07 - 10 September 2009

PROCEEDINGS

54. IWK

Internationales Wissenschaftliches Kolloquium
International Scientific Colloquium

Information Technology and Electrical Engineering - Devices and Systems, Materials and Technologies for the Future

Faculty of Electrical Engineering and Information Technology

Startseite / Index:
http://www.db-thueringen.de/servlets/DocumentServlet?id=14089

TECHNISCHE UNIVERSITÄT
ILMENAU

Impressum

Herausgeber: Der Rektor der Technischen Universität Ilmenau
Univ.-Prof. Dr. rer. nat. habil. Dr. h. c. Prof. h. c.
Peter Scharff

Redaktion: Referat Marketing
Andrea Schneider

Fakultät für Elektrotechnik und Informationstechnik
Univ.-Prof. Dr.-Ing. Frank Berger

Redaktionsschluss: 17. August 2009

Technische Realisierung (USB-Flash-Ausgabe):
Institut für Medientechnik an der TU Ilmenau
Dipl.-Ing. Christian Weigel
Dipl.-Ing. Helge Drumm

Technische Realisierung (Online-Ausgabe):
Universitätsbibliothek Ilmenau
ilmedia
Postfach 10 05 65
98684 Ilmenau

Verlag:
Verlag ISLE, Betriebsstätte des ISLE e.V.
Werner-von-Siemens-Str. 16
98693 Ilmenau

ISBN (USB-Flash-Ausgabe): 978-3-938843-45-1
ISBN (Druckausgabe der Kurzfassungen): 978-3-938843-44-4

Startseite / Index:
http://www.db-thueringen.de/servlets/DocumentServlet?id=14089

COMPARISON OF MODEL ORDER SELECTION TECHNIQUES FOR HIGH-RESOLUTION PARAMETER ESTIMATION ALGORITHMS

João Paulo C. L. da Costa[1], Arpita Thakre[2], Florian Roemer, and Martin Haardt

Ilmenau University of Technology
Communications Research Laboratory
P.O. Box 100565, D-98684 Ilmenau, Germany
{joaopaulo.dacosta,florian.roemer,martin.haardt}@tu-ilmenau.de, [2] arpita@tenet.res.in

ABSTRACT

In sensor array processing it is often required to know the number of signals received by an antenna array, since in practice only a limited number of observations is available. Robust techniques for the estimation of the model order are needed.

In this paper, we propose general application rules for the most recent model order selection techniques in the literature considering different one-dimensional scenarios. Other important contributions are a more general and improved form of the modified exponential fitting test (M-EFT) and extensions of other known model order selection techniques for the case that the number of sensors is greater than the number of snapshots.

Index Terms— Model Order Selection, Principal Component Analysis, Signal Enumeration

1. INTRODUCTION

In many signal processing applications, including radar, sonar, communications, channel modeling, medical imaging, and the estimation of the parameters of the dominant multipath components from measurements, the model order selection is a fundamental step. It allows us to separate the only noise components from the principal components applying a rank reduction of the data. For many parameter estimation techniques such a rank reduction approach is crucial.

Therefore, as important contributions in this article, we propose an improved form of the modified exponential fitting test (M-EFT) [1, 2, 3] and we compare its performance to recent model order estimation techniques available in the literature. Furthermore, after extensive simulation campaigns, we also propose rules for the application of model order selection schemes.

The remainder of this article is organized as follows. After reviewing the notation in Section 2, the data model is presented as an example for model order estimation problems in Section 3. A short description of some of the most important recent methods for model order selection is provided in Section 4. In Section 5, an extension of the modified exponential fitting test (M-EFT) is proposed. In Section 6 we perform simulations in order to compare the Probability of Detection (PoD) of the different techniques mentioned in this article. General rules for the application of state-of-the-art model order selection schemes are described in Section 7. In Section 8 conclusions are drawn.

2. NOTATION

In order to facilitate the distinction between scalars and matrices, the following notation is used: scalars are denoted as italic letters ($a, b, \ldots, A, B, \ldots, \alpha, \beta, \ldots$), column vectors as lower-case bold-face letters ($\boldsymbol{a}, \boldsymbol{b}, \ldots$) and matrices as bold-face capitals ($\boldsymbol{A}, \boldsymbol{B}, \ldots$). Lower-order parts are consistently named: the (i, j)-element of the matrix $\boldsymbol{A}$, is denoted as $a_{i,j}$.

We use the superscripts ^{T},^{H},$^{-1}$,$^+$ and * for transposition, Hermitian transposition, matrix inversion, the Moore-Penrose pseudo inverse of matrices, and complex conjugation, respectively.

3. DATA MODEL

Consider the observation of a linear mixture of d sources which can be written as

$$\boldsymbol{X} = \boldsymbol{A} \cdot \boldsymbol{S} + \boldsymbol{N}, \quad (1)$$

where the mixing matrix $\boldsymbol{A} \in \mathbb{C}^{M \times d}$ contains d linearly independent vectors $\boldsymbol{a}_i \in \mathbb{C}^{M \times 1}$, $\boldsymbol{S} \in \mathbb{C}^{d \times N}$ contains the source symbols $s_i(n)$, where $i = 1, ..., d$ and $n = 1..., N$, and the noise matrix $\boldsymbol{N} \in \mathbb{C}^{M \times N}$ contains ZMCSCG (zero-mean circularly symmetric Gaussian) elements with variance σ_n^2.

[1] João Paulo C. L. da Costa is a scholarship holder of the National Counsel of Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq) of the Brazilian Government and also a First Lieutenant of the Brazilian Army (Exército Brasileiro).

[2] Arpita Thakre from Indian Institute of Technology Madras is on leave at the Ilmenau University of Technology as a research assistant.

For the general data model of (1), it is possible to estimate the model order via eigenvalue based techniques, such as EFT [1, 2, 3], Akaike's Information theoretic Criterion (AIC) [4, 5], and the Minimum Description Length (MDL) criterion [6, 5].

On the other hand, for the techniques not based on the eigenvalues profile, but on the shift invariance equation, e.g., ESTimation ERror (ESTER) [7] and Subspace-based Automatic Model Order Selection (SAMOS) [8], we use a more specific data model, namely the superposition of d planar wavefronts received by a uniform linear array (ULA) of M sensors during N subsequent time instants. The measurement samples are given by

$$x_{m,n} = \sum_{i=1}^{d} s_i(n) \cdot e^{j\cdot(m-1)\cdot\mu_i} + n_{m,n}, \qquad (2)$$

where $m = 1, 2, \ldots, M$, $n = 1, 2, \ldots, N$, $s_i(n)$ denotes the complex amplitude of the i-th exponential at time instant n, μ_i symbolizes the spatial frequency of the i-th exponential, and $n_{m,n}$ models the additive noise component inherent in the measurement process.

Note that (1) is a special case of (2) where the i-th column of $\boldsymbol{A}$ is equal to $[1, e^{j\cdot\mu_i}, ..., e^{j\cdot(M-1)\mu_i}]^{\mathrm{T}}$.

For the computation of the eigenvalues, λ_i, we can simply apply the eigenvalue decomposition of the sample covariance matrix $\hat{\boldsymbol{R}}_{xx}$ obtained from the measurement matrix $\boldsymbol{X}$ via

$$\hat{\boldsymbol{R}}_{xx} = \frac{1}{N}\,\boldsymbol{X}\cdot\boldsymbol{X}^{H} \in \mathbb{C}^{M\mathrm{x}M}. \qquad (3)$$

In the noiseless case, we would have only d non-zero eigenvalues, however, since the data is contaminated by noise, we have $\min(M, N)$ non-zero eigenvalues. Therefore, it is our goal to estimate the number d.

In this article, we assume that the number of sources to be estimated is smaller than the number of observations N and the number of sensors M, i.e., $d \le \min\{M, N\}$.

4. MODEL ORDER SELECTION TECHNIQUES

State-of-the-art model order estimation techniques based on the eigenvalues include Akaike's Information theoretic Criterion (AIC) [4, 5], the Minimum Description Length (MDL) criterion [6, 5], the Efficient Detection Criterion (EDC) [9], the method proposed in [10] denoted here as Nadakuditi Edelman Model Order selection (NEMO) scheme, Stein's Unbiased Risk Estimator (SURE) [11] and the method proposed in [12] named here as RADOI.

In AIC, MDL and EDC, the information criterion is a function of the geometric mean, $g(k)$, and arithmetic mean, $a(k)$, of the k smallest eigenvalues of (3) respectively, and k is a candidate value for d.

In [3], we have shown modifications of AIC, MDL and EDC for the case that $M > N$, which we denote here by 1-D AIC and 1-D MDL. Basically the difference between AIC, MDL and EDC is the penalty function $p(k, N, \alpha)$, therefore these techniques can be written in the following general form

$$\hat{d} = \arg\min_k J(k) \quad \text{where} \qquad (4)$$
$$J(k) = -N(\alpha - k)\log\left(\frac{g(k)}{a(k)}\right) + p(k, N, \alpha),$$

where $\hat{d}$ represents an estimate of the model order d. The penalty functions for 1-D AIC, 1-D MDL and 1-D EDC are given by $p(k, N, \alpha) = k(2\alpha - k)$, $p(k, N, \alpha) = \frac{1}{2}k(2\alpha - k)\log(N)$, and $p(k, N, \alpha) = \frac{1}{2}k(2\alpha - k)\sqrt{N\ln(\ln N)}$ respectively. According to [3] $\alpha = \min[M, N]$, while according to [10], we should use $\alpha = M$, and $0 \le k \le \min[M, N]$. Note that we apply in this article the same penalty function used in [7] for the EDC technique. Note also that in EDC a general penalty function is proposed such that this function should obey some restrictions, thereby AIC and MDL are specific cases of EDC.

In the Nadakuditi Edelman Model Order selection (NEMO) scheme, the cost function to be minimized is:

$$\hat{d} = \arg\min_k \mathrm{NEMO}(k) \quad \text{where} \qquad (5)$$
$$\mathrm{NEMO}(k) = \frac{\psi}{4}\left[\frac{N}{M}\right]t_k^2 + 2(k+1) \quad ,$$
$$t_k = \left[(M-k)\frac{\sum_{i=k+1}^{M}\lambda_i^2}{\left(\sum_{i=k+1}^{M}\lambda_i\right)^2} - \left(1 + \frac{M}{N}\right)\right]N - \left(\frac{2}{\psi} - 1\right)\frac{M}{N} \quad ,$$

$\psi = 1$ if $\boldsymbol{X} \in \mathbb{R}^{M\times N}$ and $\psi = 2$ if $\boldsymbol{X} \in \mathbb{C}^{M\times N}$. In NEMO, we have that $0 \le k < \min(M, N)$.

In the SURE criterion, the risk $\hat{R}(k)$ must be minimized according to the following expression:

$$\hat{d} = \arg\min_k \hat{R}(k) \quad \text{where} \qquad (6)$$
$$\hat{R}(k) = (M-k)\hat{\sigma}_k^2 + 2\sigma^2 k + (\hat{\sigma}_k^4 - 2\hat{\sigma}_k^2\sigma^2 + \frac{4\hat{\sigma}_k^2\sigma^2}{N})\sum_{i=1}^{k}\frac{1}{\lambda_i}$$
$$+ \frac{4\sigma^2}{N}\sum_{i=1}^{k}\sum_{j=k+1}^{M}\frac{\lambda_i - \hat{\sigma}_k^2}{\lambda_i - \lambda_j} + \frac{2\sigma^2}{N}k(k-1)$$
$$- \frac{2\sigma^2}{N}(M-1)\sum_{i=1}^{k}(1 - \frac{\hat{\sigma}_k^2}{\lambda_i})$$

where $\hat{\sigma}_k^2 = \dfrac{1}{M-k}\displaystyle\sum_{i=r+1}^{M}\lambda_i$ and σ^2 is estimated via random matrix theory using the algorithm also derived in [11]. In order to apply SURE, it is necessary to assure that a certain percentage, e.g., 25 %, of the smallest eigenvalues is only composed of noise.

In [11], SURE outperforms the Laplace [13] and BIC [14] methods in terms of the Probability of Detection (PoD).

The RADOI model order selection scheme is an empirical approach [12], and its cost function is given by

$$\hat{d} = \arg\min_k \mathrm{RADOI}(k) \quad \text{where} \tag{7}$$

$$\mathrm{RADOI}(k) = \lambda_{k+1} \cdot \left(\sum_{i=2}^{M} \lambda_i \right)^{-1} - \xi_k \cdot \left(\sum_{i=1}^{M-1} \xi_i \right)^{-1}, \tag{8}$$

where $\xi_k = 1 - \frac{\alpha \cdot (\lambda_k - \mu_k)}{\mu_k}$, $\mu_k = \frac{1}{M-k} \cdot \sum_{i=k+1}^{M} \lambda_i$, and α is given by

$$\alpha = \left[\arg\max_k \frac{(\lambda_k - \mu_k)}{\mu_k} \right]^{-1}. \tag{9}$$

In [12], RADOI outperforms the Gerschgoerin disk estimator (GDE) criterion [15] in the presence of colored noise, while its performance in the presence of white noise is similar to the GDE criterion.

Other recent approaches based on expressions with eigenvectors are ESTimation ERror (ESTER) [7] and Subspace-based Automatic Model Order Selection (SAMOS) [8]. Note that although for many scenarios the Probability of Detection (PoD) of SAMOS is superior to the PoD of ESTER, in SAMOS there is the limitation that $1 \leq \hat{d} < \min\left(\left\lfloor \frac{M-1}{2} \right\rfloor, N\right)$, while $\hat{d}$ in ESTER can assume values such that $1 \leq \hat{d} < \min(M, N)$. $\lfloor\rfloor$ denotes the floor operator. Both ESTER and SAMOS are based on the shift invariance equations, which constrain the type of data used, in contrast to the other techniques in this article.

The ESTER's residual error for the model order estimation is given as follows:

$$\begin{aligned} \hat{d} &= \arg\max_k \frac{1}{E_{\mathrm{E}}(k)^2} \quad \text{where} \\ E_{\mathrm{E}}(k)^2 &= \| \boldsymbol{J}_1 \cdot \boldsymbol{U}_k \cdot \hat{\boldsymbol{\Psi}} - \boldsymbol{J}_2 \cdot \boldsymbol{U}_k \|_2^2 \quad \text{where} \\ \hat{\boldsymbol{\Psi}} &= (\boldsymbol{J}_2 \cdot \boldsymbol{U}_k)^+ \cdot \boldsymbol{J}_1 \cdot \boldsymbol{U}_k \end{aligned} \tag{10}$$

where $\boldsymbol{J}_1 \in \mathbb{R}^{M-1 \times M}$ and $\boldsymbol{J}_2 \in \mathbb{R}^{M-1 \times M}$ are the selection matrices for the first $M-1$ sensors and for the last $M-1$ sensors respectively, and $\boldsymbol{U}_k$ represents the k first left singular eigenvectors of $\boldsymbol{X}$.

In SAMOS, the matrix $\boldsymbol{U}_k^{tb} = [\boldsymbol{J}_2 \boldsymbol{U}_k \;\; \boldsymbol{J}_1 \boldsymbol{U}_k] \in \mathbb{C}^{M-1 \times 2 \cdot d}$ is defined. Without noise $\boldsymbol{U}_k^{tb}$ is a rank-d matrix, and the last d singular values λ_{d+1}^{tb},...,λ_{2d}^{tb} are equal to zero. Therefore, the model order selection expression for SAMOS is given by:

$$\begin{aligned} \hat{d} &= \arg\max_k \frac{1}{E_{\mathrm{S}}(k)} \quad \text{where} \\ E_{\mathrm{S}}(k) &= \frac{1}{k} \sum_{i=k+1}^{2k} \lambda_i^{tb} \quad \text{where} \end{aligned} \tag{11}$$

where λ_i^{tb} denotes the i-th singular value of $\boldsymbol{U}_k^{tb}$.

5. EXTENSION OF THE MODIFIED EXPONENTIAL FITTING TEST

AIC and MDL often fail when the number of independent temporal snapshots N is small, in contrast to EDC and mainly NEMO, whose PoD is very high in such a case. The Modified Exponential Fitting Test (M-EFT) [3], an improved version of the Exponential Fitting Test (EFT) [1, 2], has also a very high PoD for such a scenario. EFT is based on the observation that, in the noise-only case, the profile of the ordered eigenvalues can be well approximated by a decaying exponential.

Let λ_i be the i-th eigenvalue of the sample covariance matrix in (3). The exponential model may be expressed as

$$\mathrm{E}\{\lambda_i\} = \mathrm{E}\{\lambda_1\} \cdot q(\alpha, \beta)^{i-1}, \tag{12}$$

where $\mathrm{E}\{\cdot\}$ is the expectation operator and we assume that the eigenvalues are sorted so that λ_1 is the largest. The term $q(\alpha, \beta)$ for the M-EFT is given by

$$q(\alpha, \beta) = \exp\left\{ -\sqrt{\frac{30}{\alpha^2 + 2} - \sqrt{\frac{900}{(\alpha^2 + 2)^2} - \frac{720\alpha}{\beta(\alpha^4 + \alpha^2 - 2)}}} \right\}, \tag{13}$$

so that $0 < q(\alpha, \beta) < 1$ and where $\alpha = \min\{M, N\}$ and $\beta = \max\{M, N\}$.

Three fundamental equations are necessary for the derivation of M-EFT. The first one is the assumption of the exponential profile approximation in (12). The second is the sum of the expectation of the eigenvalues in

$$\sum_{i=1}^{\alpha} \mathrm{E}\{\lambda_i\} = M \cdot \sigma_n^2. \tag{14}$$

The last fundamental equation is the expectation of the square of the eigenvalues in

$$\sum_{i=1}^{\alpha} \mathrm{E}\{\lambda_i^2\} = \frac{M}{N} \cdot (M + N + \gamma) \cdot \sigma_n^4. \tag{15}$$

In case of real-valued noise, we set $\gamma = 1$, otherwise $\gamma = 0$. Basically (15) is modified here in contrast to [1, 2, 3] due to the fact that $\mathrm{E}\{n_{m,n} \cdot n_{m,n}^* \cdot n_{m,n} \cdot n_{m,n}^*\} = 3$ for real-valued noise, and is equal to 2 for complex-valued noise.

Therefore, in contrast to (13), we obtain q from the following equation

$$\begin{aligned} &\Big(C_1(M,N) - 1\Big) \cdot q^{\alpha+1} + \Big(C_1(M,N) + 1\Big) \cdot q^{\alpha} \\ &-\Big(C_1(M,N) + 1\Big) \cdot q + 1 - C_1(M,N) = 0, \end{aligned} \tag{16}$$

where $C_1(M,N) = \frac{M+N+\gamma}{M \cdot N}$. To find a closed-form expression for the rate q in (13), equation (16) was solved in [3] using an approximation. Alternatively, in this variation of M-EFT, we can solve (16) using numerical methods. Moreover, we extend (16) also to the real-valued case, which can be accomplished by replacing $C_1(M,N) = \frac{M+N}{M \cdot N}$, as used in [3] by the more general expression $C_1(M,N) = \frac{M+N+\gamma}{M \cdot N}$.

Similarly to the EFT, in our proposed variation of the M-EFT, it is assumed that the smallest eigenvalue is a noise eigenvalue, and then the prediction of the noise eigenvalues is performed by using the set of the previous eigenvalues. Thereby, it is only possible to obtain $\alpha - 1$ predicted eigenvalues.

The decision if a certain eigenvalue is a noise eigenvalue is performed by computing its relative distance with respect to the predicted eigenvalue. For this relative distance a threshold is defined as function of the probability of false alarm.

Finally the estimation of the model order stops when the relative distance between the eigenvalue and its prediction is outside the threshold region, otherwise, the next larger eigenvalue is tested.

6. SIMULATION RESULTS

In this section we present simulation results demonstrating the performance of the proposed methods considering the data model from (2). Following the CFAR (Constant False-Alarm Rate) approach, the probability of false alarm is set to a constant for all signal to noise ratios. For simplicity, we set $P_{\text{fa}} = 10^{-4}$ for all threshold values computed for M-EFT and the proposed variation. It is therefore instructive to compare the probability of correct detection, i.e., $\Pr\left(\hat{d} = d\right)$ versus the SNR. We also assume that the noise samples are zero mean circularly symmetric complex Gaussian distributed and mutually independent with variance equal to σ_n^2. The spatial frequencies μ_i are drawn from a uniform distribution in $[-\pi, \pi]$. The source symbols are zero mean i.i.d. circularly symmetric complex Gaussian distributed with power equal to σ_s^2 for all the sources. The SNR at the receiver can then be defined as

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right). \tag{17}$$

In Figure 1, we consider real-valued Gaussian data and noise, and as expected the extension of the M-EFT, represented in the legend as M-EFT II, outperforms M-EFT.

Several recent model order selection techniques are compared in Figure 2, and particularly in this scenario we are interested in the case of having a small number of both sensors and snapshots. We can see that the two versions of M-EFT and the traditional EFT outperform all the other techniques. The 1-D AIC has a better performance than AIC, and also than the other techniques

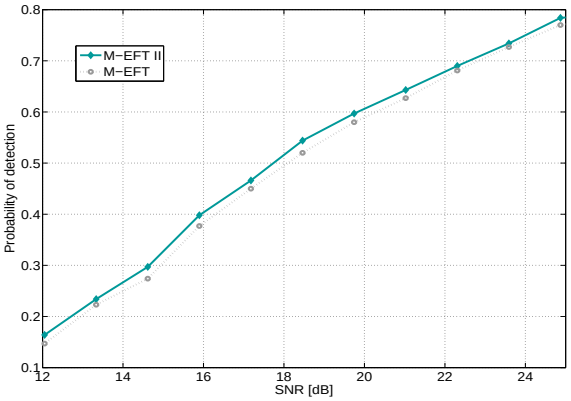

Fig. 1. Comparing the M-EFT II and M-EFT for the case of real noise. Probability of detection vs. SNR for an array of size $M_1 = 5$. The number of snapshots N is set to 6 and the number of sources $d = 3$.

in the literature. The same behavior is observed for the 1-D MDL and MDL.

In Figure 3, we increase the number of sensors to 100, and a very significant difference of performance between the M-EFT techniques and all the other techniques is observed. For example, NEMO is 5 dB far away from the M-EFT techniques. Furthermore, 1-D AIC and 1-D MDL have a better performance than AIC and MDL, while no improvement is observed for 1-D EDC compared to the EDC.

7. COMPARISON OF STATE-OF-ART SCHEMES

After extensive simulation campaigns, we have obtained the following results for different one-dimensional scenarios:

1.1) If $M > N$, the two versions of M-EFT outperform all the other model order estimation techniques. As an exception for this rule, the scenario described in rule 1.3) is shown;

1.2) If $M \leq N$ and considering a high number of samples, N, many solutions can be applied, e.g., M-EFT and its variation proposed here, the traditional EFT, AIC, and MDL. However, since threshold coefficients are not required for AIC and MDL, they turn out to be the best option.

1.3) If $N \leq 8$ and $M \leq 10$, or even smaller, then the traditional EFT and the NEMO slightly outperforms the two versions of M-EFT. Otherwise, the versions of M-EFT have a better PoD.

In addition, for real valued data, the proposed variation of M-EFT should be used instead of M-EFT and for complex value data, both can be applied.

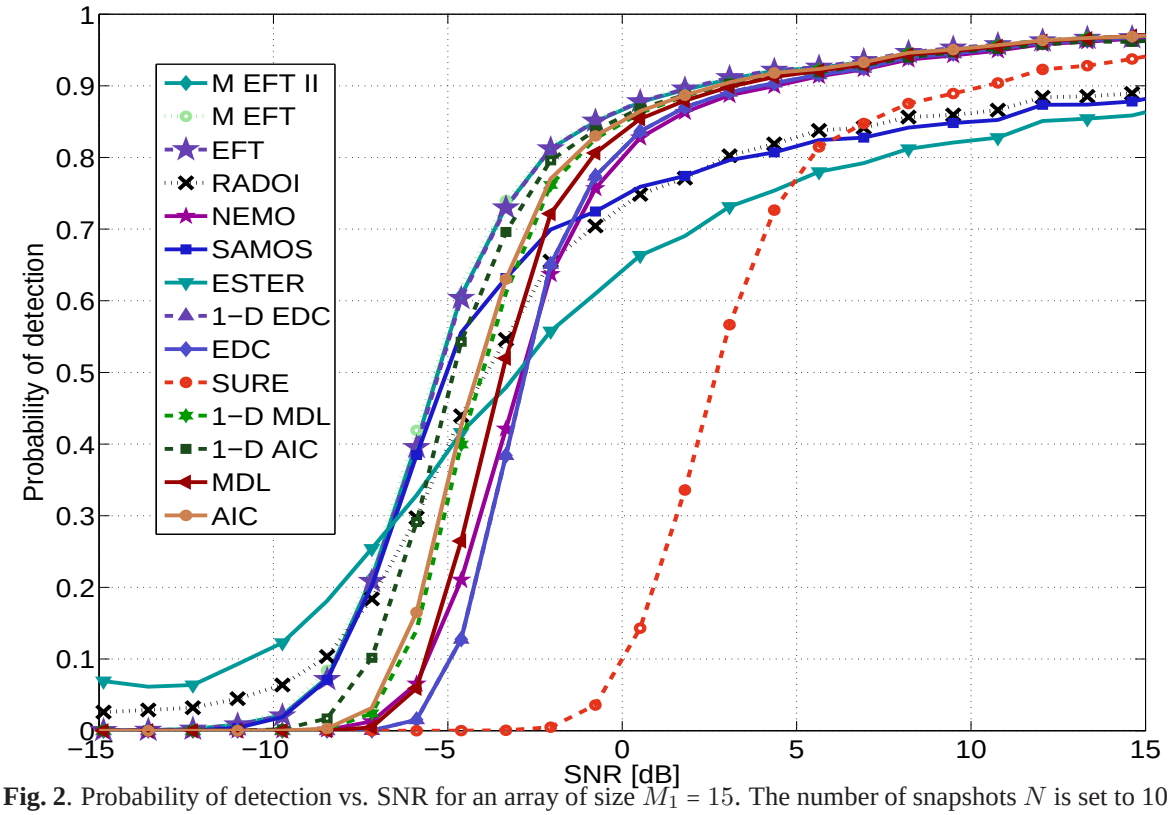

Fig. 2. Probability of detection vs. SNR for an array of size $M_1 = 15$. The number of snapshots N is set to 10 and the number of sources $d = 3$.

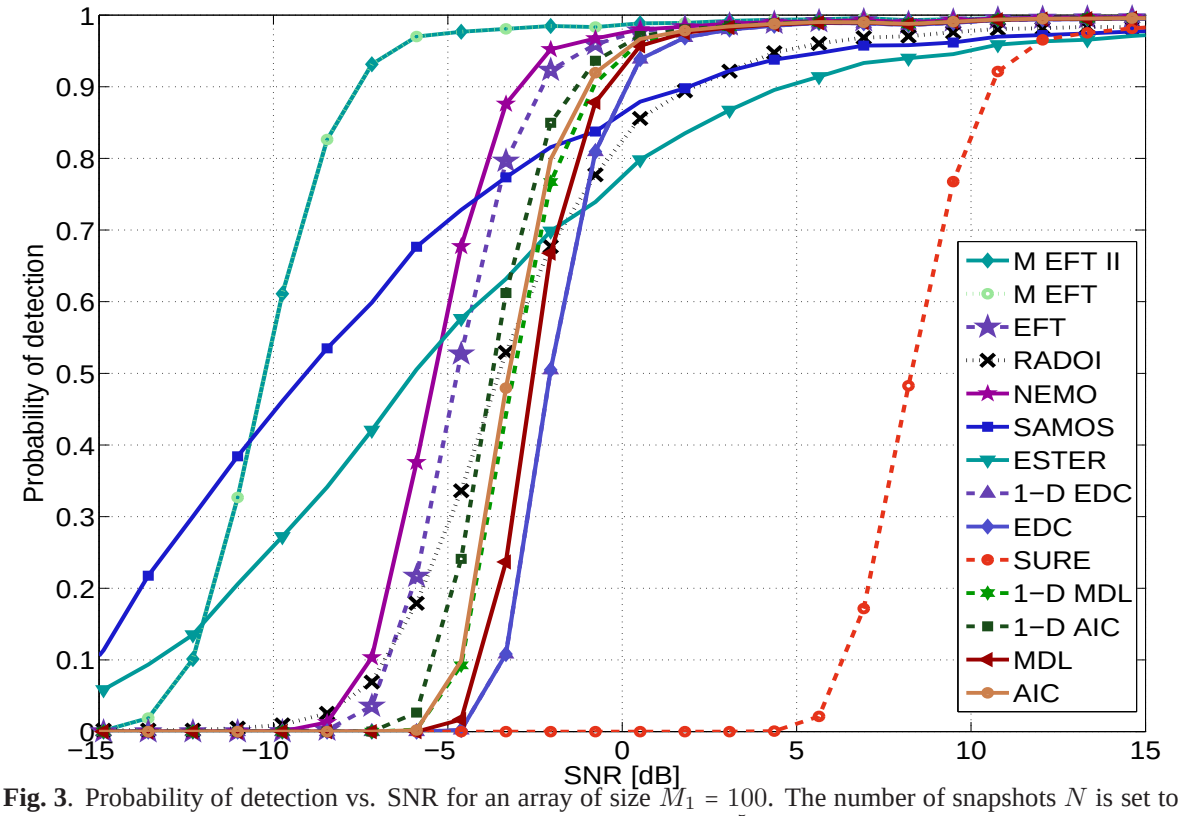

Fig. 3. Probability of detection vs. SNR for an array of size $M_1 = 100$. The number of snapshots N is set to 10 and the number of sources $d = 3$. In this scenario the P_{fa} is set to 10^{-5}.

8. CONCLUSIONS

In this article, we present general rules for selecting the best model order selection techniques for different scenarios. We propose also a variation of M-EFT, whose PoD is higher for real valued data if compared to the M-EFT.

9. REFERENCES

[1] 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)*, May 1996, vol. 5, pp. 2463 – 2466.

[2] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, "Model order selection for short data: An exponential fitting test (EFT)," *EURASIP Journal on Applied Signal Processing*, 2007, Special Issue on Advances in Subspace-based Techniques for Signal Processing and Communications.

[3] J. P. C. L. da Costa, M. Haardt, F. Römer, and G. Del Galdo, "Enhanced model order estimation using higher-order arrays," in *Proc. 40th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2007.

[4] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. on Automatic Control*, vol. AC-19, pp. 716–723, 1974.

[5] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-33, pp. 387–392, 1985.

[6] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Trans. on Information Theory*, vol. 44, pp. 2743–2760, 1998.

[7] R. Badeau, B. David, and G. Richard, "Selecting the modeling order for the esprit high resolution method: an alternative approach," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*, Montreal, Canada, May 2004.

[8] J.-M. Papy, L. De Lathauwer, and S. Van Huffel, "A shift invariance-based order-selection technique for exponential data modelling," *IEEE Signal Processing Letters*, vol. 14, pp. 473–476, July 2007.

[9] P.R. Krishnaiah L.C. Zhao and Z.D.Bai, "On detection of the number of signals in presence of white noise," *Journal of Multivariate Analysis*, vol. 20, pp. 1–25, 1986.

[10] Raj R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *IEEE Transactions of Signal Processing*, vol. 56, pp. 2625–2638, July 2008.

[11] M.O. Ulfarsson and V. Solo, "Rank selection in noisy PCA with SURE and random matrix theory," in *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, Las Vegas, USA, Apr. 2008.

[12] E. Radoi and A. Quinquis, "A new method for estimating the number of harmonic components in noise with application in high resolution radar," *EURASIP Journal on Applied Signal Processing*, pp. 1177–1188, 2004.

[13] T. Minka, "Automatic choice of dimensionality for pca," *Advances in Neural Information Processing Systems*, pp. 598–604, 2000.

[14] R. Kass and A. Raftery, "Bayes factors," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.

[15] 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.