ON THE USE OF ORDER SELECTION RULES FOR ACCURATE PARAMETER ESTIMATION IN THRESHOLD REGION

Kefei Liu[1], *João Paulo C. L. da Costa*[2], *H.C. So*[1],
Florian Roemer[3], *Lei Huang*[4], *and Rafael Timóteo de Sousa Jr.*[2]

[1]Department of Electronic Engineering, City University of Hong Kong, Hong Kong
[2]Department of Electrical Engineering, University of Brasília, Brasília, Brazil
[3]Digital Broadcasting Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany
[4]Department of Electronic and Information Engineering,
Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China
E-mail: kefeilau@gmail.com, jpdacosta@unb.br, hcso@ee.cityu.edu.hk,
Florian.Roemer@ieee.org, lhuang@hitsz.edu.cn, desousa1961@gmail.com

ABSTRACT

Finding the number of signals is crucial to parametric direction-of-arrival (DOA) estimation methods such as MUSIC and ESPRIT. In challenging scenarios such as low signal-to-noise ratio (SNR) and/or presence of closely-spaced sources, only part of the parameters can be accurately estimated while others cannot. The number of former estimates is termed as the effective model order (EMO). We first propose a procedure to determine the EMO via Monte Carlo simulation. Ideally an order selection rule should return a source number estimate equal to EMO, since using an overestimated signal number larger than the EMO in a parameter estimator introduces inaccurate parameter estimates, which is a waste of resources in some applications, while using an underestimate renders some strong signals being treated as noise, which causes an accuracy loss in their parameter estimates. We propose to combine an under-enumerator with an over-enumerator for accurate parameter estimation in the threshold region. Simulations results using the combination of the Baysian information criterion with Akaike information criterion in ESPRIT show that our proposal retains the benefit of the under-enumerators with only accurate estimates while remarkably improves the estimation accuracy.

***Index Terms*—** Order selection, parameter estimation, threshold region, effective model order, joint detection and estimation, array processing

1. INTRODUCTION

Parametric methods for direction-of-arrival (DOA) estimation, such as maximum-likelihood estimator (MLE), estimation of signal parameters via rotational invariance techniques (ESPRIT) [1] and multiple signal classification (MUSIC) [2], rely on the knowledge of the number of signals to properly work. In challenging scenarios such as low signal-to-noise ratio (SNR) and/or presence of closely-spaced sources, using the true signal number for DOA estimation leads to the threshold effect which is characterized by a mean square estimation error (MSE) significantly larger than the Cramér-Rao bound. In general, the sources have distinct powers and/or unequally spaced angles. Therefore, the estimation errors among sources are unevenly distributed: some estimates are centered about the true parameters and carry small or local errors [3], whereas other estimates carry large or global errors [3]. The number of former estimates is referred to as the effective model order (EMO).

In this work, we investigate the optimal choice of the number of signals in the threshold region in the context of joint signal number and parameter estimation. In the threshold region, existing order selection rules tend to underestimate the correct number of signals. To handle this, an empirical detection method biased toward over-enumeration is designed in [4] to cater to radar imagery applications, where it is preferable to overestimate the number of harmonic components than underestimation. However using an overestimated signal number in a parameter estimator introduces useless inaccurate parameter estimates.

We therefore focus on the complementary application scenarios where under-enumeration is preferred. First we explicitly define the EMO and determine it via the Monte Carlo simulation. In applications where inaccurate estimates entail high price and are undesirable, overestimation of the EMO should be avoided, because using an overestimated signal number larger than the EMO in a parameter estimator returns a mixture of both accurate and inaccurate estimates. Therefore, the order selection rules that are prone to underestimating the number of signals, referred to as under-enumerators for short, such as the Baysian information criterion (BIC) [5,6], heavily-penalized efficient detection criteria (EDC) [7], and estimation error (ESTER) [8], are good choices since empirically they rarely overestimate the EMO in the threshold region.

However, the under-enumerators may underestimate the EMO to various extents and with various probabilities. And using an underestimated signal number less than EMO for parameter estimation renders some strong signals being treated as noise, which harms the estimation of other strong signals and results in an accuracy loss in the parameter estimates. We propose to combine the under-enumerators with the order selection rules that tend to overestimate the number of signals, referred to as over-enumerators, for parameter estimation in the threshold region. Such a scheme retains the benefit of the under-enumerators with only accurate estimates while remarkably improves the estimation accuracy.

2. DATA MODEL AND PROBLEM STATEMENT

Consider a scenario where d far-field narrowband emitting sources are impinging on a uniform linear array (ULA) of M sensors. The complex baseband output of the receive antennas is expressed as

$$\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S} + \boldsymbol{Z}, \tag{1}$$

where $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_d)] \in \mathbb{C}^{M\times d}$ is the array steering matrix, with $\boldsymbol{a}(\theta_i) = [1, \ldots, \; e^{j(M-1)2\pi r \sin(\theta_i)/\lambda}]^{\mathrm{T}}$, $i = 1, \ldots, d$, being the array steering vector. Here, $\theta_i \in [-\pi/2, \pi/2]$ is the DOA of the i-th source, measured relative to the array normal direction, r is the inter-element spacing of the receiving antenna array, λ denotes the carrier wavelength, and $^{\mathrm{T}}$ stands for the transpose. For notational convenience, we denote $\mu_i \triangleq 2\pi r \sin(\theta_i)/\lambda$ and let $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_d]$ be the parameter vector that collects all spatial frequencies. The $\boldsymbol{S} = [\boldsymbol{s}_1^{\mathrm{T}}, \ldots, \boldsymbol{s}_d^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{d\times N}$ contains the samples of all sources, with N being the number of snapshots, and $\boldsymbol{Z} \in \mathbb{C}^{M\times N}$ is the noise matrix collecting uncorrelated ZMCSCG samples with variance of σ_z^2.

Hereafter, we assume that (i) $r = \lambda/2$ and hence $\mu_i = \pi \sin(\theta_i) \in [-\pi, \pi]$; (ii) the i-th source samples $\boldsymbol{s}_i$, $i = 1, \ldots, d$, are generated from an independent ZMCSCG random process with variance of $\sigma_{s_i}^2$. The noise is assumed to be uncorrelated with the signal. The SNR is defined as $\mathrm{SNR} = \frac{1}{d}\sum_{i=1}^{d} \sigma_{s_i}^2/\sigma_z^2$.

Given the noisy measurement $\boldsymbol{X}$, we desire to estimate the spacial frequencies $\mu_1, \ldots, \mu_d$.

2.1. Review of State-of-the-art Order Selection Rules

The most commonly used order selection rules are the information theoretic criterion (ITC) approaches which include the Akaike information criterion (AIC) [6, 9], BIC [5, 6], and EDC [7]. In ITCs, the number of signals is determined by minimizing the penalized negative log-likelihood function. The AIC, BIC and EDC as members of the family of ITCs, have a common form of cost function, namely [6, 7],

$$\mathrm{ITC}(k) = -2 \log\left(p_k\left(\boldsymbol{X}, \hat{\boldsymbol{\theta}}^{(k)}\right)\right) + C(N) \cdot \nu \tag{2}$$

but with different penalty coefficients for penalizing overfitting of the model:

$$\mathrm{AIC}: \quad C(N) = 2 \tag{3}$$

$$\mathrm{BIC}: \quad C(N) = \log(N) \tag{4}$$

$$\mathrm{EDC}: \quad C(N) = \sqrt{N \cdot \log(\log(N))} \tag{5}$$

where k is a candidate value for the estimated number of signals, $p_k(\boldsymbol{X}, \hat{\boldsymbol{\theta}}^{(k)})$ is the likelihood function with $\hat{\boldsymbol{\theta}}^{(k)}$ being the maximum likelihood estimate of the parameter vector of the k-th model, namely, $\boldsymbol{\theta}^{(k)}$, and ν is the number of free parameters in $\boldsymbol{\theta}^{(k)}$. In the context of sensor array processing, the likelihood function in (2) is expressed in terms of ratio of the arithmetic and geometric means of the sample eigenvalues [10]. We see that in AIC a lighter penalty than BIC is imposed, whereas in EDC, a heavier penalty is employed. Therefore the AIC is more inclined to overestimate the number of signals, whereas the latter two are apt to underestimation.

The ITCs are derived based on large-sample asymptotics, and perform well particularly when the number of samples N is much larger than that of sensors M. For order selection rules that are applicable to small sample scenarios, the reader is referred to [11–13]. In addition, some detection methods, e.g., the subspace method ESTER [8] and the empirically derived detection method termed here as discriminant function for eigenvalue classification (DFEC) [4], are applicable in both large and small sample scenarios.

2.2. Integration of Order Selection with Parameter Estimation

Since the ESPRIT is a closed-form parameter estimator that is accurate and computationally efficient, it is used for parameter estimation throughout the paper. Note the proposed approach can be generalized to other parametric estimation methods such as MLE and MUSIC in a straightforward way.

The ESPRIT algorithm uses the signal subspace to estimate the spatial frequencies. The first step is to compute the eigenvalue decomposition of the sample covariance matrix

$$\hat{\boldsymbol{R}}_{\mathrm{xx}} = \frac{1}{N}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}} \in \mathbb{C}^{M\times M}, \tag{6}$$

where $^{\mathrm{H}}$ represents the Hermitian transpose. The $\hat{d}$ eigenvectors associated to the $\hat{d}$ largest eigenvalues, where $\hat{d}$ denotes an signal number estimated by an order selection rule, are assumed to form the signal subspace $\boldsymbol{U}_s$.

The shift invariance equation then takes the following form:

$$\boldsymbol{J}_1\boldsymbol{U}_s\boldsymbol{\Psi} \quad \approx \quad \boldsymbol{J}_2\boldsymbol{U}_s, \tag{7}$$

where $\boldsymbol{\Phi}_{\hat{d}} \in \mathbb{C}^{\hat{d}\times\hat{d}}$ is the unknown matrix to be solved, and $\boldsymbol{J}_1 \in \mathbb{R}^{(M-1)\times M}$ (resp. $\boldsymbol{J}_2 \in \mathbb{R}^{(M-1)\times M}$) is the selection matrix formed by the first (resp. last) $(M-1)$ rows of an $M\times M$ identity matrix. These sets of equations are overdetermined and can be solved by the least squares method.

From $\boldsymbol{\Psi}$, the spatial frequencies are estimated as

$$\hat{\mu}_i = \arg(\lambda_i) \qquad i = 1, \ldots, \hat{d}, \tag{8}$$

where λ_i denotes the i-th eigenvalue of $\boldsymbol{\Psi}$.

In order to evaluate the performance and calculate the root MSE (RMSE), we use the greedy algorithm to pair the estimated and true spatial frequencies. Denote $\tilde{d} = \min(d, \hat{d})$, after pairing up we have $(\hat{\mu}_1, \mu_{i_1}), \ldots, (\hat{\mu}_{\tilde{d}}, \mu_{i_{\tilde{d}}})$, where $|\hat{\mu}_1 - \mu_{i_1}| \le \cdots \le |\hat{\mu}_{\tilde{d}} - \mu_{i_{\tilde{d}}}|$. The MSE of the parameter estimates is defined as

$$\mathrm{RMSE}(\hat{d}) = \sqrt{\frac{1}{\tilde{d}}\sum_{k=1}^{\tilde{d}} |\hat{\mu}_k - \mu_{i_k}|^2}. \tag{9}$$

2.3. Problem with Use of Single Order Selection Rule for Parameter Estimation in Threshold Region

In the threshold region, only part of the signal parameters can be accurately estimated. The traditionally well-performing order selection rules that target for identifying the correct number of signals in the asymptotic region may not be a good choice in the threshold region.

To illustrate this, we consider a scenario where $d = 5$ sources with equal powers but different angular separations impinge on a ULA of $M = 10$ elements each collecting $N = 100$ snapshots, with $\boldsymbol{\mu} = [47, 125, -11, 38, 135]^\circ$. Random Gaussian white noise with a fixed SNR of = 5 dB is added. The parameter estimates and the estimation errors for a single but representative noisy realization are displayed in Table 1, where the second to fifth rows respectively correspond to the results obtained when the correct number of signals, namely, $d = 5$, the number of signals estimated by the AIC, BIC and EDC, denoted as $\hat{d}_{\mathrm{AIC}}$, $\hat{d}_{\mathrm{BIC}}$ and $\hat{d}_{\mathrm{EDC}}$, are respectively used in ESPRIT for DOA estimation. The numbers in the parentheses denote the estimation errors of the individual estimates.

From the second row, we see that among the five estimates, only three, namely, the ones corresponding to $\mu_3 = -11^\circ$, $\mu_4 = 38^\circ$ and $\mu_5 = 135^\circ$ are accurate (effective) whereas other the two, namely, the ones associated to $\mu_1 = 47^\circ$ and $\mu_2 = 125^\circ$ carry large estimation errors. This is because two pairs of sources, namely, $(\mu_1, \mu_4) = (47, 38)^\circ$ and $(\mu_2, \mu_5) = (125, 135)^\circ$ are closely spaced, and for a high noise level, the separation between each pair of closely-spaced sources falls below the resolution of ESPRIT and becomes invisible and they merge to one. The effective number of signals, namely, the EMO, is hence equal to 3.

In the third row of Table 1, the AIC overestimates the EMO by 1, which results in a mixture of 3 accurate parameter estimates and 1 inaccurate estimate. This is undesirable in applications where inaccurate estimates entail extra costs.

On the other hand, the BIC correctly estimates the EMO and the EDC underestimates the EMO by one, in both cases the inaccurate estimates have been completely excluded. However, in the last column of Table 1 we see that the RMSE of the resultant estimates by passing $\hat{d}_{\mathrm{EDC}} = 2$ to ESPRIT is 3.91°, which is clearly larger than that of the two most accurate estimates obtained by passing $d = 5$ to ESPRIT, which is only 2.08°. We call this as the accuracy loss. This is because using $\hat{d}_{\mathrm{EDC}} = 2$ in ESPRIT renders one effective (strong) signal being treated as noise, which harms the estimation of other two strong signals.

Note that for BIC, under-enumeration of the EMO also occurs although less commonly than EDC. Therefore, it suffers the accuracy loss as well. In Section 3, we propose a simple method to reduce the accuracy loss.

3. PROPOSED APPROACH

Given the number of sensors and samples, number of sources d, signal powers and DOAs of d sources, and noise level, the EMO can be determined by the Monte Carlo simulation. Denote the d frequency estimates obtained by passing the correct number of signals d to ESPRIT and ordered in terms of accuracy as $\hat{\mu}_1, \ldots, \hat{\mu}_d$, with $|\hat{\mu}_1 - \mu_{i_1}| \leq \cdots \leq |\hat{\mu}_d - \mu_{i_d}|$, where $(\hat{\mu}_1, \mu_{i_1}), \ldots, (\hat{\mu}_d, \mu_{i_d})$ are the estimated-true frequency pairs. The effective number of signals, denoted by d_{eff}, is determined as

$$d_{\mathrm{eff}} = \max_{\ell \in \{1,\ldots,d\}} \ell \quad \text{subject to} \quad \mathrm{RMSE}(\ell) \leq \eta, \tag{10}$$

with

$$\mathrm{RMSE}(\ell) = \sqrt{\frac{\sum_{n=1}^{N_{\mathrm{T}}} \sum_{k=1}^{\ell} |\hat{\mu}_k^{(n)} - \mu_{i_k}^{(n)}|^2}{N_{\mathrm{T}} \times \ell}}, \tag{11}$$

where N_{T} is the number of trials in the Monte Carlo simulation, $\{\hat{\mu}_k^{(n)}\}_{k=1}^{d}$ denote the d frequency estimates in the n-th trial, $n = 1, \ldots, N_{\mathrm{T}}$, and η is the predefined threshold estimation error. To reduce the statistical error, N_{T} should be sufficiently large, typically of the order of 1000.

A close-to-generic settings of η is the estimation error corresponding to the threshold point. Fig. 1 shows the identified EMO for various SNRs by Monte Carlo simulation. Note that the EMO is equal to the number of sources at a low noise level in the asymptotic region, and gradually decreases as the noise power increases in the threshold region, and finally down to zero at a high noise level in the no information region, also known as permanent state of futility.

As shown in Section 2.3, integration of under-enumerators with ESPRIT suffers from an accuracy loss in the parameter estimates to various extents. To reduce the accuracy loss, we can resort to an over-enumerator. Looking again at Table 1, we see that the MSE of the two most accurate estimates obtained by passing $\hat{d}_{\mathrm{AIC}} = 4$ to ESPRIT improves in accuracy over the ones obtained by passing $\hat{d}_{\mathrm{EDC}} = 2$ to ESPRIT. This is because, different from the latter, in the former only one non-identifiable signal is treated as noise and this does not harm the estimation of strong components. This result is presented in the following conjecture whose proof is left as future work while numerical simulations are used as corroborating evidence.

Conjecture 1 Consider two signal number underestimates $\hat{d}_1 < \hat{d}_2 \leq d$. If $\hat{d}_1 < d_{\mathrm{eff}} \leq \hat{d}_2$, the parameter estimates obtained by passing $\hat{d}_1$ to ESPRIT is less accurate (have larger MSE) than the $\hat{d}_1$ most accurate parameter estimates obtained by passing $\hat{d}_2$ to ESPRIT.

In order to reduce the accuracy loss of the under-enumerator while keeping its advantage with only accurate estimates, we propose to combine it with an over-enumerator in the following steps.

1) Choose a proper over-enumerator to obtain a signal number estimate, denoted as $\hat{d}_{\mathrm{ub}}$. Let $\hat{\boldsymbol{\mu}}_{\mathrm{ub}} = [\hat{\mu}_1, \ldots, \hat{\mu}_{\hat{d}_{\mathrm{ub}}}]$ collect the $\hat{d}_{\mathrm{ub}}$ parameter estimates obtained by passing $\hat{d}_{\mathrm{ub}}$ to ESPRIT. Normally $\hat{d}_{\mathrm{ub}} \geq d_{\mathrm{eff}}$, and when $\hat{d}_{\mathrm{ub}} > d_{\mathrm{eff}}$, $\hat{\boldsymbol{\mu}}_{\mathrm{ub}}$ is a mixture of accurate and inaccurate parameter estimates.
2) Use an under-enumerator to obtain a signal number estimate, denoted as $\hat{d}_{\mathrm{lb}}$.
3) For $\hat{d}_{\mathrm{lb}} > 0$, identify the $\hat{d}_{\mathrm{lb}}$ most accurate components of $\hat{\boldsymbol{\mu}}$ via the minimum reconstruction-error criteria. Specifically, let $\hat{\boldsymbol{\mu}}_{\mathrm{lb}}^{(k)}$ represent a $\hat{d}_{\mathrm{lb}}$-combination of the $\hat{d}_{\mathrm{lb}}$ components in $\hat{\boldsymbol{\mu}}$. For the k-th combination, $k = 1, \ldots, \binom{\hat{d}_{\mathrm{ub}}}{\hat{d}_{\mathrm{lb}}}$, we first reconstruct $\boldsymbol{A}\left(\hat{\boldsymbol{\mu}}_{\mathrm{lb}}^{(k)}\right)$ and then find the estimate of $\boldsymbol{S}$ as $\hat{\boldsymbol{S}}^{(k)} = \left[\boldsymbol{A}\left(\hat{\boldsymbol{\mu}}_{\mathrm{lb}}^{(k)}\right)\right]^{\dagger} \cdot \boldsymbol{X}$. The $\hat{d}_{\mathrm{lb}}$ most accurate estimates are determined as

$$\hat{\boldsymbol{\mu}}_{\mathrm{lb}} = \underset{k=1,\ldots,\binom{\hat{d}_{\mathrm{ub}}}{\hat{d}_{\mathrm{lb}}}}{\arg\min} \left\| \boldsymbol{X} - \boldsymbol{A}\left(\hat{\boldsymbol{\mu}}_{\mathrm{lb}}^{(k)}\right) \cdot \hat{\boldsymbol{S}}^{(k)} \right\|_{\mathrm{F}}^{2}, \tag{12}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix.

In general $\hat{d}_{\mathrm{ub}} \leq d$, and hence the number of combinations $\binom{\hat{d}_{\mathrm{ub}}}{\hat{d}_{\mathrm{lb}}}$ increases at a rate no more than $\left(\hat{d}_{\mathrm{ub}} - \hat{d}_{\mathrm{lb}}\right)$-order polynomially with d. Since typically $\left(\hat{d}_{\mathrm{ub}} - \hat{d}_{\mathrm{lb}}\right)$ is small, the above algorithm can be implemented at a reasonable computational cost.

3.1. Choice of Under- and Over-enumerators

The choice of a good under-enumerator depends on the accuracy requirement which is application dependent. The higher the estimation accuracy requirement is, the more conservative enumerator should be chosen. Using an enumerator that is more conservative in selecting the number of signals such as EDC leads to a decrease in the number of available estimates, but meanwhile an improvement in the estimation accuracy is expected.

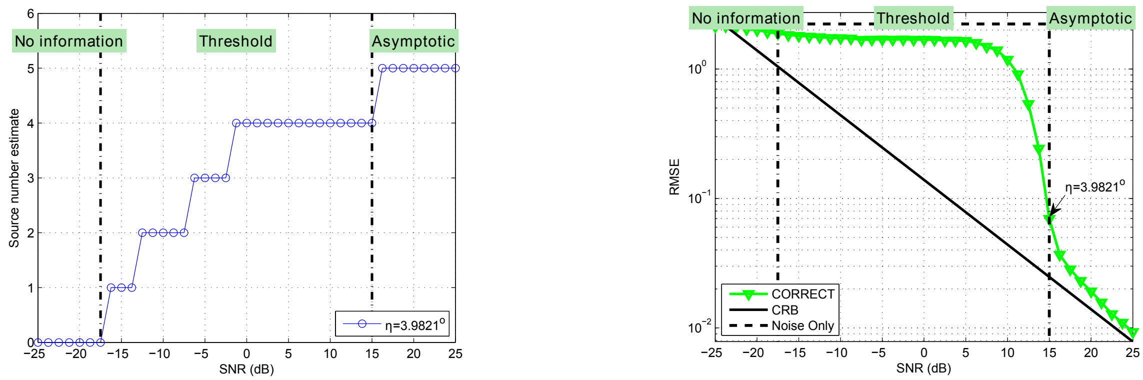
On the other hand, according to Conjecture 1, a good over-enumerator should be able to return a signal number estimate that satisfies $d_{\mathrm{eff}} \leq \hat{d}_{\mathrm{ub}} \leq d$. However, this condition can be relaxed to $d_{\mathrm{eff}} \leq \hat{d}_{\mathrm{ub}} \leq d + \Delta$, where Δ is a small positive integer, e.g., 1 or 2. It is because simulations show that the accuracy of the final $\hat{d}_{\mathrm{lb}}$ estimates obtained in the proposed algorithm is much less sensitive to overestimation of the EMO than to underestimation of it by the chosen over-enumerator, even if $\hat{d}_{\mathrm{ub}}$ slightly exceeds d. In this sense, the AIC is a good option for the over-enumerator.

4. SIMULATION RESULTS

The BIC and EDC are used as under-enumerators, while for the over-enumerator two schemes are employed: 1) AIC; 2) Combination

Table 1. Frequency estimation error in degrees (°) in ESPRIT when the correct number of signals, the number of signals estimated by AIC and BIC are used. M = 10, N = 100, d = 5, $\boldsymbol{\mu}$ = $[47, 125, -11, 38, 135]^\circ$, and equal-power sources with SNR = 5 dB.

μ_i	47	125	-11	38	135	RMSE	Best $\hat{d}_{\mathrm{BIC}}$	Best $\hat{d}_{\mathrm{EDC}}$
$\hat{\mu}_i$ (d = 5)	-74.5 (-121.5)	94.1 (-30.9)	-12.8 (-1.8)	40.3 (2.3)	130.4 (-4.6)	56.12	3.15	2.08
$\hat{\mu}_i$ ($\hat{d}_{\mathrm{AIC}}$ = 4)	×	101.8 (-23.2)	-11.6 (-0.6)	41.3 (3.3)	130.1 (-4.9)	11.97	3.43	2.35
$\hat{\mu}_i$ ($\hat{d}_{\mathrm{BIC}}$ = 3)	×	×	-11.6 (-0.6)	42.1 (4.1)	130.8 (-4.2)	3.42	3.42	2.96
$\hat{\mu}_i$ ($\hat{d}_{\mathrm{EDC}}$ = 2)	×	×	×	34.5 (-3.5)	130.7 (-4.3)	3.91	×	3.91



(a) Effective model order

(b) RMSE of frequency estimates by ESPRIT using $\hat{d} = d$.

Fig. 1. Determination of effective model order. M = 10, N = 50, d = 5. $\boldsymbol{\theta} = [17, -29, 56, 22, 55]^\circ$ and $\eta = 2.57^\circ$.

of AIC and DFEC, namely, choose the estimate of AIC or DFEC, whichever is larger. The DFEC is an empirical order selection rule that has been shown to be good at detecting weak signals [4].

We consider a scenario where the sources have equal powers and closed-spaced sources are present, with $\boldsymbol{\theta} = [17, -29, 56, 22, 55]^\circ$. The numbers of sensors and snapshots are M = 10, N = 100. For widely-spaced sources with distinct powers, similar results are obtained. Therefore, they are not included here to save space. The results are plotted for various SNRs, and for each SNR, 50000 independent realizations are conducted.

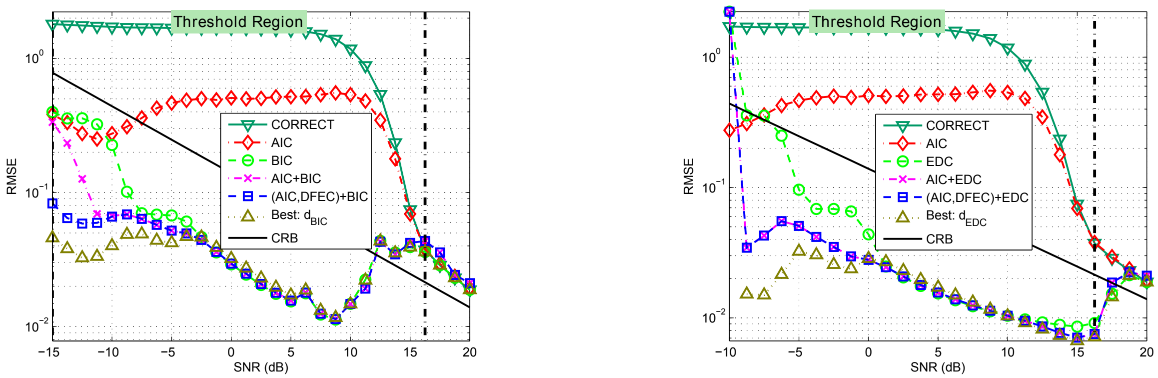
Fig. 2(a) and (b) show the RMSEs of the frequency estimates obtained by ESPRIT, where BIC and EDC are respectively used as the under-enumerator. We see that using the correct signal number for parameter estimation in the threshold region results in large RMSE of the estimated parameters that is far above the CRB, which implies that inaccurate estimates are present.

In Fig. 3, the EMO and the mean values of the signal number estimate by various order selection rules are shown and compared. We see that the AIC underestimates the correct number of signals but overestimates the EMO. Correspondingly, in Fig. 2, the RMSE of the parameter estimates obtained using the estimated signal number by AIC, denoted as $\hat{d}_{\mathrm{AIC}}$, in ESPRIT drops down because part of inaccurate estimates have been excluded from its estimates, but may be still too large to satisfy the accuracy requirement in applications.

On the other hand, the BIC underestimates the EMO, and therefore, the RMSE drops down below the CRB since inaccurate estimates have been excluded. Nevertheless, in Fig. 2(a) we see that, the RMSE of the obtained estimates by passing $\hat{d}_{\mathrm{BIC}}$ to ESPRIT is not as small as that of the "BEST: $\hat{d}_{\mathrm{BIC}}$", which corresponds to the $\hat{d}_{\mathrm{BIC}}$ most accurate components of $\hat{\boldsymbol{\mu}}$, particularly in the lower half of the threshold region. Here $\hat{\boldsymbol{\mu}} = [\hat{\mu}_1, \ldots, \hat{\mu}_d]$ denotes the vector of parameter estimates obtained by passing the correct number of signals d to ESPRIT. As mentioned before, this is because using an underestimated number of signals less than EMO renders strong signals being treated as noise and this harms the estimation of other strong signals which results in an accuracy loss.

By combining the BIC with AIC, the RMSE is significantly reduced, by a factor up to 78%. In addition, by combining the BIC with both AIC and DFEC, the RMSE is further reduced, particularly in the lower half of the threshold region, where the DFEC is more apt to overestimate the EMO than the AIC, as shown in Fig. 3. This is consistent with [4], showing the superiority of DFEC in detecting weak signals.

Similar observations are also obtained when the EDC is used as the under-enumerator. However, note that since the EDC is more conservative in selecting the number of signals than BIC, the RMSE of the $\hat{d}_{\mathrm{EDC}}$ most accurate components of $\hat{\boldsymbol{\mu}}$ is smaller than that of the $\hat{d}_{\mathrm{BIC}}$ most accurate ones, which is seen by comparing the "BEST: $\hat{d}_{\mathrm{BIC}}$" curve in Fig. 2(a) and "BEST: $\hat{d}_{\mathrm{EDC}}$" curve in Fig. 2(b). Nevertheless, the accuracy loss due to underestimation of number of signals is more significant, which is reflected by the enlarged gap between the RMSE of the $\hat{d}_{\mathrm{EDC}}$ estimates obtained by passing $\hat{d}_{\mathrm{EDC}}$ to ESPRIT and that of the "BEST: $\hat{d}_{\mathrm{EDC}}$" curve and almost cancels out all the potential gains made by reducing the number of estimates through the use of the more conservative EDC. By combining the EDC with AIC, the performance gap becomes much smaller, with the RMSE being reduced by up to 90% compared to the 78% for the BIC case. Moreover, the resultant $\hat{d}_{\mathrm{EDC}}$ estimates with improved accuracy are more accurate than the accuracy-improved $\hat{d}_{\mathrm{BIC}}$ estimates in Fig. 2(a).



(a) BIC

(b) EDC

Fig. 2. RMSE of frequency estimates versus SNR for $d = 5$ equal-power sources among which closely-spaced sources are present. $M = 10$, $N = 100$, $d = 5$. $\boldsymbol{\theta} = [17, -29, 56, 22, 55]^{\circ}$.

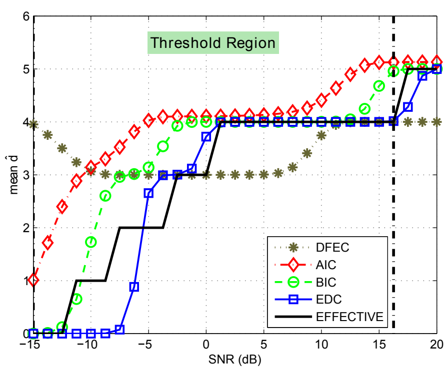


Fig. 3. EMO and mean values of signal number estimate by under- and over-enumerators. The parameter settings are the same as those in Fig. 2.

5. CONCLUSION

In challenging scenarios with low SNRs or the presence of closely-spaced sources, only part of signal parameters can be accurately estimated. The number of such parameter estimates is called the effective model order (EMO). In applications where false alarm and inaccurate estimates entail much higher price than missed detection, the EMO is preferred over the correct model order for parameter estimation, since using the latter in a parameter estimator introduces inaccurate estimates which incurs extra costs. However, using an underestimated signal numbers in a parameter estimator causes some strong signals to be treated as noise and hence a loss in estimation accuracy. We propose to combine the under- and over-enumerators for parameter estimation in the threshold region. Such a scheme enjoys the benefit of the under-enumerators with only accurate estimates while significantly alleviates the accuracy loss.

6. REFERENCES

[1] R. Roy and T. Kailath, "ESPRIT - Estimation of signal parameters via rotational invariance techniques," *IEEE Trans Acoust*, vol. ASSP-37, pp. 984–995, Jul. 1989.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[3] C. D. Richmond, "Capon algorithm mean-squared error threshold SNR prediction and probability of resolution," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2748–2764, Aug. 2005.

[4] E. Radoi and A. Quinquis, "A new method for estimating the number of harmonic components in noise with application in high resolution radar," *EURASIP J. Appl. Signal Process.*, vol. 8, pp. 1177–1188, 2004.

[5] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, pp. 461–464, 1978.

[6] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, 2004.

[7] Z. D. Bai, P. R. Krishnaiah, and L. C. Zhao, "On rates of convergence of efficient detection criteria in signal processing with white noise," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 380–388, 1989.

[8] R. Badeau, B. David, and G. Richard, "Selecting the modeling order for the ESPRIT high resolution method: An alternative approach," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP '04)*, Montreal, Canada, May 2004, vol. 2, pp. 1025–1028.

[9] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[10] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans Acoust*, vol. 33, no. 2, pp. 387–392, 1985.

[11] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, "Model order selection for short data: An exponential fitting test (EFT)," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. Article ID 71953, pp. 54–64, 2007.

[12] R. R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2625–2638, 2008.

[13] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, 2009.