AN EMPIRICAL EIGENVALUE-THRESHOLD TEST FOR SPARSITY LEVEL ESTIMATION FROM COMPRESSED MEASUREMENTS

A. Lavrenko, F. Römer, G. Del Galdo, R. Thomä

Ilmenau University of Technology
Institute for Information Technology
Helmholzplatz 2, 98693, Ilmenau, Germany
anastasia.lavrenko@tu-ilmenau.de

O. Arikan

Bilkent University
Electrical and Electronics Eng. Dep.
TR-06800 Bilkent, Ankara, Turkey
oarikan@ee.bilkent.edu.tr

ABSTRACT

Compressed sensing allows for a significant reduction of the number of measurements when the signal of interest is of a sparse nature. Most computationally efficient algorithms for signal recovery rely on some knowledge of the sparsity level, i.e., the number of non-zero elements. However, the sparsity level is often not known a priori and can even vary with time. In this contribution we show that it is possible to estimate the sparsity level directly in the compressed domain, provided that multiple independent observations are available. In fact, one can use classical model order selection algorithms for this purpose. Nevertheless, due to the influence of the measurement process they may not perform satisfactorily in the compressed sensing setup. To overcome this drawback, we propose an approach which exploits the empirical distributions of the noise eigenvalues. We demonstrate its superior performance compared to state-of-the-art model order estimation algorithms numerically.

Index Terms— Compressed sensing, sparsity level, detection, model order selection

1. INTRODUCTION

Compressed sensing (CS) is a recently emerged paradigm that provides a framework to simultaneously compress sparse signals while measuring them. Most of the theoretical bounds derived within CS are expressed in terms of the dimensionality of the problem, including the *sparsity level* of the signal, i.e., the number of non-zero elements in a proper representation. Moreover, the vast majority of efficient reconstruction methods, like greedy algorithms for example, rely on a priori knowledge of the signal sparsity as well. However, in practical applications such information is rarely available beforehand. One way to tackle this problem is to use cross-validation as in [1]. Unfortunately, this requires performing multiple signal reconstructions, at a significant cost in terms of computational complexity. Therefore, a method to estimate the sparsity level efficiently directly from the measurements would be highly desirable.

Some initial steps to show that classical signal processing problems such as detection, classification and estimation can be performed directly in the compressed domain were made in [2]. In [3] a compressive subspace detector is proposed, where the sparsity level is known a priori. A close relation between sparse signal reconstruction and parameter estimation with model order selection has been discussed in [4], where the sparsity-promoting regularization parameter (which influences the model order of the sparse solution) is chosen according to classical information criteria. However, the specific task of detecting the sparsity level from the compressed measurements, to the best of our knowledge, has not been analyzed yet.

In this contribution, by deriving an equivalent signal model, we show that classical model order selection algorithms (MOS) based on the analysis of the sample covariance matrix can be applied. However, under a strong limitation on the sample size, the performance of the available MOS algorithms depends on the knowledge of the noise model and may deteriorate significantly when the actual noise statistics are different. In this contribution, we propose an alternative approach that explicitly accounts for the measurement process. It does so by exploiting an empirical distribution of the noise eigenvalues obtained during a training period, i.e., when only noise is received. Numerical comparison of the proposed algorithm, which we refer to as empirical eigenvalue-threshold test (EET), with state-of-the-art MOS algorithms shows that EET performs better for small sample sizes and a low SNR.

It is worth noting that the equivalent signal model that allows for classical MOS (and the EET) is based on the availability of multiple snapshots of the mixture of signals and the fact that the signals are incoherent (which implies that they must change in time, e.g., be randomly modulated signals). Although, the general CS setup does not impose any restrictions on the signal but its sparsity, there are applications where the aforementioned assumptions hold. Examples of such applications include sub-Nyquist sampling of multiband signals, compressive signal localization, and radar signal processing.

The remainder of the paper is organized as follows: a compressed sensing data model is introduced in Section 2, followed by the analysis of an eigenvalue-based sparsity level estimation in Section 3. The proposed empirical eigenvalue-threshold test (EET) is described in Section 4. Section 5 presents numerical results for a comparison between the proposed EET algorithm with state-of-the-art MOS schemes. Finally, Section 6 concludes the paper.

2. DATA MODEL AND PROBLEM FORMULATION

We consider a discrete compressed sensing formulation of the following form

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}^{\mathrm{T}} \cdot \boldsymbol{s}(t) + \boldsymbol{n}_{\mathrm{y}}(t) = \boldsymbol{\Phi}^{\mathrm{T}} \cdot \boldsymbol{A} \cdot \boldsymbol{x}(t) + \boldsymbol{n}_{\mathrm{y}}(t), \quad (1)$$

where $\boldsymbol{y}(t) \in \mathbb{C}^{M\times 1}$ are the compressed observations at the time t of a signal $\boldsymbol{s}(t) \in \mathbb{C}^{K\times 1}$ that is sparse in a basis $\boldsymbol{A} \in \mathbb{C}^{K\times K}$ with coefficients $\boldsymbol{x}(t) \in \mathbb{C}^{K\times 1}$, i.e., $\boldsymbol{x}(t)$ contains $N \ll K$ non-zeros only. We assume that the support, i.e., the positions of the non-zero elements in $\boldsymbol{x}(t)$ is constant over a certain observation time window and that the different sequences in the vector $\boldsymbol{x}(t)$ are incoherent to each other (as it is the case, e.g., for randomly modulated signals). The matrix $\boldsymbol{\Phi} \in \mathbb{C}^{K\times M}$ in (1) is the measurement matrix with $K > M$, where $(\cdot)^{\mathrm{T}}$ denotes matrix transpose, and $\boldsymbol{n}_{\mathrm{y}}(t) \in \mathbb{C}^{M\times 1}$ represents the additive noise.

In the CS setting there are different types of noise. As discussed in [5], we could have "signal noise" that is added to $\boldsymbol{s}(t)$ (or, equivalently to $\boldsymbol{x}(t)$) or "measurement noise" that is added to $\boldsymbol{y}(t)$. In the considered applications, e.g. CS for multiband signal acquisition, the received signal inevitably contains both of them. Therefore, we model the noise $\boldsymbol{n}_{\mathrm{y}}(t)$ as

$$\boldsymbol{n}_{\mathrm{y}}(t) = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{n}_{\mathrm{s}}(t) + \boldsymbol{n}_{\mathrm{m}}(t), \quad (2)$$

where $\boldsymbol{n}_{\mathrm{s}}(t) \in \mathbb{C}^{K\times 1}$ and $\boldsymbol{n}_{\mathrm{m}}(t) \in \mathbb{C}^{M\times 1}$.

Introducing a short-hand notation for the sensing matrix according to $\boldsymbol{B} = \boldsymbol{\Phi}^{\mathrm{T}} \cdot \boldsymbol{A} \in \mathbb{C}^{M\times K}$, (1) becomes

$$\boldsymbol{y}(t) = \boldsymbol{B} \cdot \boldsymbol{x}(t) + \boldsymbol{n}_{\mathrm{y}}(t). \quad (3)$$

We are interested in estimating the sparsity order N from the compressed observations $\boldsymbol{y}$ directly.

3. EIGENVALUE BASED SPARSITY LEVEL ESTIMATION

To this end, we consider the covariance matrix $\boldsymbol{R}_{\mathrm{y}}$ which is defined as

$$\boldsymbol{R}_{\mathrm{y}} = \mathrm{E}\{\boldsymbol{y}(t)\boldsymbol{y}(t)^{\mathrm{H}}\}, \quad (4)$$

where $(\cdot)^{\mathrm{H}}$ denotes Hermitian transpose. Inserting (3) into (4) we obtain

$$\boldsymbol{R}_{\mathrm{y}} = \boldsymbol{B} \cdot \boldsymbol{R}_{\mathrm{x}} \cdot \boldsymbol{B}^{\mathrm{H}} + \boldsymbol{R}_{\mathrm{n_y}}, \quad (5)$$

with $\boldsymbol{R}_{\mathrm{x}}$ being the covariance matrix of $\boldsymbol{x}$ and $\boldsymbol{R}_{\mathrm{n_y}}$ the noise covariance of $\boldsymbol{n}_{\mathrm{y}}$.

Assuming the signal and the measurement noise to be independent random processes, the noise covariance $\boldsymbol{R}_{\mathrm{n_y}}$ can be written as

$$\boldsymbol{R}_{\mathrm{n_y}} = \boldsymbol{\Phi}^{\mathrm{T}} \cdot \boldsymbol{R}_{\mathrm{n_s}} \cdot \boldsymbol{\Phi}^{*} + \boldsymbol{R}_{\mathrm{n_m}}, \quad (6)$$

where * represents complex conjugation, while $\boldsymbol{R}_{\mathrm{n_s}}$ and $\boldsymbol{R}_{\mathrm{n_m}}$ are the covariance matrices of signal and measurement noise, respectively. Equations (2) and (6) show that the noise covariance $\boldsymbol{R}_{\mathrm{n_y}}$ will depend on the measurement matrix $\boldsymbol{\Phi}$, the signal noise covariance matrix $\boldsymbol{R}_{\mathrm{n_s}}$ and the measurement noise covariance matrix $\boldsymbol{R}_{\mathrm{n_m}}$.

When the covariance matrix $\boldsymbol{R}_{\mathrm{n_y}}$ is fully known at the receiver, we can perform prewhitening to the output vector $\boldsymbol{y}(t)$. After the prewhitening stage, our observation model (3) is transformed into

$$\boldsymbol{z}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{n}_{\mathrm{z}}(t), \quad (7)$$

where $\boldsymbol{C} = \left(\boldsymbol{R}_{\mathrm{n_y}}\right)^{-1/2}\boldsymbol{B}$ and $\boldsymbol{n}_{\mathrm{z}}(t)$ is a white noise vector.

Due to the prewhitening stage, the covariance matrix of the whitened observations $\boldsymbol{z}$ is given by

$$\boldsymbol{R}_{\mathrm{z}} = \boldsymbol{C}\boldsymbol{R}_{\mathrm{x}}\boldsymbol{C}^{\mathrm{H}} + \mathbf{I}_M, \quad (8)$$

where $\boldsymbol{R}_{\mathrm{x}} \in \mathbb{C}^{K\times K}$ is a covariance matrix of the input signal $\boldsymbol{x}(t)$. Note that under the assumptions on the $\boldsymbol{x}(t)$ described in Section 2 and since $\boldsymbol{x}(t)$ is N-sparse, the rank of $\boldsymbol{R}_{\mathrm{x}}$ is only $N \ll K$. Let $\lambda_{\mathrm{z},1} \geq \lambda_{\mathrm{z},2} \geq \ldots \geq \lambda_{\mathrm{z},M}$ denote the ordered set of eigenvalues of $\boldsymbol{R}_{\mathrm{z}}$. We then have

$$\lambda_{\mathrm{z},m} = \begin{cases} \lambda_{\mathrm{s},m} + 1, & 1 \leq m \leq N \\ \\ 1, & N+1 \leq m \leq M, \end{cases} \quad (9)$$

where $\lambda_{\mathrm{s},m}$ denotes the ordered set of N non-zero eigenvalues of the "signal" component of $\boldsymbol{R}_{\mathrm{z}}$ given by $\boldsymbol{C}\boldsymbol{R}_{\mathrm{x}}\boldsymbol{C}^{\mathrm{H}}$. The concrete values of $\lambda_{\mathrm{s},m}$ depend on the correlation between the different signals in $\boldsymbol{x}(t)$ as well as the matrix $\boldsymbol{C}$. Based on (9), the sparsity level N would simply be given by the number of eigenvalues that are greater than one.

However, the covariance matrix $\boldsymbol{R}_{\mathrm{z}}$ is not known in practice, but it has to be estimated. Given a limited number of snapshots $t = 1, 2, \ldots, T$, let us denote $\boldsymbol{Z} = [\boldsymbol{z}(1), \boldsymbol{z}(2), \cdots, \boldsymbol{z}(T)] \in \mathbb{C}^{M\times T}$, $\boldsymbol{X} = [\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(T)] \in \mathbb{C}^{K\times T}$, and $\boldsymbol{N}_{\mathrm{z}} = [\boldsymbol{n}_{\mathrm{z}}(1), \boldsymbol{n}_{\mathrm{z}}(2), \cdots, \boldsymbol{n}_{\mathrm{z}}(T)] \in \mathbb{C}^{M\times T}$. The covariance matrix $\boldsymbol{R}_{\mathrm{z}}$ can be estimated from $\boldsymbol{Z}$ as

$$\hat{\boldsymbol{R}}_{\mathrm{z}} = \frac{1}{T}\boldsymbol{Z} \cdot \boldsymbol{Z}^{\mathrm{H}} = \boldsymbol{C}\hat{\boldsymbol{R}}_{\mathrm{x}}\boldsymbol{C}^{\mathrm{H}} + \hat{\boldsymbol{R}}_{\mathrm{n_z}} + \hat{\boldsymbol{R}}_{\mathrm{x,n_z}}, \quad (10)$$

where $\hat{\boldsymbol{R}}_{\mathrm{n_z}} = \frac{1}{T}\boldsymbol{N}_{\mathrm{z}}\boldsymbol{N}_{\mathrm{z}}^{\mathrm{H}}$ is the sample noise covariance matrix and $\hat{\boldsymbol{R}}_{\mathrm{x,n_z}}$ is a cross term defined as

$$\hat{\boldsymbol{R}}_{\mathrm{x,n_z}} = \frac{1}{T}\Big((\boldsymbol{C}\boldsymbol{X})\boldsymbol{N}_{\mathrm{z}}^{\mathrm{H}} + \boldsymbol{N}_{\mathrm{z}}\left(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{C}^{\mathrm{H}}\right)\Big). \quad (11)$$

Let the eigenvalues of the sample covariance matrix $\hat{\boldsymbol{R}}_{\mathrm{z}}$ be given by $\hat{\lambda}_{\mathrm{z},m}$, $m = 1, 2, \ldots, M$. Due to the limited number of observations, the estimated eigenvalues $\hat{\lambda}_{\mathrm{z},m}$ differ significantly from the ideal eigenvalue profile shown in (9). Firstly, since $\hat{\boldsymbol{R}}_{\mathrm{n_z}} \neq \boldsymbol{I}_M$, the noise eigenvalues are not equal to one but vary around one (which leads to a decaying profile in the ordered set of eigenvalues). Secondly, the cross term $\hat{\boldsymbol{R}}_{\mathrm{x,n_z}}$ between the signal and the noise becomes non-vanishing.

At this point classical model order selection algorithms (MOS) as, for instance, [6–8] can be applied in order to discriminate between the signal and noise eigenvalues. However, such algorithms heavily rely on the assumption that the noise $\boldsymbol{n}_{\mathrm{z}}(t)$ is indeed white. In order to perform prewhitening according to (7), the noise covariance matrix $\boldsymbol{R}_{\mathrm{n_y}}$ has to be known. For instance, if both the signal and the measurement noise from (2) are known to be white with elements that have known common variances σ_{s}^2 and σ_{m}^2 for $\boldsymbol{n}_{\mathrm{s}}$ and $\boldsymbol{n}_{\mathrm{m}}$, respectively, $\boldsymbol{R}_{\mathrm{n_y}}$ can be computed simply as

$$\boldsymbol{R}_{\mathrm{n_y}} = \sigma_{\mathrm{s}}^2 \cdot \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{\Phi}^* + \sigma_{\mathrm{m}}^2 \cdot \mathbf{I}_M, \tag{12}$$

where $\mathbf{I}_M$ being an $M \times M$ identity matrix. However, in a more general case, e.g., when the noise statistics is not known a priori, $\boldsymbol{R}_{\mathrm{n_y}}$ has to be estimated in advance. Practically, this would require collecting a training set $\boldsymbol{N}_{\mathrm{y}}^{\mathrm{tr}} = [\boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(1), \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(2), \cdots, \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(L_{\mathrm{tr}})] \in \mathbb{C}^{M \times L_{\mathrm{tr}}}$ of noise samples. The set $\boldsymbol{N}_{\mathrm{y}}^{\mathrm{tr}}$ can be obtained during a calibration stage from a portion of the data that is known to contain only noise and no signal. In the following section we propose an approach for sparsity level detection that makes use of these training data for estimation of the noise eigenvalues distribution.

4. EMPIRICAL EIGENVALUE-THRESHOLD TEST

We formulate the sparsity level estimation problem as a set of binary hypothesis tests. For each test eigenvalue $\hat{\lambda}_{\mathrm{z},m}$ of the sample covariance matrix $\hat{\boldsymbol{R}}_{\mathrm{z}}$, the following hypothesis are tested

$$\begin{aligned} \mathcal{H}_{0,m}: &\quad \hat{\lambda}_{\mathrm{z},m} \in \mathcal{S}_{\mathrm{n}} \\ \mathcal{H}_{1,m}: &\quad \hat{\lambda}_{\mathrm{z},m} \in \mathcal{S}_{\mathrm{nx}}, \end{aligned} \tag{13}$$

where $\mathcal{S}_{\mathrm{n}}$ and $\mathcal{S}_{\mathrm{nx}}$ are sets of noise only and noise plus signal eigenvalues, respectively. Taking into account that the test eigenvalues $\hat{\lambda}_{\mathrm{z},m}$ are sorted in a descending order, the sparsity level then is estimated simply as

$$\hat{N} = \max_{m:\{\hat{\lambda}_{\mathrm{z},m} \in \mathcal{S}_{\mathrm{nx}}\}} (m). \tag{14}$$

To differentiate between the two hypotheses, a classical Neyman-Pearson (NP)-based detector can be used. The NP detector maximizes the probability of correct detection P_{d} for a fixed probability of false alarm P_{fa}. Let us denote the desired probability of false alarm as α. The decision rule for (13) can then be formulated as

$$\lambda_{\mathrm{z},m} \underset{\mathcal{H}_{0,m}}{\overset{\mathcal{H}_{1,m}}{\gtrless}} \eta_m, \quad \text{where} \quad \eta_m = \bar{F}_{\mathcal{H}_0,m}(\alpha), \tag{15}$$

and $\bar{F}_{\mathcal{H}_0,m}$ is a complementary cumulative distribution function (CCDF) of the probability density function (PDF) $f_{\mathcal{H}_0,m}(\lambda_{\mathrm{z},m})$ corresponding to the hypotheses $\mathcal{H}_{0,m}$.

The direct usage of (15) requires the knowledge of the PDFs $f_{\mathcal{H}_0,m}(\lambda_{\mathrm{z},m})$. There is a large amount of results available for the asymptotic distributions of the sample eigenvalues $\lambda_{\mathrm{z},m}$ under the hypothesis $\mathcal{H}_{0,m}$ ("noise only") for the case of white Gaussian noise [7, 9, 10]. Recent achievements in random matrix theory allowed to extend some of the results available for the white noise to the case of colored noise as well [11]. However, these asymptotic expressions are derived based on the limit theorems as of certain parameters tend to infinity (for instance M or T, or both of them). The performance of the algorithms based on such asymptotic estimates deteriorates for limited signal dimensions. Therefore, we propose to use actual noise samples obtained during the training period (as discussed in the end of Section 3) for the calculation of the empirical distribution of the noise eigenvalues as an approximation of $f_{\mathcal{H}_0,m}(\lambda_{\mathrm{z},m})$. Hence, it explicitly accounts for both the actual signal dimensions and measurement process.

In this way, during the training period, a set of L noise eigenvalue profiles $\hat{\boldsymbol{\lambda}}_{\mathrm{n_z}}^{(\ell)} \in \mathbb{R}^M$ is obtained from the $\hat{\boldsymbol{R}}_{\mathrm{n_y}}^{(\ell)} = \frac{1}{T}[\boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(1), \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(2), \cdots, \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(T)][\boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(1), \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(2), \cdots, \boldsymbol{n}_{\mathrm{y}}^{\mathrm{tr}}(T)]^{\mathrm{H}}$ where $\ell = 1, 2, \ldots, L$ and $L = \lfloor L_{\mathrm{tr}}/T \rfloor$. These are stacked into one vector $\boldsymbol{\xi} = \left[\hat{\boldsymbol{\lambda}}_{\mathrm{n_z}}^{(1)^{\mathrm{T}}}, \hat{\boldsymbol{\lambda}}_{\mathrm{n_z}}^{(2)^{\mathrm{T}}}, \ldots, \hat{\boldsymbol{\lambda}}_{\mathrm{n_z}}^{(L)^{\mathrm{T}}}\right]^{\mathrm{T}} \in \mathbb{R}^{ML}$. Let us denote $\tau = (\max_j(\xi_j) - \min_j(\xi_j))/Q$, where $Q \in \mathbb{N}$. The empirical distribution of the noise eigenvalues $\hat{\lambda}_{\mathrm{n_z}}$ is then estimated from $\boldsymbol{\xi}$ as

$$\hat{f}(\hat{\lambda}_{\mathrm{n_z}}) = \sum_{q=1}^{Q} P_q \delta\left(\lambda_{\mathrm{n_z}} - (q - 0.5)\tau - \xi_{\min}\right), \tag{16}$$

where $\delta(\lambda_{\mathrm{n_z}})$ is a Dirac delta function, $\xi_{\min} = \min_j(\xi_j)$ and P_q is

$$\begin{aligned} P_q &= \Pr[\xi_{\min} + \tau(q-1) \leq \hat{\lambda}_{\mathrm{n_z}} < \xi_{\min} + \tau q] = \\ &= \frac{1}{ML} \sum_{\{j:\xi_{\min} + \tau(q-1) \leq \xi_j < \xi_{\min} + \tau q\}} 1, \end{aligned} \tag{17}$$

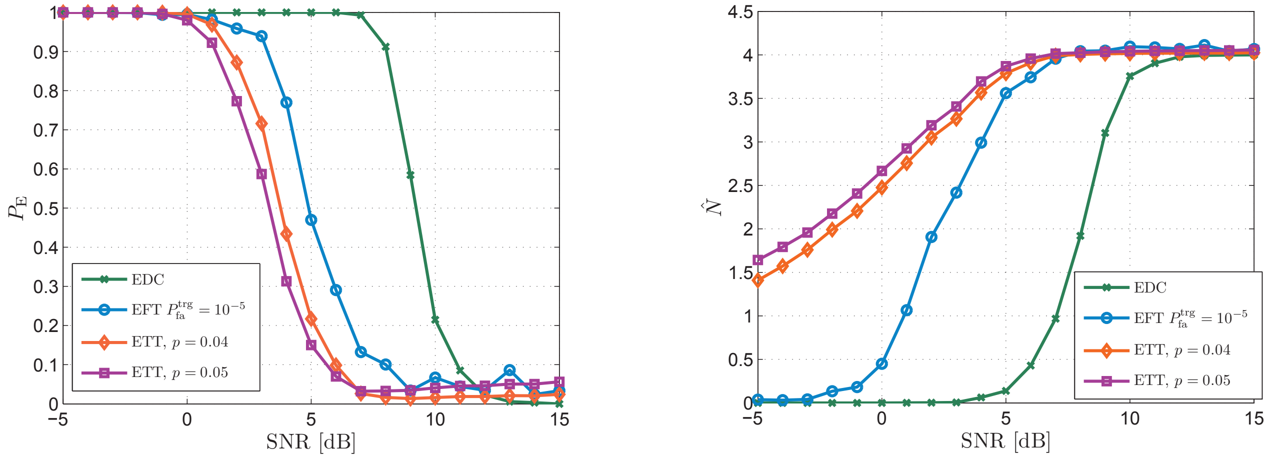
with $q = 1, 2, \ldots, Q$ and $j = 1, 2, \ldots, ML$.

A unified threshold $\eta_m = \eta$ for a decision rule in (15) is derived by setting a parameter p so that

$$\eta = \xi_{\min} + (j_\eta - 0.5)\tau, \tag{18}$$

where

$$j_\eta = \arg \min_{i=1,2,\ldots,Q} \left| \left(\sum_{q=i}^{Q} P_q \right) - p \right|. \tag{19}$$



(a) Probability of wrong estimation P_E

(b) Mean estimated sparsity level $\hat{N}$

Fig. 1: Probability of wrong estimation P_E (a) and estimated sparsity $\hat{N}$ (b) as functions of the SNR for $T = M$

The parameter p can be seen as an analog of the parameter α from (15). It asymptotically approaches the true probability of false alarm with increasing L and increasing number of snapshots T.

5. NUMERICAL RESULTS

For comparison of the proposed approach with the classical MOS algorithms, we performed a series of Monte-Carlo simulations for the following tests:

- the information-theoretic-based Efficient Detection Criterion (EDC) [6],
- the Exponential Fitting Test (EFT) which exploits the exponential profile of the ordered noise eigenvalues learned from synthetically created noise samples [8],
- the proposed Empirical Eigenvalue Threshold (EET) test described in Section 4.

Throughout the simulations, both the signal noise $n_s(t)$ and measurement noise n_m were modeled as i.i.d. circularly symmetric complex Gaussian white noise with variances $\sigma_s^2 = \sigma_m^2 = \sigma_0^2$, where the total SNR is defined as $1/((K+M)\sigma_0^2)$. The matrix $\boldsymbol{B}$ from (3) was chosen randomly with entries drawn from an i.i.d. $\mathcal{CN}(0, 1/K)$ distribution. The values of the parameters K, M and N are listed in Table 1, where the number of snapshots T used for calculation of the covariance matrix $\hat{\boldsymbol{R}}_z$ was equal to M.

To assess how often the aforementioned algorithms obtain the correct result, we calculate the probability of wrong estimation P_{er}, which is given as the percentage of the trials when $\hat{N} \neq N$. Additionally, in order to obtain deeper insight into the nature of the error (i.e., whether the test tends to over- or underestimate), we calculate mean estimated sparsity $\hat{N}$ and a posteriori probabilities of false alarm P_{fa}^{a} and mis-detection P_{md}^{a} defined as the percentage of the trials when $\hat{N} > N$ and $\hat{N} < N$, respectively.

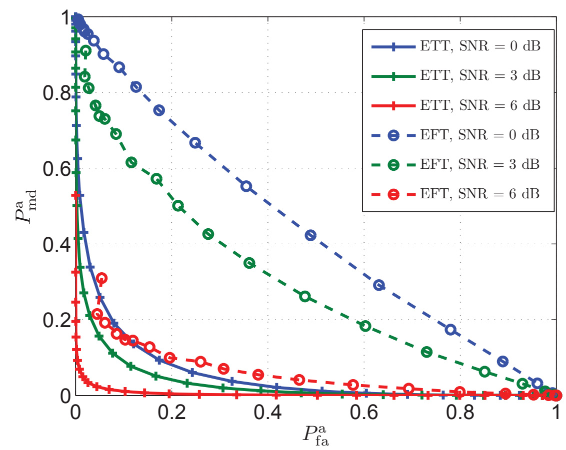


Fig. 2: Operating characteristic P_{md}^{a} vs P_{fa}^{a} for $T = M$

Figure 1 shows the probability of wrong estimation P_{er} and the mean estimated sparsity $\hat{N}$ as functions of the SNR for the two considered MOS algorithms and the proposed EET algorithm with $p = 0.04$ and $p = 0.05$, where the parameter p was tuned heuristically. From Figure 1a it is seen that the proposed EET algorithm outperforms both EDC and EFT in the low SNR regime. According to Figure 1b, all three considered algorithms tend to underestimate the sparsity level in the low SNR regime with proposed ETT test providing significantly better performance.

In order to compare operating characteristic of the EFT and ETT within a wide range of parameters p and F_{fa}^{trg} (specified in Table 1), Figure 2 presents a posteriori probabilities of false alarm and mis-detection. It shows that in the considered

Parameter	K	M	N	p	$P_{\mathrm{fa}}^{\mathrm{trg}}$
Value	100	20	4	[0.01 0.3]	$[10^{-8}\ 10^{-1}]$

Table 1: List of parameters used for simulations.

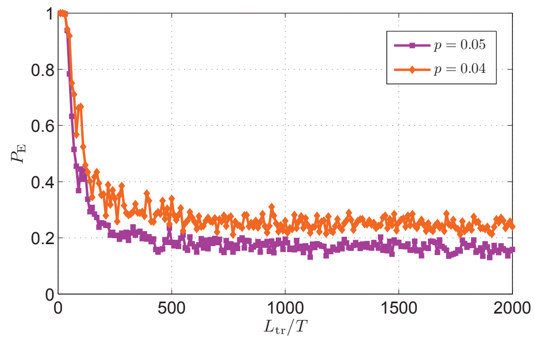


Fig. 3: Probability of wrong estimation P_{E} as a function of L for $T = M$, SNR $= 5$ dB

low SNR regime (SNR$<$ 6 dB) ETT provides significantly lower probability of the sparsity level underestimation for the fixed probability of its overestimation. Although, the strategy for finding an optimal value of the parameter p requires further investigation,

Note that for the previous results, the number of observations L_{tr} used to obtain the training statistics for the prewhitening stage and the EET was fixed to $L_{\mathrm{tr}} = 500T$. Thus, Figure 3 demonstrates the influence of the size of the training set on the performance of the EET. It shows that the probability of error decreases with increasing L_{tr} but only mildly. This means that our test requires only a small number of training samples to obtain the suitable thresholds.

6. CONCLUSION

In this paper we examined the problem of the estimation of the sparsity level from the analysis of the compressed covariance matrix. Working in the compressed domain has the advantage that no additional signal reconstruction is necessary. By deriving an equivalent system model, we show that state-of-the-art model order selection schemes can be applied, provided that several snapshots of the incoherent in the sparse domain signals are available. However, this techniques are impaired by how the measurement process influences the distribution of the noise eigenvalues. As a solution, we propose the EET algorithm which exploits the empirical distribution of the noise eigenvalues obtained during a training period. Numerical comparisons of the proposed algorithm with state-of-the-art model order selection schemes reveal its superiority in terms of the probability of wrong estimation and the mean estimation error for a low number of snapshots and a low SNR.

REFERENCES

[1] P. T. Boufounos, M. F. Duarte, and R. G. Baraniuk, “Sparse signal reconstruction from noisy compressive measurements using cross validation,” in *Proceedings of the IEEE Workshop on Statistical Signal Processing (SSP)*, Madison, WI, Aug. 2007, pp. 299–303.

[2] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, “Signal processing with compressive measurements,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 445–460, 2010.

[3] Z. Wang, G.R. Arce, and B.M. Sadler, “Subspace compressive detection for sparse signals,” in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, March 2008, pp. 3873–3876.

[4] C. D. Austin, R. L. Moses, J. N. Ash, and E. Ertin, “On the relation between sparse reconstruction and parameter estimation with model order selection,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 3, pp. 560–570, 2010.

[5] M. A. Davenport, J. N. Laska, J. R. Treichler, and R. G. Baraniuk, “The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 9, pp. 4628–4642, 2012.

[6] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai, “On detection of the number of signals in presence of white noise,” *Journal of multivariate analysis*, vol. 20, no. 1, pp. 1–25, 1986.

[7] W. Chen, K. M. Wong, and J. P Reilly, “Detection of the number of signals: A predicted eigen-threshold approach,” *Signal Processing, IEEE Transactions on*, vol. 39, no. 5, pp. 1088–1098, 1991.

[8] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, “Model order selection for short data: An exponential fitting test (eft),” *EURASIP Journal on Advances in Signal Processing*, vol. 2007, 2006.

[9] R.R. Nadakuditi and A. Edelman, “Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 7, pp. 2625–2638, July 2008.

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

[11] R. R. Nadakuditi and J. W. Silverstein, “Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, pp. 468–480, 2010.