3. EIGENVALUE PROFILE OF THE CORRELATION MATRIX UNDER THE NOISE-ONLY ASSUMPTION

As the noise eigenvalues are no longer equal for a small sample size it is necessary to identify the mean profile of the decreasing noise eigenvalues. We therefore consider the eigenvalue profile of the sample covariance matrix for the noise-only situation $\hat{\mathbf{R}}_n = (1/N) \sum_{t=1}^N \mathbf{n}(t) \cdot \mathbf{n}(t)^H$. The distribution of the matrix $\hat{\mathbf{R}}_n$ is a Wishart distribution [26] with N degrees of freedom. This distribution can be seen as a multivariate generalization of the χ^2 distribution. It depends on N , M , and σ^2 and is sometimes denoted by $W_M(N, \sigma^2 \mathbf{I})$. In order to establish the mean profile of the ordered eigenvalues (denoted as $\lambda_1, \dots, \lambda_M$) the joint probability of an ordered M -tuple has to be known. The joint distribution of the ordered eigenvalues is then [26]

$$\begin{aligned} p(\lambda_1, \dots, \lambda_M) &= \alpha \left(-\frac{1}{2\sigma^2} \sum_{i=1}^M \lambda_i \right) \left(\prod_{i=1}^M \lambda_i \right)^{(1/2)(N-M-1)} \prod_{i>j} (\lambda_j - \lambda_i), \end{aligned} \quad (17)$$

where α is a normalization coefficient. The distribution of each eigenvalue can be found in [27], but this requires zonal polynomials and, to our knowledge, produces unusable results.

Instead we use an alternative approach which consists of finding an approximation of this profile by conserving the first two moments of the trace of the error covariance matrix defined by $\Psi = \hat{\mathbf{R}}_n - \mathbf{R}_n = \hat{\mathbf{R}}_n - E\{\hat{\mathbf{R}}_n\} = \hat{\mathbf{R}}_n - \sigma^2 \mathbf{I}$. It follows from $E\{\text{tr}[\Psi]\} = \mathbf{0}$ that, in a first approximation,

$$M\sigma^2 = \sum_{i=1}^M \lambda_i. \quad (18)$$

Using the definition of the error covariance matrix Ψ , the element Ψ_{ij} can be expressed as

$$\Psi_{ij} = \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) - \sigma^2 \delta_{ij}. \quad (19)$$

Consequently, $E[\|\Psi_{ij}\|^2]$ is obtained as follows:

$$\begin{aligned} E[\|\Psi_{ij}\|^2] &= E \left[\left\| \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) - \sigma^2 \delta_{ij} \right\|^2 \right] \\ &= E \left[\left\| \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) \right\|^2 \right] + E[\|\sigma^2 \delta_{ij}\|^2] \\ &\quad + E \left[-2\Re \left\{ \sigma^2 \delta_{ij} \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) \right\} \right], \end{aligned} \quad (20)$$

where $\Re\{\cdot\}$ represents the real part of a complex value.

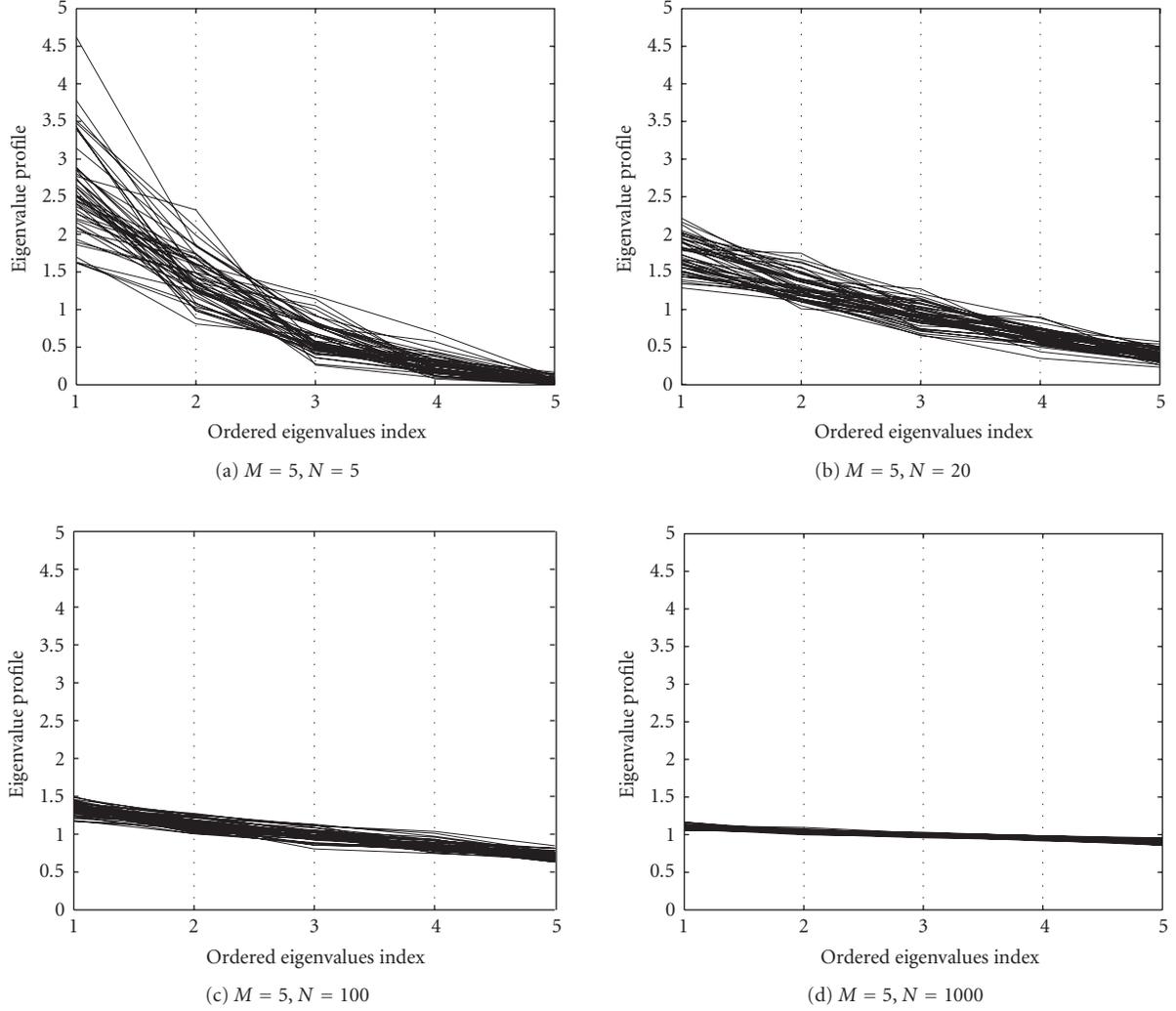


FIGURE 1: Profile of the ordered eigenvalues under the noise-only assumption for 50 independent trials, with $M = 5$ and various values of N .

Let us now derive each term of (20):

$$\begin{aligned}
 E \left[\left\| \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) - \sigma^2 \delta_{ij} \right\|^2 \right] &= \frac{1}{N^2} N \sigma^4 = \frac{\sigma^4}{N}, \\
 E \left[\|\sigma^2 \delta_{ij}\|^2 \right] &= \sigma^4 \delta_{ij}, \\
 E \left[-2\Re \left\{ \sigma^2 \delta_{ij} \frac{1}{N} \sum_{t=1}^N n_i(t) \cdot n_j^*(t) \right\} \right] & \\
 &= -\frac{2\sigma^2 \delta_{ij}}{N} E \left[\Re \left\{ \sum_{t=1}^N n_i(t) \cdot n_j^*(t) \right\} \right] \\
 &= -\frac{2\sigma^2 \delta_{ij}}{N} \left(\frac{N\sigma^2}{2} \right) = -\sigma^4 \delta_{ij}.
 \end{aligned} \tag{21}$$

Finally,

$$E \left[\|\Psi_{ij}\|^2 \right] = \frac{\sigma^4}{N} + \sigma^4 \delta_{ij} - \sigma^4 \delta_{ij} = \frac{\sigma^4}{N}. \tag{22}$$

Since the trace of a matrix remains unchanged when the base changes, it follows that

$$\sum_{i,j} E \left\{ \|\Psi_{ij}\|^2 \right\} = E \left\{ \text{tr} (\hat{\mathbf{R}}_n - \mathbf{R}_n)^2 \right\} = M^2 \frac{\sigma^4}{N} \tag{23}$$

and, in a first approximation,

$$M^2 \frac{\sigma^4}{N} = \sum_{i=1}^M (\lambda_i - \sigma^2)^2. \tag{24}$$

From both simulation results shown in Figure 1, and experimental results reported in literature (e.g., see [21]) the decreasing model of the noise-only eigenvalues can be seen to be approximately exponential. The decreasing model retained for the approximation is

$$\lambda_i = \lambda_1 r_{M,N}^{i-1}, \tag{25}$$

with $0 < r_{M,N} < 1$. Of course, $r_{M,N}$ depends on M and N , but is denoted by r for simplicity. From (18) we get

$$\lambda_1 = M \frac{1-r}{1-r^M} \sigma^2 = MJ_M \sigma^2, \quad (26)$$

where

$$J_M = \frac{1-r}{1-r^M}. \quad (27)$$

Considering that $(\lambda_i - \sigma^2) = (MJ_M r^{i-1} - 1)\sigma^2$, the relation (23) gives

$$\frac{M+N}{MN} = \frac{(1-r)(1+r^M)}{(1-r^M)(1+r)}. \quad (28)$$

We therefore set $r = e^{-2a}$ ($a > 0$), leading to the re-expression of (28) as

$$\frac{M \cdot \tanh(a) - \tanh(Ma)}{M \cdot \tanh(Ma)} = \frac{1}{N}, \quad (29)$$

where $\tanh(\cdot)$ is the hyperbolic tangent function. An order-4 expansion gives the following biquadratic equation in a :

$$a^4 - \frac{15}{M^2+2}a^2 + \frac{45M}{N(M^2+1)(M^2+2)} = 0 \quad (30)$$

for which the positive solution is given by

$$\begin{aligned} a(M, N) &= \sqrt{\frac{1}{2} \left\{ \frac{15}{M^2+2} - \sqrt{\frac{225}{(M^2+2)^2} - \frac{180M}{N(M^2+1)(M^2+2)}} \right\}}. \end{aligned} \quad (31)$$

As the calculation of the noise-only eigenvalue profile takes into account the number of snapshots, this profile is valid for all sample sizes, with the exponential profile tending to a horizontal profile as the noise eigenvalues become equal.

4. A RECURSIVE EXPONENTIAL FITTING TEST (EFT)

4.1. Test principle

The expressions for the noise-only eigenvalue profile can now be extended to the case where the observations consist of d noncoherent sources corrupted by additive noise. Under these conditions the covariance matrix can be broken down into two complementary subspaces: the source subspace \mathcal{E}_s of dimension d , and the noise subspace \mathcal{E}_n of dimension $Q = M - d$. Consequently, the profile established in the previous section still holds for the Q noise eigenvalues, and the theoretical noise eigenvalues can be found by replacing M with Q in the previous expressions for the noise-only eigenvalue profile.

The proposed test then finds the highest dimension P of the candidate noise subspace, such that the profile of these P

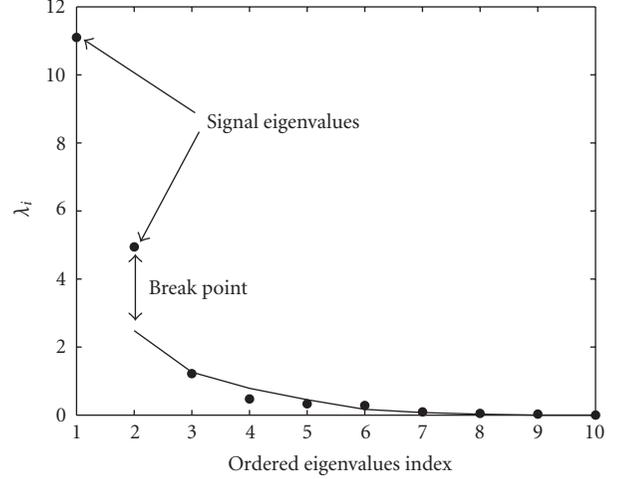


FIGURE 2: Profile of ordered noise eigenvalues in the presence of 2 sources, and 10 sensors. The ordered profile of the observed eigenvalue is seen to break from the noise eigenvalue distribution, when there are sources present.

candidate noise eigenvalues is compatible with the theoretical noise eigenvalue profile. The main idea of the test is to detect the eigenvalue index at which a break occurs between the profile of the observed eigenvalues and the theoretical noise eigenvalue profile provided by the exponential model. Figure 2 shows how a break point appears between the signal eigenvalues and the theoretical noise eigenvalue profile, while the observed noise eigenvalues are seen to fit the theoretical profile.

Firstly, an eigen-decomposition of the sample covariance matrix is performed and the resulting eigenvalues $\lambda_1, \dots, \lambda_M$, which we call the observed eigenvalues, are arranged in order of decreasing size. Beginning with the smallest observed eigenvalue λ_M , this is assumed to be a noise eigenvalue, giving the initial candidate noise subspace dimension $P = 1$. Then using λ_M , $P = 1$, and the prediction equation (32) we find the next eigenvalue of the theoretical noise eigenvalue profile $\hat{\lambda}_{M-1}$:

$$\begin{aligned} \hat{\lambda}_{M-P} &= (P+1)J_{P+1}\hat{\sigma}^2, \quad \text{with } J_{P+1} = \frac{1-r_{P+1,N}}{1-(r_{P+1,N})^{P+1}}, \\ \hat{\sigma}^2 &= \frac{1}{P+1} \sum_{i=0}^P \lambda_{M-i}. \end{aligned} \quad (32)$$

Now taking both λ_M and $\hat{\lambda}_{M-1}$ to be noise eigenvalues, corresponding to a candidate noise subspace dimension $P = 2$, (32) is applied again to predict $\hat{\lambda}_{M-2}$.

These steps are then repeated, and for each step the candidate noise subspace dimension P is increased by one. Then taking all the previously estimated noise eigenvalues, the next noise eigenvalue in the theoretical profile $\hat{\lambda}_{M-P}$ is found. This process is continued until $P = M - 1$, and we now have the M eigenvalues of the theoretical noise-only profile, $\hat{\lambda}_1, \dots, \hat{\lambda}_M$, where $(\hat{\lambda}_M = \lambda_M)$.

We define the following two hypotheses:

$$\begin{aligned} H_{P+1} : \lambda_{M-P} \text{ is a noise eigenvalue,} \\ \bar{H}_{P+1} : \lambda_{M-P} \text{ is a signal eigenvalue.} \end{aligned} \quad (33)$$

Then, starting with the smallest eigenvalue pair (that are not equal) $\hat{\lambda}_{M-1}$ and λ_{M-1} , the relative distance between each of the theoretical noise eigenvalues and the corresponding observed eigenvalue is found, and compared to the threshold found for that eigenvalue index, (34) and (35),

$$H_{P+1} : \left| \frac{\lambda_{M-P} - \hat{\lambda}_{M-P}}{\hat{\lambda}_{M-P}} \right| \leq \eta_P, \quad (34)$$

$$\bar{H}_{P+1} : \left| \frac{\lambda_{M-P} - \hat{\lambda}_{M-P}}{\hat{\lambda}_{M-P}} \right| > \eta_P. \quad (35)$$

If the relative difference between the theoretical noise eigenvalue and the observed eigenvalue is less than (or equal to) the corresponding threshold, the observed eigenvalue matches the theoretical noise-only eigenvalue profile, and so it is deemed to be a noise eigenvalue, which is the case shown by (34).

We then compare the next eigenvalues $\hat{\lambda}_{M-2}$ and λ_{M-2} in the same manner. This process continues until we find a pair of eigenvalues, $\hat{\lambda}_{M-P}$ and λ_{M-P} whose relative difference is greater than the corresponding threshold, as shown in (35). When this happens the observed eigenvalue is taken to correspond to a signal eigenvalue and so the test stops here. The estimated dimension of the noise subspace \hat{P} is the value P where the test stops, that is, when the hypothesis given in (35) is chosen over that in (34). The estimated model order is then given by $\hat{d} = M - \hat{P}$.

Note on the complexity

The proposed EFT method requires calculation of the sample correlation matrix for each set of observations. An eigenvalue decomposition of this matrix must then be performed and the smallest of the observed eigenvalues is used to predict the theoretical noise-only eigenvalue profile. The computational cost of the EFT method is of the same order as those of the AIC and MDL tests. Compared to the methods proposed in [9, 23] the computational complexity of the proposed algorithm is much lower due to the fact that both these algorithms rely on initially finding a maximum likelihood estimate of the direction of arrival for each proposed number of sources. This estimation step greatly increases the computational complexity and necessitates the introduction of computational cost reduction techniques. Moreover, the PDL proposed in [23] requires the calculation of the sample covariance matrix and its eigen-decomposition at each individual snapshot.

4.2. Computation of thresholds

The comparison thresholds are closely related to the statistical distribution of the prediction error and are determined

to respect a preset probability of false alarm P_{fa} . The P_{fa} is the probability of the method mistakenly determining that a source is present, and is defined as

$$P_{fa} = \Pr[\hat{d} > d_0 \mid d = d_0] \quad \text{for } d_0 = 0, 1, 2, \dots, M-1. \quad (36)$$

For the noise-only case $d = 0$, and the expression for P_{fa} can be decomposed as follows:

$$P_{fa} = \Pr[\hat{d} > 0 \mid d = 0] = \sum_{i=1}^{M-1} \Pr[\hat{d} = i \mid d = 0] = \sum_{p=1}^{M-1} P_{fa}^{(p)}, \quad (37)$$

where $P_{fa}^{(p)} = \Pr[\hat{d} = M - P \mid d = 0]$ is the contribution of P th step to the total false alarm.

Reexpressing (34) and (35) we get

$$H_{P+1} : Q(P) = \left| \frac{\lambda_{M-P}}{\sum_{i=M-P}^M \lambda_i} \right| \leq (\eta_P + 1)J_{P+1}, \quad (38)$$

$$\bar{H}_{P+1} : Q(P) = \left| \frac{\lambda_{M-P}}{\sum_{i=M-P}^M \lambda_i} \right| > (\eta_P + 1)J_{P+1},$$

resulting in the following expression for $P_{fa}^{(p)}$ in the noise-only situation:

$$P_{fa}^{(p)} = \Pr[Q(P) > (\eta_P + 1)J_{P+1} \mid d = 0]. \quad (39)$$

Then, denoting the distribution of $Q(P)$ as $h_p(q)$ the threshold η_P is defined by the following integral equation:

$$P_{fa}^{(M-P)} = \int_{J_{P+1}(\eta_P + 1)}^{+\infty} h_p(q) dq. \quad (40)$$

Solution of this equation in order to find η_P is reliant on knowledge of the distribution $h_p(q)$. For $P = M$ and $P = M - 1$ the distribution is known as given in [8], but is unusable in our application. To our knowledge, this statistical distribution is not known for other values of P . Hence, numerical methods must instead be used in order to solve for η_P .

4.3. Threshold determination by Monte Carlo methods

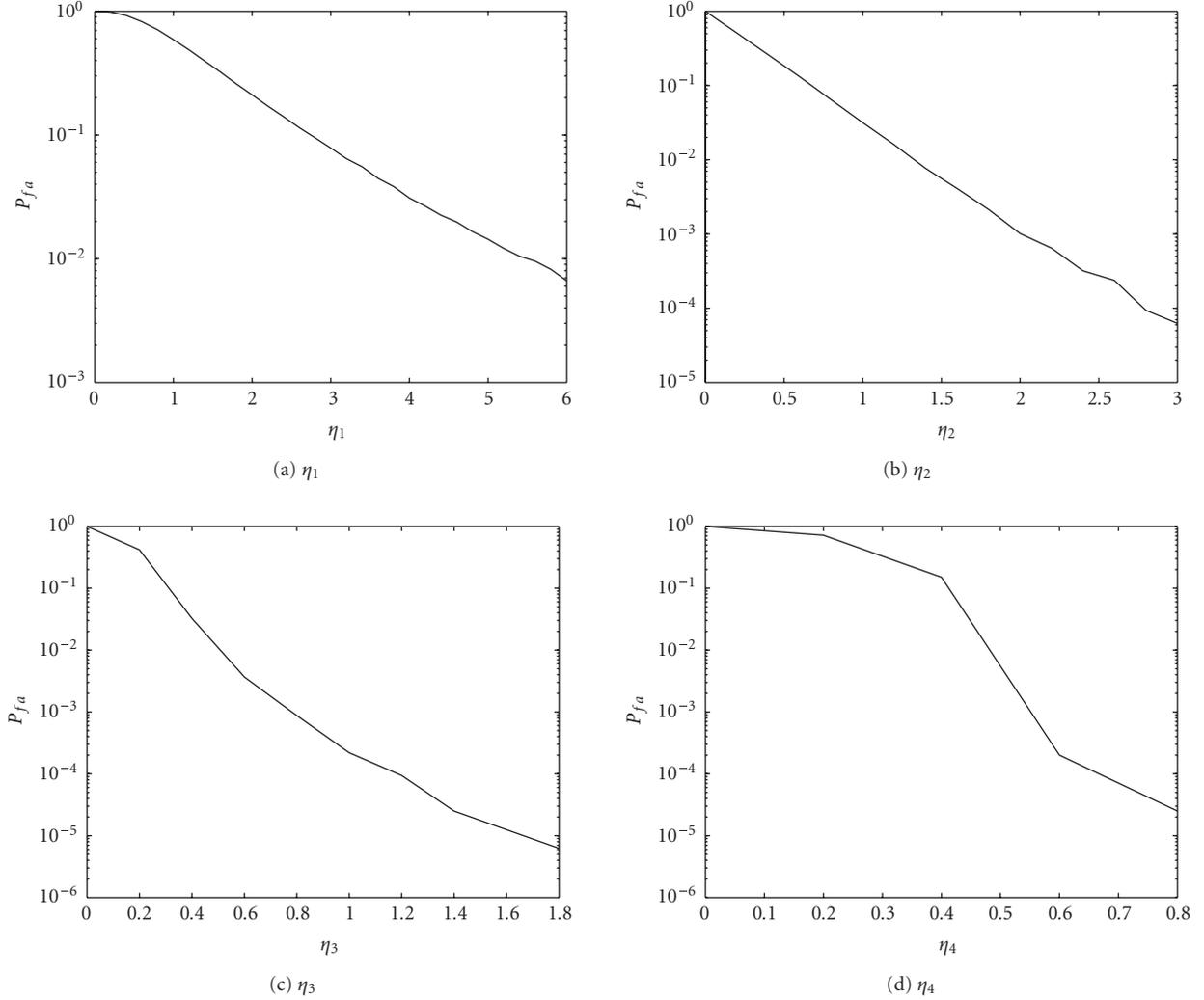
Using $I = P_{fa}^{(M-P)}$ for the sake of notational simplicity, we rewrite equation (40) as

$$I = \int_{\mathcal{D}} p(\lambda_1, \dots, \lambda_M) \prod_{i=1}^M d\lambda_i = E[1_{\mathcal{D}}], \quad (41)$$

where \mathcal{D} is the domain of integration defined as follows:

$$\mathcal{D} = \{0 < \lambda_M < \dots < \lambda_1 < \infty \mid Q(P) > J_{P+1}(\eta_P + 1)\}, \quad (42)$$

and $1_{\mathcal{D}}(\lambda_1, \dots, \lambda_M)$ is the indicator function over the domain \mathcal{D} . The value of the indicator function is unity if the eigenvalues belong to \mathcal{D} and zero otherwise. Equation (41) can then be estimated by Monte Carlo simulations, in which the steps are

FIGURE 3: Thresholds computation for $M = 5$ and $N = 10$.

- (i) generation of q noise-only sample correlation matrices, where q is the number of the Monte Carlo trials to be run;
- (ii) computation of the ordered eigenvalues for each of these q matrices: $(\lambda_{1,j}, \dots, \lambda_{M,j})$ $1 \leq j \leq q$;
- (iii) estimation of I by $\hat{I} = (1/q) \sum_{j=1}^q \mathbf{1}_{\mathcal{D}}(\lambda_{1,j}, \dots, \lambda_{M,j})$.

As the P_{fa} is usually very small, q must be statistically determined in order to obtain a predefined precision for the estimation of I . Because of the central limit theorem, \hat{I} follows a Gaussian law. Consequently, denoting the standard deviation of \hat{I} as σ , we can say $\Pr[(\sqrt{q}/\sigma)|I - \hat{I}| < 1.96] = 0.95$, where $\Pr[x < y]$ is the probability that $x < y$. Then, as $\sigma^2 = E[(\mathbf{1}_{\mathcal{D}}(\cdot))^2] - I^2 = I - I^2 \approx I$, we obtain $\sigma = \sqrt{I}$.

Application

For $M = 5$ sensors and a false alarm probability of 1%, identically distributed over the $M - 1$ steps of the test, $I = P_{fa}^{(M-P)} = 0.01/4 = 0.0025$ and $1 \leq P \leq 4$. With a probability

of 95%, I is estimated with an accuracy of 10% if $q = 160000$. In Figure 3 we have plotted the $P_{fa}^{(M-P)}$ versus η_p . From this, η_p is selected for each P and for a given P_{fa} .

5. PERFORMANCE AND COMPARISON WITH CLASSICAL TESTS

In order to evaluate the test performance in white Gaussian complex noise, computed simulations have been performed with a uniform linear array of five omnidirectional sensors. The distance between adjacent sensors is half a wavelength. The number of snapshots is $N = 6$. All the simulations have been performed with 1000 Monte Carlo simulations. Two sources of the same power impinge on the array at -10° and $+10^\circ$. The SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10} \left(\frac{\sigma_s^2}{\sigma^2} \right), \quad (43)$$

where σ_s^2 is the power of one of the sources and σ^2 is the noise power.

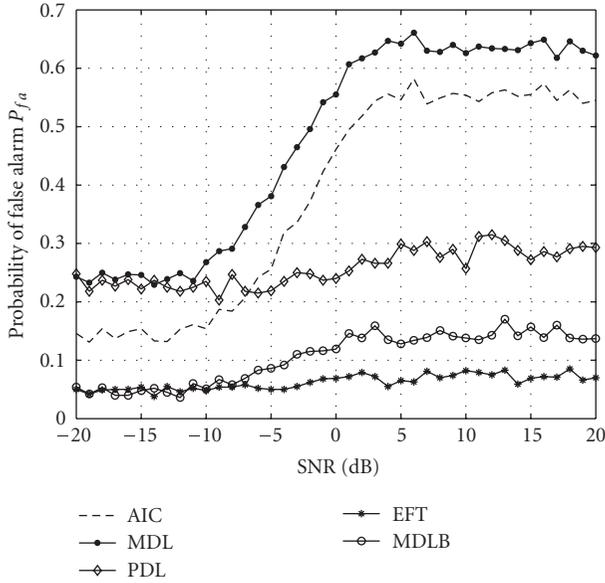


FIGURE 4: Comparison of the probability of false alarm for the EFT (predefined $P_{fa} = 10\%$), the MDL, the AIC, the PDL, and the MDLB.

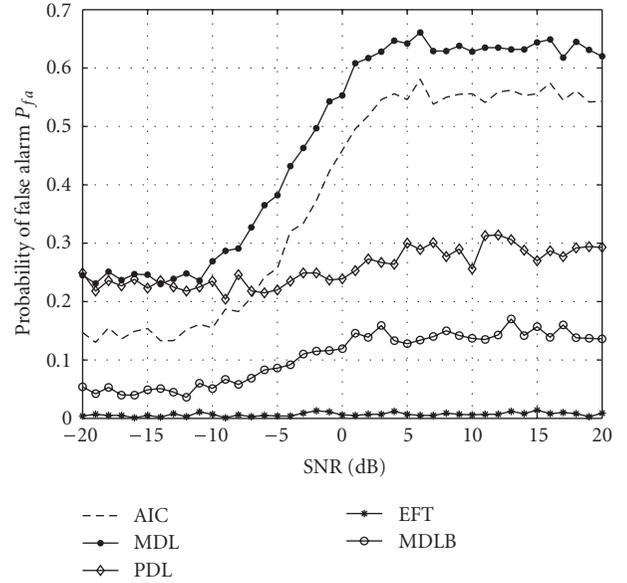


FIGURE 6: Comparison of the probability of false alarm for the EFT (predefined $P_{fa} = 1\%$), the MDL, the AIC, the PDL, and the MDLB.

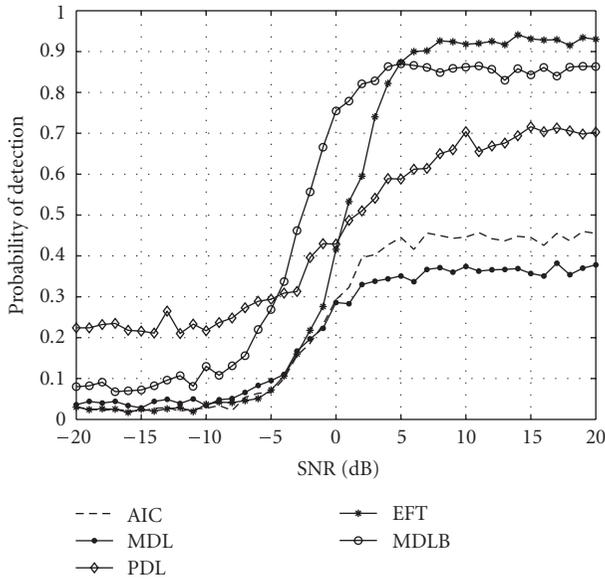


FIGURE 5: Probability of detection for the EFT (predefined $P_{fa} = 10\%$), the MDL, the AIC, the PDL, and the MDLB.

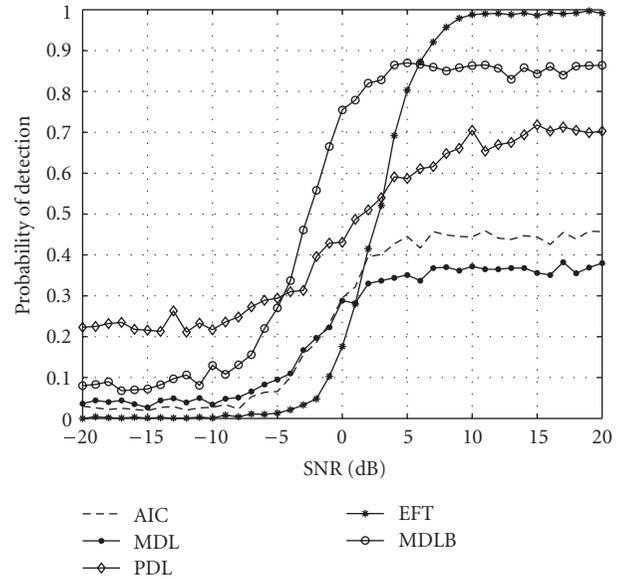


FIGURE 7: Probability of detection for the EFT (predefined $P_{fa} = 1\%$), the MDL, the AIC, the PDL, and the MDLB.

For various SNR, all the criteria, AIC, MDL, EFT, PDL, MDLB are applied. The EFT test has firstly been designed for a $P_{fa} = 10\%$. In such a configuration, the thresholds of the EFT test are $\eta_1 = 26.3990$, $\eta_2 = 3.6367$, $\eta_3 = 1.2383$, and $\eta_4 = 0.6336$. In Figure 4 we have reported the probability of false alarm versus SNR for AIC, MDL, EFT, PDL, and MDLB. As expected the P_{fa} of EFT is 10% and we observe that the uncontrolled P_{fa} of other tests is significantly higher, except for the MDLB which is about 10% when the SNR is lower than -4 dB. In Figure 5 we have reported the probability of correct detection versus SNR for the same tests. We observe that only the EFT and MDLB tests give good results

both in terms of probability of correct detection and probability of false alarm. When the SNR is lower than 5 dB, the MDLB gives the best probability of detection and acceptable results for the probability of false alarm, but requires an important computational complexity. When the SNR is greater than 5 dB, the EFT outperforms all the other tests in terms of P_d with a P_{fa} still lower than 10%.

Now, if the $P_{fa} = 1\%$, the thresholds of the EFT test are $\eta_1 = 88.5464$, $\eta_2 = 6.5121$, $\eta_3 = 2.1086$, and $\eta_4 = 1.1050$. We observe in Figure 6 that the P_{fa} of the EFT is always well controlled. In Figure 7 we observe that even with such a disadvantageous constraint for EFT, this last gives better results

than the classical tests in terms of correct probability of detection P_d for SNR higher than 7 dB.

We can note that the P_d of classical tests has drastically decreased when the noise eigenvalues are not closely clustered.

6. CONCLUSION

We have proposed a new test for model order selection based on the geometrical profile of noise-only eigenvalues. We have shown that noise eigenvalues for white Gaussian noise fit an exponential law whose parameters have been predicted. Contrary to traditional algorithms, this test performs well when there is a small number of snapshots used for the estimation of the correlation matrix. Another important advantage over classical tests is that the false alarm probability can be adjusted by a predetermined threshold. Moreover, the computational cost of the EFT method is of the same order as those of the AIC and MDL.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their helpful suggestions that considerably improved the quality of the paper. This work has been partly funded by the European Network of Excellence NEWCOM.

REFERENCES

- [1] Y. Q. Yin and P. R. Krishnaiah, "On some nonparametric methods for detection of the number of signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 11, pp. 1533–1538, 1987.
- [2] L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 3, pp. 350–355, 1987.
- [3] T. W. Anderson, "Asymptotic theory for principal component analysis," *Annals of Mathematical Statistics*, vol. 34, pp. 122–148, 1963.
- [4] A. T. James, "Test of equality of latent roots of the covariance matrix," *Journal of Multivariate Analysis*, pp. 205–218, 1969.
- [5] W. Chen, K. M. Wong, and J. Reilly, "Detection of the number of signals: a predicted eigen-threshold approach," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1088–1098, 1991.
- [6] P. Stoica and Y. Selén, "Model-order selection: a review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.
- [7] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [8] J. Rissanen, "Modeling by shortest data description length," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.
- [10] M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 8, pp. 1190–1196, 1989.
- [11] M. Kaveh, H. Wang, and H. Hung, "On the theoretical performance of a class of estimators of the number of narrow-band sources," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 9, pp. 1350–1352, 1987.
- [12] K. M. Wong, Q.-T. Zhang, J. Reilly, and P. Yip, "On information theoretic criteria for determining the number of signals in high resolution array processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 11, pp. 1959–1971, 1990.
- [13] Q. Wu and D. Fuhrmann, "A parametric method for determining the number of signals in narrow-band direction finding," *IEEE Transactions on Signal Processing*, vol. 39, no. 8, pp. 1848–1857, 1991.
- [14] P. M. Djurić, "Model selection based on asymptotic Bayes theory," in *Proceedings of the 7th IEEE SP Workshop on Statistical Signal and Array Processing*, pp. 7–10, Quebec City, Quebec, Canada, June 1994.
- [15] W. B. Bishop and P. M. Djurić, "Model order selection of damped sinusoids in noise by predictive densities," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 611–619, 1996.
- [16] H. L. Van Trees, *Optimum Array Processing*, vol. 4 of *Detection, Estimation and Modulation Theory*, John Wiley & Sons, New York, NY, USA, 2002.
- [17] A. Di, "Multiple source location - a matrix decomposition approach," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 5, pp. 1086–1091, 1985.
- [18] H.-T. Wu, J.-F. Yang, and F.-K. Chen, "Source number estimators using transformed Gerschgorin radii," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1325–1333, 1995.
- [19] A. P. Liavas and P. A. Regalia, "On the behavior of information theoretic criteria for model order selection," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1689–1695, 2001.
- [20] A. Quinlan, J.-P. Barbot, and P. Larzabal, "Automatic determination of the number of targets present when using the time reversal operator," *The Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 2220–2225, 2006.
- [21] M. Tanter, J.-L. Thomas, and M. Fink, "Time reversal and the inverse filter," *The Journal of the Acoustical Society of America*, vol. 108, no. 1, pp. 223–234, 2000.
- [22] J. Grouffaud, P. Larzabal, and H. Clergeot, "Some properties of ordered eigenvalues of a Wishart matrix: application in detection test and model order selection," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '96)*, vol. 5, pp. 2463–2466, Atlanta, Ga, USA, May 1996.
- [23] S. Valaee and P. Kabal, "An information theoretic approach to source enumeration in array signal processing," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1171–1178, 2004.
- [24] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on first-order perturbations," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2758–2770, 1994.
- [25] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.

- [26] N. L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*, chapter 38-39, John Wiley & Sons, New York, NY, USA, 1972.
- [27] P. R. Krishnaiah and F. J. Schurmann, "On the evaluation of some distribution that arise in simultaneous tests of the equality of the latents roots of the covariance matrix," *Journal of Multivariate Analysis*, vol. 4, pp. 265–282, 1974.

Angela Quinlan studied engineering in Trinity College Dublin, Ireland, and received her B.S. degree in electronic engineering in June 2002. In October 2002, she began her Ph.D. in the Department of Electronics in Trinity College. From 2004 to 2005, she completed a year working with the Signal Processing and Information Team at the SATIE Laboratory at the École Normale Supérieure de Cachan. She will defend her Ph.D. thesis in June 2006. Her research interests include statistical signal and array processing, audio source localization, audio time reversal focusing, and estimator performance bounds.



Jean-Pierre Barbot was born in Tours, France, in 1964. He received the Agrégation degree in electrical engineering from the École Normale Supérieure de Cachan, France, in 1988. In 1995, he received the Ph.D. degree from the University Paris-Sud, France. From 1990 to 1994, he worked in the CNET (France Telecom), Paris, France, where he was involved in indoor propagation studies for mobile communications.



From 1994 to 1996, he was a teacher in the Electrical Engineering Department of the Institut Universitaire de Technologie of Velizy, University of Versailles Saint-Quentin, France, and worked in the CETP Laboratory, UMR-8639 CNRS, University of Versailles Saint-Quentin, France, on propagation studies. In 1996, he joined the Electrical Engineering Department of the École Normale Supérieure de Cachan, France, as an Associate Professor. He teaches electronics, microwaves, signal processing, and numerical communications. He is a Member of the Signal Processing Team of SATIE Laboratory, UMR CNRS, École Normale Supérieure de Cachan, France. His current research interest is in the field of estimation applied to numerical communications and radar. He is a Member of the European Network of Excellence NEWCOM.

Pascal Larzabal was born in the Basque country in the south of France in 1962. He received the Agrégation degree in electrical engineering and the Ph.D. degree from École Normale Supérieure de Cachan, France, in 1988 and 1992, respectively, and the Habilitation à Diriger les Recherches degree from École Normale Supérieure de Cachan, France, in 1998. He is now a Professor at the Institut Universitaire de Technologie de Cachan, University of Paris-Sud, France, where he is the Head of the Electrical Engineering Department. He teaches electronics, signal processing and control. He is the Head of the Signal Processing Team of SATIE Laboratory, UMR CNRS, École Normale Supérieure de Cachan, where his research concerns estimation in array processing and spectral analysis for communications and radar. He is a Member of the European Network of Excellence NEWCOM.



Martin Haardt has been a Full Professor and Head of the Communications Research Laboratory at Ilmenau University of Technology, Germany, since 2001. After studying electrical engineering at the Ruhr-University Bochum, Germany, and at Purdue University, USA, he received his Diplom-Ingenieur (M.S.) degree from the Ruhr-University Bochum in 1991 and his Doktor-Ingenieur (Ph.D.) degree from Munich University of Technology in 1996. In 1997, he joined Siemens Mobile Networks in Munich, Germany, where he was responsible for strategic research for third-generation mobile radio systems. From 1998 to 2001, he was the Director for International Projects and University Cooperations in the mobile infrastructure business of Siemens, where his work focused on mobile communications beyond 3G. During his time at Siemens, he also taught in an international master program at Munich University of Technology. He received the Vodafone Award for outstanding research in mobile communications, the ITG Best Paper Award from the VDE, and the Rohde and Schwarz Outstanding Dissertation Award. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, and numerical linear algebra. He is a Senior Member of the IEEE. Since 2002, he has served as an Associate Editor for the IEEE Transactions on Signal Processing.

