

Article

Simple and Efficient Algorithm for Improving the MDL Estimator of the Number of Sources

Dayan A. Guimarães * and Rausley A. A. de Souza *

National Institute of Telecommunications (Inatel), Av. João de Camargo, 510, Santa Rita do Sapucaí MG 37540-000, Brazil

* Authors to whom correspondence should be addressed; E-Mails: dayan@inatel.br (D.A.G.); rausley@inatel.br (R.A.A.S.); Tel.: +55-353471-9227 (D.A.G.); Fax: +55-353471-9314 (D.A.G.).

External Editor: Markus Sigrist

Received: 28 July 2014; in revised form: 11 September 2014 / Accepted: 8 October 2014 /

Published: 17 October 2014

Abstract: We propose a simple algorithm for improving the MDL (minimum description length) estimator of the number of sources of signals impinging on multiple sensors. The algorithm is based on the norms of vectors whose elements are the normalized and nonlinearly scaled eigenvalues of the received signal covariance matrix and the corresponding normalized indexes. Such norms are used to discriminate the largest eigenvalues from the remaining ones, thus allowing for the estimation of the number of sources. The MDL estimate is used as the input data of the algorithm. Numerical results unveil that the so-called norm-based improved MDL (iMDL) algorithm can achieve performances that are better than those achieved by the MDL estimator alone. Comparisons are also made with the well-known AIC (Akaike information criterion) estimator and with a recently-proposed estimator based on the random matrix theory (RMT). It is shown that our algorithm can also outperform the AIC and the RMT-based estimator in some situations.

Keywords: AIC; estimation of the number of sources; iMDL; MDL; RMT

1. Introduction

The estimation of the number of sources of signals impinging on multiple sensors is a fundamental problem in communications and signal processing. This number is important in itself in a huge number of applications [1–9], e.g., to determine the approximate number of neurons responding to some stimulus in the brain [10], to estimate the number of active muscles during an action for identifying the action and determining pathologies [11], to find the number of chemical elements in a mixture [12], to determine the number of speakers in a room with background noise and reverberation [13] or, yet, in a cognitive radio network, to estimate the number of active radio transmitters in a given area [14] or to find the best spectrum opportunities [15]. In other applications, it is used as the input for subsequent estimations, e.g., for some direction of arrival (DoA) estimation and source separation methods in array processing [16,17], for detecting the number of signals in a radar measurement [18] and cognitive radio applications [19] and for estimating the parameters in Markov models [20]. Some DoA estimators do not demand the prior knowledge of the number of sources and apply, as well, to the estimation of this number [7]. Some source separation algorithms also do not need prior knowledge of the number of sources and can also be used for estimating this number [8,9].

In spite of the multitude of techniques available for solving the problem of estimating the number of sources, most of them adopt information theoretic approaches, such as the minimum description length (MDL) and the Akaike information criterion (AIC) [1–5,13,14,21]. The AIC achieves high detection performance (*i.e.*, high probability of correctly estimating the number of sources) at low signal-to-noise ratio (SNR) regimes, but it is not consistent. The MDL estimator is consistent, but degrades its performance at low SNR regimes when compared with the MDL. Nevertheless, over the years, the MDL has been considered a standard when assessing the performance of algorithms for estimating the number of sources [2].

Recently, the random matrix theory (RMT) approach has been proposed to tackle the estimation of the number of source problem [2,22]. In [2], an RMT-based sequential hypothesis test is proposed, claiming high detection performance at low SNR, similar to the AIC estimator, and near consistency at large sample sizes, similar to the MDL estimator. In [22], an improved RMT-based algorithm is proposed. It is also based on a sequential hypothesis test, but the test statistics were derived so as to consider the interaction between signal and noise eigenvalues in the small sample size situation, which is not considered in the original algorithm of [2]. By combining the original algorithm with the new one, improved performance is achieved for both the low and the high SNR regimes.

In this paper, we propose a new algorithm for improving the performance of the MDL estimator. We call it norm-based improved MDL (iMDL). Our algorithm enjoys low complexity, yet exhibits consistency and outperforms the AIC estimator and the RMT-based estimator of [2] in some situations. The performance of the iMDL is always better than or equal to the performance achieved by the MDL estimator alone.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation related to eigenvalue-based source estimators and describes briefly the AIC-, the MDL- and the RMT-based estimators. The proposed iMDL algorithm is described in Section 3. The simulation setup, numerical results and discussions concerning the influence of the system parameters on performance are shown

in Section 4. Finally, in Section 5, concluding remarks are drawn and opportunities for future work are highlighted.

2. Eigenvalue-Based Estimators

In this paper, we focus on methods that use the eigenvalues of the received signal covariance matrix for estimating the number of sources. We first address the problem formulation and then briefly describe the AIC, the MDL and the RMT-based estimators.

2.1. Problem Formulation

Let there be an array with m sensors (antennas, for example), each one collecting n samples of the received signal from p sources (transmitters). These samples are arranged in a matrix $\mathbf{Y} \in \mathbb{C}^{m \times n}$, and the samples from the p sources are arranged in a matrix $\mathbf{X} \in \mathbb{C}^{p \times n}$. Let $\mathbf{H} \in \mathbb{C}^{m \times p}$ be the channel matrix with elements $\{h_{ij}\}$, $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, p$, representing the channel gains between the j -th source and the i -th sensor. Finally, let $\mathbf{V} \in \mathbb{C}^{m \times n}$ be a matrix of additive Gaussian noise samples, distributed $\mathcal{N}(0, \sigma^2 \mathbf{I}_m)$ and independent of the signal samples. The matrix of received samples is then:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V} \quad (1)$$

In this scenario, we consider only those methods for estimating the number of sources that use the ordered eigenvalues $\{\lambda_1 \geq \lambda_2 \geq \dots \lambda_m\}$ of the received signal population covariance matrix \mathbf{R} , for which the maximum likelihood estimate is the sample covariance matrix:

$$\hat{\mathbf{R}} = \frac{1}{n} \mathbf{Y}\mathbf{Y}^\dagger \quad (2)$$

where † means complex conjugate and transpose. It can be shown [1] that \mathbf{R} can be written as:

$$\mathbf{R} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^\dagger] = \mathbf{H}\mathbf{R}_x\mathbf{H}^\dagger + \sigma^2\mathbf{I} \quad (3)$$

where \mathbf{R}_x is the transmitted signal population covariance matrix, \mathbf{I} is the identity matrix, σ^2 is the noise variance and $\mathbb{E}[\cdot]$ is the expectation operator. If \mathbf{H} is full column rank, the rank of $\mathbf{H}\mathbf{R}_x\mathbf{H}^\dagger$ is p , which means that the $m - p$ smallest eigenvalues of $\mathbf{H}\mathbf{R}_x\mathbf{H}^\dagger$ are equal to zero. Therefore, the $m - p$ smallest eigenvalues of \mathbf{R} are equal to σ^2 . Then, it is possible to estimate the number of sources p from the multiplicity of the smallest eigenvalues of \mathbf{R} . Instead of \mathbf{R} , in practice, its estimate $\hat{\mathbf{R}}$ is computed using a finite number of samples, and the resulting eigenvalues are all different with probability one. In this case, the classification of the eigenvalues in two groups (the largest p and the smallest $m - p$) is not trivial, representing a challenge for the estimation of the number of sources.

The fundamental difference among the eigenvalue-based methods for estimating the number of sources resides in the way that the eigenvalues are classified. In the following two subsections, we describe three of these methods: the AIC, the MDL and the RMT-based algorithm proposed in [2].

2.2. The AIC and the MDL Estimators

The AIC and the MDL estimators have been derived from information theoretic foundations, which renders these estimators the classification as information-theoretic methods [1]. The AIC has been derived from the Kullback–Leibler information concept [23], and the MDL has been derived from a Bayesian approach [24,25]. By departing from different roots, they have ended up in very similar structures; they are also referred to as maximum penalized likelihood methods, for which a compensation or penalization term is subtracted from the likelihood function of the parameters to be estimated. Specifically, the criteria for AIC and MDL are, respectively [1]:

$$\text{AIC}(k) = -2n \ln \frac{\prod_{i=k+1}^m \lambda_i}{\left(\frac{1}{m-k} \sum_{i=k+1}^m \lambda_i\right)^{m-k}} + 2k(2m-k) \quad (4)$$

$$\text{MDL}(k) = -n \ln \frac{\prod_{i=k+1}^m \lambda_i}{\left(\frac{1}{m-k} \sum_{i=k+1}^m \lambda_i\right)^{m-k}} + \frac{1}{2}k(2m-k) \ln n \quad (5)$$

for $k = 0, 1, \dots, m-1$. For both, the estimate \hat{p} of the number of sources is the value of k , which minimizes the criterion. Notice that, apart from a constant, both estimators have a common first term and different penalization terms. These penalization terms act by compensating, in different ways, the overestimation of the model order (the number of sources), which is known to be produced by the first term [5].

2.3. The RMT-Based Estimator

The algorithm proposed in [2] is based on a sequence of hypothesis tests. In each step, the significance of the k -th eigenvalue of the covariance matrix of the signal received by m sensors is tested. For $k = 1, 2, \dots, \min(m, n) - 1$, it is tested if:

$$\lambda_k > \hat{\sigma}^2(k) C_{n,m,k}(\alpha) \quad (6)$$

where $\hat{\sigma}^2(k)$ is the noise variance estimate on step k and $C_{n,m,k}(\alpha)$ is a parameter that depends on the confidence level $\alpha \ll 1$, as determined by the user. The value of α is referred to as the asymptotic false alarm (overestimation) probability.

The computation of $C_{n,m,k}(\alpha)$ involves the calculation of the inverse of a Tracy–Widom distribution, which is a quite difficult process, since this distribution does not have a closed-form expression. Nevertheless, in [26], the authors of [2] provide MATLAB routines capable of inverting numerically the Tracy–Widom distribution, which facilitates the computer implementation of the RMT-based algorithm.

Then, the RMT-based estimate of the number of sources can be written as [2]:

$$\hat{p} = \arg \min_k \{ \lambda_k < \hat{\sigma}^2(k) C_{n,m,k}(\alpha) \} - 1 \quad (7)$$

Notice that this method suffers from the need of computing a threshold that depends on an estimate of the noise variance, but on the other hand, permits control of the overestimation probability, which can be advantageous in some applications. A noise variance estimation process is also suggested in [2].

3. The Proposed iMDL Algorithm

Assume that the ordered eigenvalues of $\hat{\mathbf{R}}$ and the corresponding indexes are normalized, so that both are placed in the interval $[0, 1]$, that is, for $i = 1, \dots, m$,

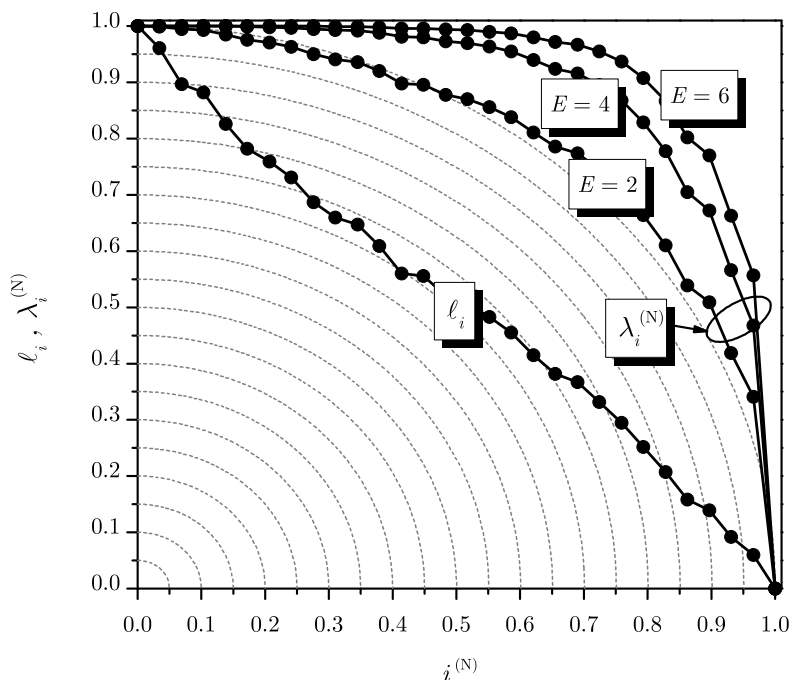
$$\ell_i = \frac{\lambda_i - \lambda_m}{\lambda_1 - \lambda_m} \quad (8a)$$

$$i^{(N)} = \frac{i - 1}{m - 1} \quad (8b)$$

$$\lambda_i^{(N)} = \sqrt{1 - (1 - \ell_i)^E} \quad (8c)$$

where ℓ_i and $i^{(N)}$ are the i -th normalized eigenvalues and indexes, respectively. Notice in Equation (8) that the normalized eigenvalues $\{\ell_i\}$ are further modified to $\{\lambda_i^{(N)}\}$ by a nonlinear operation whose role is explained with the help of an example: Figure 1 shows the normalized eigenvalues $\{\ell_i\}$ and $\{\lambda_i^{(N)}\}$ for some values of the exponent E , assuming $m = 30$, $p = 0$ and $n = 5000$. Since $p = 0$, as n grows, the curve for $\{\ell_i\}$ tends to become a straight line. The role of the nonlinear operation on $\{\ell_i\}$ is to bend the curve according to the value of E . For $E = 2$ and large n , the curve for $\{\lambda_i^{(N)}\}$ tends to lay on the unit semicircle. For $E > 2$, the curves for $\{\lambda_i^{(N)}\}$ are further bent.

Figure 1. Graphical representation of the normalized eigenvalues and indexes.

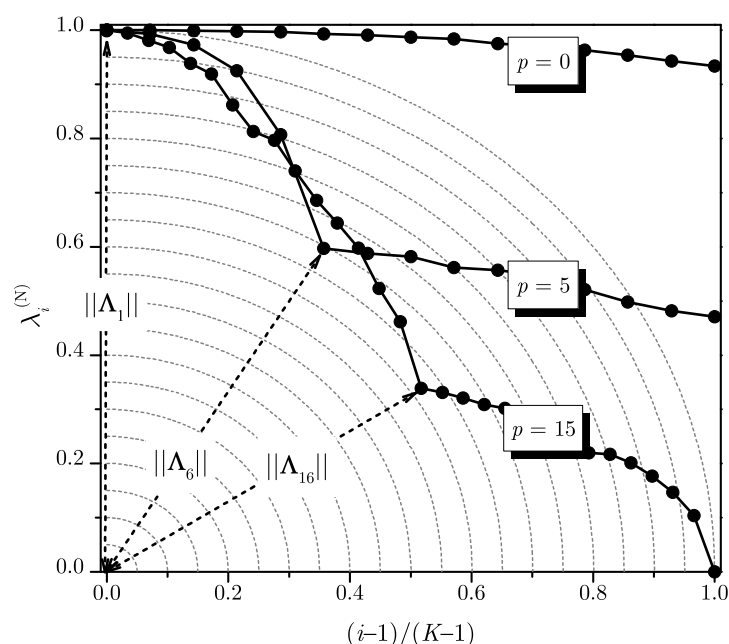


Now, define a vector $\mathbf{\Lambda}_i = [\lambda_i^{(N)} i^{(N)}]^T$. Notice in Figure 1 that, for $E > 2$, $\|\mathbf{\Lambda}_1\| < \|\mathbf{\Lambda}_i\|$ for $i \neq 1$, where $\|\mathbf{\Lambda}_i\|$ is the Euclidean norm of $\mathbf{\Lambda}_i$. For $p > 0$, the smallest $m - p$ eigenvalues of $\hat{\mathbf{R}}$ tend to be equal to the noise variance, and an inflection point in the curve will become evident, at the transition from the largest p and the smallest $m - p$ eigenvalues, as illustrated in Figure 2. This will tend to make $\|\mathbf{\Lambda}_{p+1}\| < \|\mathbf{\Lambda}_i\|$ for $i \neq p + 1$. Then, $\|\mathbf{\Lambda}_i\|$ can be used to estimate the number of sources.

For a small number of sources, however, the placement of the inflection point more to the left will lead to a high chance of having $\|\mathbf{\Lambda}_i\| < \|\mathbf{\Lambda}_{p+1}\|$ for $i > p + 1$. To avoid this, only a subset of vectors

of the set $\{\Lambda_i\}$ should be tested while searching for the smallest norm. Figure 2 illustrates this heuristic for a subset size $K = 15$, with $p = 0$ and $p = 5$, assuming $m = 30$, $n = 5000$ and $E = 5$. Notice that searching within a subset of a size smaller than m is equivalent to pushing the inflection point of the curves to the right, reducing the chance of overestimating the number of sources. If the number of sources increases, the entire set with $K = m$ vectors is used; see an example for $p = 15$ in Figure 2. This heuristic is considered in our algorithm by using the MDL estimate as the reference: we set $K = m$ if the MDL estimate is greater than $m/3$ and set $K = \min \{\lfloor m/3 \rfloor, 3(\hat{p}_{\text{MDL}} + 1)\}$ otherwise. The reasoning behind this heuristic goes as follows: During the development of our algorithm, we realized that the value of K that maximizes the average probability of correct detection, measured over several different sets of system parameters, could not be $K = m$ for any number of sources p . Then, through exhaustive search based on the probability of correct detection, we have empirically found that two values of K would suffice: $K = m/3$ for small values of p and $K = m$ for larger values of p . The reference to distinguish between small values and larger ones has been empirically set to $m/3$ (also through exhaustive search). However, if the number of sensors m is large and the MDL estimate is by far smaller than $m/3$, the subset size had to be set smaller than $m/3$ to increase the probability of correct detection for low values of p . This is the reason for choosing $K = \min \{\lfloor m/3 \rfloor, 3(\hat{p}_{\text{MDL}} + 1)\}$, where \hat{p}_{MDL} is the initial estimate of p in our algorithm, obtained from Equation (5).

Figure 2. Normalized eigenvalues and indexes according to the iMDL algorithm.



The iMDL algorithm is summarized in Algorithm 1.

The choice of the bending exponent E in Equation (8) can be made via simulation, aiming at maximizing the probability of correctly estimating the number of sources, P_c , for a specific set of system parameters. In a scenario of more practical significance, E , can be found as the value that maximizes the average P_c over several sets of parameters. For example, combining the parameters $m = 10, 15, 20, 50$; $n = 50, 100, 200, 500, 1000, 50,000$; $\text{SNR} = -5 \text{ dB}, 0 \text{ dB}, 5 \text{ dB}, 8 \text{ dB}, 10 \text{ dB}$; $p = 2, 5, 10, 15$, the optimal exponent $E = 5$ was found. The numerical results shown hereafter consider $E = 5$.

Algorithm 1 The proposed iMDL algorithm.

```

Compute  $\hat{p}_{\text{MDL}}$  using (5)
if  $\hat{p}_{\text{MDL}} > m/3$  then
     $K = m$ 
else
     $K = \min \{ \lfloor m/3 \rfloor, 3(\hat{p}_{\text{MDL}} + 1) \}$ 
end if
for  $i = 1 \dots m$  do
    Compute  $\lambda_i^{(N)}$  using (8)
end for
for  $j = 1 \dots K$  do
    Compute  $\Lambda_j = [\lambda_j^{(N)} \frac{j-1}{K-1}]^T$ 
end for
Compute  $\hat{p} = \max\{\hat{p}_{\text{MDL}}, \arg \min_j \|\Lambda_j\| - 1\}$ 

```

4. Simulation Setup and Numerical Results

In this section, we present the simulation setup and numerical results comparing the iMDL, the MDL, the AIC and the RMT-based estimators. Each point in the subsequent graphs was generated from 5000 Monte Carlo events. In each event, a new matrix $\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V}$ was generated. The entries in \mathbf{X} are independent and identically distributed (i.i.d.) complex Gaussian variates that simulate Gaussian-distributed and uncorrelated transmitted signals. This situation arises, for example, in wireless communications, where the envelope of most modulated signals are Gaussian-distributed, for instance the amplitude of a multi-carrier signal, such as orthogonal frequency-division multiplexing (OFDM), with a large number of subcarriers, which is the preferred modulation technique in most modern wireless technologies, including several digital television standards. The entries in \mathbf{H} are also i.i.d. complex Gaussian, to simulate a slow flat Rayleigh fading channel, which is constant during each detection interval, changing independently from one interval to the next. The entries in \mathbf{V} are also i.i.d. complex Gaussian, representing the additive thermal noise present at the receiver inputs. Assuming unitary total transmit power, the received SNR is given by $\text{tr}[\mathbf{H}^\dagger \mathbf{H}] / (mp\sigma^2)$, where $\text{tr}[\cdot]$ is the trace of the underlying matrix. To simulate noise uncertainty in the noise variance estimate for the RMT-based algorithm, the noise variance is made a uniform random variable in $[\sigma^2/1.05, 1.05\sigma^2] = [\sigma^2/\rho, \rho\sigma^2]$, where $\rho = 10^{\mu/10}$ and μ is the degree of noise uncertainty, in dB. This situation is denoted by RMT2 in the graphs. Notice that this degree of uncertainty corresponds to $\mu \approx 0.2$ dB, a value that can be considered realistic from a practical standpoint, since typical worst-case values of μ are on the order of 1 dB or even more [27].

In order to simulate signal sources with variable power, we assume that the signal strengths are uniform random variables in $[0.4, 1]$. Additionally, for the RMT-based estimator, the asymptotic false alarm (overestimation) probability is adjusted to $\alpha = 0.1\%$. In all figures, the probability of correctly estimating the number of sources is denoted by P_c , that is, $P_c = \Pr\{\hat{p} = p\}$, where $\Pr\{\cdot\}$ is the probability of a given event. This probability is usually referred to as the probability of detection in the related literature. Similarly, the probability of overestimating the number of sources is denoted by

P_{oe} , that is, $P_{oe} = \Pr\{\hat{p} > p\}$. These probabilities were considered as the performance measurements of the estimators. In our simulations, P_c has been estimated as the ratio between the number of correct detections concerning the number of sources and the total number of test events, in our case 5,000 events for each point in the graphs. In a similar manner, P_{oe} has been estimated in our simulations as the ratio between the number of overestimated detections about the number of sources and the total number of test events.

Figure 3 shows results for P_c and P_{oe} as a function of the number of samples n collected by each sensor, for $m = 30$ sensors, $p = 3$ signals sources, $\text{SNR} = -12$ dB, $\text{SNR} = -10$ dB and $\text{SNR} = -8$ dB. Notice that, with perfectly known noise variance, the RMT-based estimator achieves high detection performance at low SNR, similar to the AIC, and near consistency at large sample sizes, similar to the MDL (which is known to be consistent). However, the RMT-based estimator unveils severe inconsistency at large sample sizes in the presence of a small noise uncertainty $\mu = 0.2$ dB (RMT2 curves). The AIC estimator reaches higher detection performance with a smaller number of samples, but clearly, it is inconsistent at large sample sizes, having a non-negligible P_{oe} for $n \gg 1$. Notice that the proposed iMDL estimator performs better than the MDL estimator for all considered values of n and SNR. It also clearly appears to be a consistent estimator, though we are not able to prove it due to difficult mathematical tractability. Notice also that an increase in the SNR reduces the number of samples necessary for a target P_c , which is a characteristic of all estimators. For example, the number of samples necessary for a target $P_c \approx 0.7$ in the iMDL estimator decreases from 3000 ($\text{SNR} = -12$ dB) to 1200 ($\text{SNR} = -10$ dB) and to 480 ($\text{SNR} = -8$ dB).

Figure 4 depicts numerical results for P_c and P_{oe} as a function of the SNR, for $p = 1$, $p = 3$ and $p = 20$ signals sources, $m = 30$ sensors and $n = 1000$ samples. Again, this figure shows the superior detection capability of the iMDL algorithm when compared with the MDL estimator. Notice that the MDL estimator reaches the same performance of the proposed iMDL algorithm only for $p = 1$. However, this is not meant to state that, if $p = 1$, the MDL and the iMDL estimators will always achieve the same performance, since this depends on the remaining system parameters.

The advantages and drawbacks of the estimators as perceived from Figure 3 are also observable in Figure 4. Particularly, the MDL and the iMDL achieve $P_c = 1$ for high values of SNR. The RMT-based estimator achieves $P_c \approx 1$ in this situation; in fact, it achieves $P_c \approx 1 - \alpha$. However, when there is uncertainty in the noise variance (RMT2 curves), which reflects a practical scenario, the RMT-based estimator produces a non-negligible P_{oe} , preventing it from achieving $P_c \approx 1$. Due to inconsistency, the AIC also produces a non-negligible P_{oe} , leading to $P_c \rightarrow (1 - P_{oe})$ as $\text{SNR} \rightarrow \infty$.

Figure 5 shows results for P_c and P_{oe} as a function of the number of sources p , considering $m = 30$ sensors, $n = 1000$ samples, $\text{SNR} = -10$ dB, $\text{SNR} = -8$ dB and $\text{SNR} = 8$ dB. Similar to the previous results, Figure 5 unveils the superior performance of the iMDL when compared with the MDL, mainly in the low SNR regime, *i.e.*, for $\text{SNR} = -10$ dB and $\text{SNR} = -8$ dB. The iMDL and the MDL performances are almost the same in the high SNR regime, *i.e.*, for $\text{SNR} = 8$ dB. Similarly to the case of $p = 1$ in Figure 4, this is not meant to state that, if the SNR is high, the MDL and the iMDL estimators will always achieve the same performance, for it depends on the remaining system parameters.

Figure 3. P_c and P_{oe} against n for $p = 3$, $m = 30$, SNR = -12 dB (top); SNR = -10 dB (middle) and SNR = -8 dB (bottom).

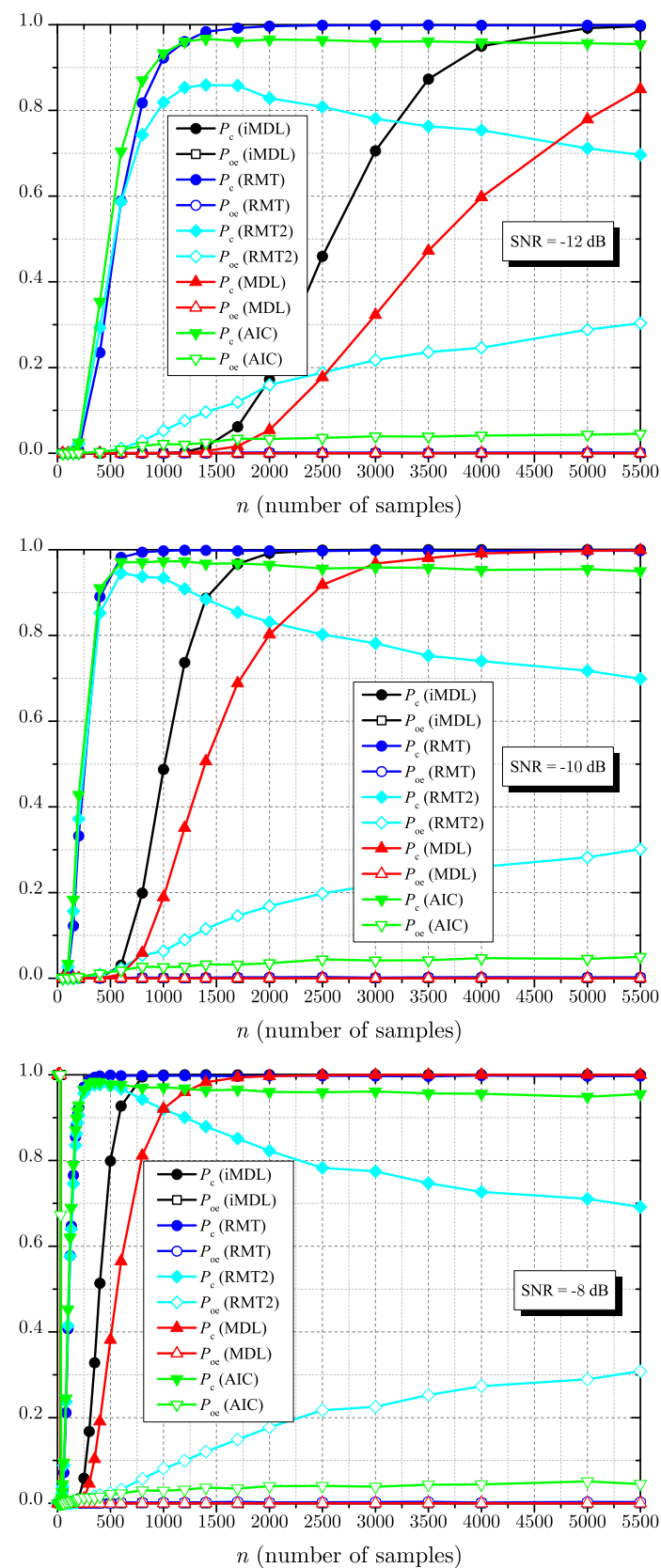


Figure 4. P_c and P_{oe} against SNR for $m = 30$, $n = 1000$, $p = 1$ (top); $p = 3$ (middle) and $p = 20$ (bottom).

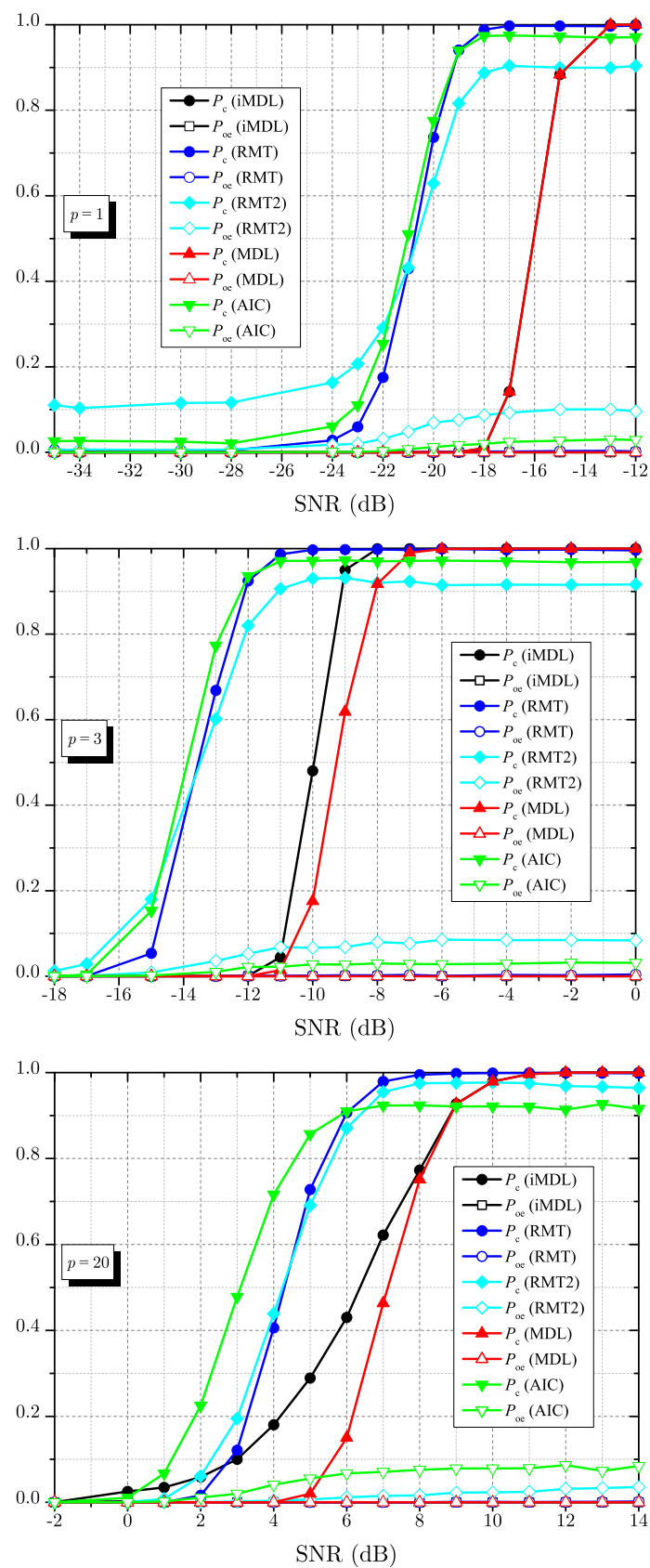


Figure 5. P_c and P_{oc} against p for $m = 30$, $n = 1000$, SNR = -10 dB (top); SNR = -8 dB (middle) and SNR = 8 dB (bottom).

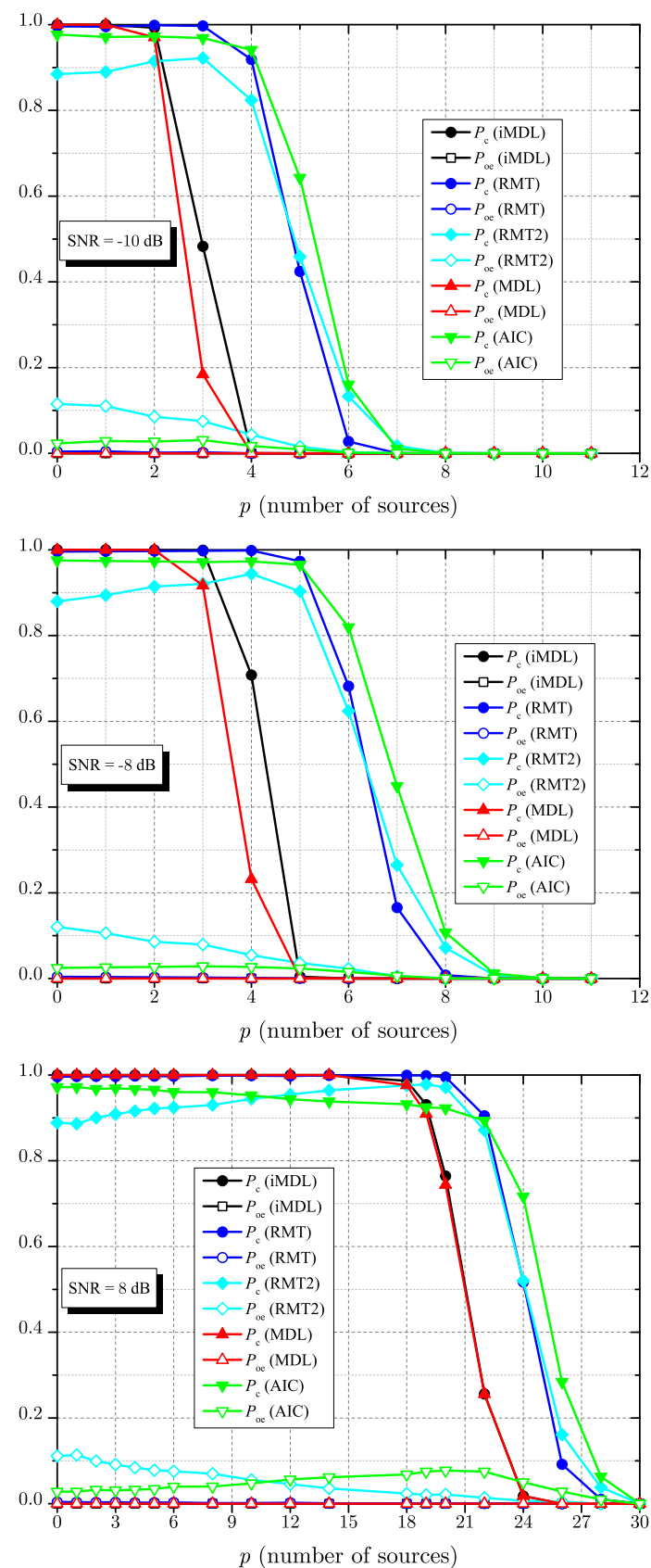
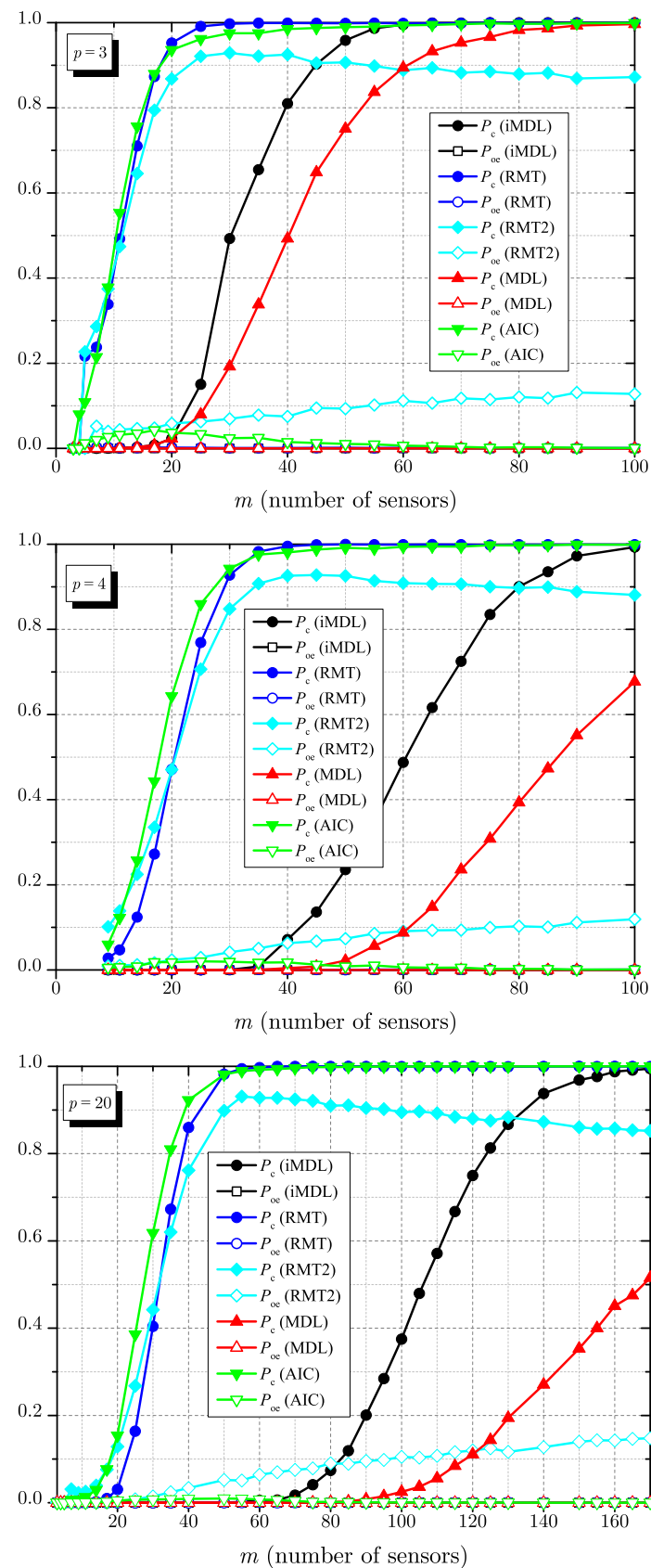


Figure 6. P_c and P_{oe} against m for SNR = −10 dB, $n = 1000$, $p = 3$ (top); $p = 4$ (middle) and $p = 10$ (bottom).



Still referring to Figure 5, for a smaller number of sources, both the MDL and the iMDL achieve $P_c = 1$. In this situation, the RMT-based estimator achieves $P_c \approx 1$ when the noise variance is assumed to be perfectly known, but its performance is degraded if noise uncertainty takes place (RMT2 curves). The maximum estimated number of sources with high detection probability clearly varies from one estimator to another and with the SNR, with the advantage of the RMT and the AIC. We can also see in Figure 5 that, as the SNR increases, the relative differences among these maximum numbers of sources diminish.

In Figure 6, the performance of all estimators are plotted as a function of the number of sensors m , considering $n = 1000$ samples, $\text{SNR} = -10$ dB, $p = 3$, $p = 4$ and $p = 10$ signals sources. One more time, we can see the superiority of the iMDL when compared with the MDL, especially for $p = 10$, *i.e.*, with a large number of signal sources. The advantages and drawbacks of the AIC and the RMT-based estimators can again be observed, now with variable m .

All of the results reported here considered sets of different values of the main system parameters, namely the number of sources p , the number of sensors m , the number of samples n , the signal-to-noise ratio SNR and the noise uncertainty μ . Obviously, different values of these parameters can lead to different performances. As a consequence, a myriad of different scenarios may be exercised. We attest, however, that, for a bunch of other parameters different from those adopted in this paper, we still have verified the superiority of the iMDL against the MDL estimator or at least an equal performance in some few cases.

5. Conclusions

This paper proposed an empirical algorithm for improving the performance of the MDL estimator of the number of sources of signals impinging on multiple sensors. The so-called improved MDL (iMDL) uses the MDL estimate as the input data of the algorithm. As shown by the numerical results, the proposed iMDL algorithm achieves performances that are better than or equal to those achieved by the MDL estimator alone. In some situations, the proposed algorithm outperforms the well-known AIC and the RMT-based estimators. Additionally, in contrast to the AIC and to the RMT-based estimator in the presence of noise variance uncertainty, the iMDL is shown to be consistent, though this has not been supported by a mathematical proof. Furthermore, being an empirical proposition, the iMDL is not guaranteed to be superior to the MDL in all possible scenarios, though it has been shown to be indeed superior in a large number of other situations not shown here. We attest that none of these situations has produced an iMDL performance worse than the MDL performance. Last, but not least, it is worth mentioning that our results can be easily reproduced due to the low complexity of the iMDL algorithm.

Besides wireless communications, scenarios that are more applicable to radar and sonar applications could be investigated, as well. They were put aside due to the limited space, but represent important opportunities for further studies and, hopefully, to broaden the applicability of the proposed algorithm.

The continuity of the work suggests comparing the iMDL, proposed in this paper, with the algorithm proposed in [28] with respect to the robustness against the deviation from the spatially white noise assumption of the original MDL and including the new RMT-based estimator of [22] in the comparisons. Additionally, from an implementation point of view, it will be important to perform measurements with real signals using a test-bed with different types of the sources.

Author Contributions

Both authors have equally contributed to the development of the paper and the related research.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. Wax, M.; Kailath, T. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 387–392.
2. Kritchman, S.; Nadler, B. Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory. *IEEE Trans. Signal Process.* **2009**, *57*, 3930–3941.
3. Fishler, E.; Grosman, M.; Messer, H. Determining the number of discrete alphabet sources from sensor data. *EURASIP J. Appl. Signal Process.* **2005**, *2004*, 4–12.
4. Fishler, E.; Poor, H.V. Estimation of the number of sources in unbalanced arrays via information theoretic criteria. *IEEE Trans. Signal Process.* **2005**, *53*, 3543–3553.
5. Stoica, P.; Selen, Y. Model-order selection: A review of information criterion rules. *IEEE Signal Process. Mag.* **2004**, *21*, doi:10.1109/MSP.2004.1311138.
6. Bruckstein, A.; Shan, T.; Kailath, T. The Resolution of Overlapping Echoes. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *ASSP-33*, 1357–1367.
7. Atashbar, M.; Kahaei, M. Direction-Of-Arrival Estimation Using AMLSS Method. *IEEE Latin Am. Trans.* **2012**, *10*, 2053–2058.
8. Naik, G.R.; Kumar, D.K. Dimensional reduction using Blind Source Separation for identifying sources. *Int. J. Innov. Comput. Inf. Control* **2011**, *7*, 989–1000.
9. Naik, G.R.; Kumar, D.K. Determining Number of Independent Sources in Undercomplete Mixture. *EURASIP J. Adv. Signal Process.* **2009**, *2009*, 1–5.
10. Makeig, S.; Bell, A.J.; Jung, T.P.; Sejnowski, T.J. Independent Component Analysis of Electroencephalographic Data. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1996; pp. 145–151.
11. Naik, G.; Kumar, D.; Weghorn, H. ICA based identification of sources in sEMG. In Proceedings of the 3rd International Conference on Intelligent Sensors, Sensor Networks and Information, 2007, Melbourne, Australia, 3–6 December 2007; pp. 619–624.

12. Kritchman, S.; Nadler, B. Determining the number of components in a factor model from limited noisy data. *Chem. Int. Lab. Syst.* **2008**, *94*, 19–32.
13. Swamy, R.K.; Murty, S.R.; Yegnanarayana, B. Determining number of speakers from multispeaker speech signals using excitation source information. *IEEE Trans. Signal Process.* **2007**, *14*, 481–484.
14. Chiani, M.; Win, M.Z. Estimating the number of Signals observed by multiple sensors. In Proceedings of the International Workshop on Cognitive Information Processing, Elba Island, Tuscany, Italy, 14–16 June 2010.
15. Braga, A.J.; de Souza, R.A.A.; Guimaraes, D.A. Eigenvalue-based techniques for continuous sensing model in MIMO CR networks. In Proceedings of the 2013 IEEE Latin-America Conference on Communications (LATINCOM), Santiago, Chile, 24–26 November 2013; pp. 1–6.
16. Araki, S.; Sawada, H.; Mukai, R.; Makino, S. DOA Estimation for Multiple Sparse Sources with Arbitrarily Arranged Multiple Sensors. *Signal Process. Syst.* **2011**, *63*, 265–275.
17. Lee, J.Y.; Harlow, Z.; Collier, T.C.; Taylor, C.E.; Yao, K. A beamforming method for multiple source DOA estimation, spectrum separation and localization from field data. In Proceedings of the the 11th international conference on Information Processing in Sensor Networks, Beijing, China, 16–19 April 2012; pp. 123–124.
18. Chen, P.; Wicks, M.; Adve, R. Development of a statistical procedure for detecting the number of signals in a radar measurement. *IEE Proc. Radar Sonar Navig.* **2001**, *148*, 219–226.
19. Ghasemi, A.; Sousa, E. Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs. *IEEE Commun. Mag.* **2008**, *46*, 32–39.
20. Gillblad, D.; Steinert, R.; Ferreira, D. Estimating the Parameters of Randomly Interleaved Markov Models. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW'09), Miami, FL, USA, 6–9 December 2009; pp. 308–313.
21. Cheng, W.; Zhang, Z.; Cao, H.; He, Z.; Zhu, G. A Comparative Study of Information-Based Source Number Estimation Methods and Experimental Validations on Mechanical Systems. *Sensors* **2014**, *14*, 7625–7646.
22. Yi, H. Local Signal Search Enhanced RMT Estimator for Estimating the Number of Signals Based on Random Matrix Theory. Available online: <http://arxiv.org/abs/1405.4713> (accessed on 14 October 2014).
23. Akaike, H. Information Theory and Extension of the Maximum Likelihood Principle. In *Selected Papers of Hirotugu Akaike*; Parzen, E., Tanabe, K., Kitagawa, G. Eds.; Springer New York: New York, NY, USA, 1998; pp. 199–213.
24. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464.
25. Rissanen, J. Modeling by shortest data description. *Automatica* **1978**, *14*, 465–471.
26. Nadler, B. On the Distribution of the Ratio of the Largest Eigenvalue to the Trace of a Wishart Matrix. Available online: http://www.wisdom.weizmann.ac.il/~nadler/Wishart_Ratio_Trace/TW_ratio.html#Code (accessed on 14 October 2014).
27. Wei, L.; Tirkkonen, O. Spectrum Sensing in the Presence of Multiple Primary Users. *IEEE Trans. Commun.* **2012**, *60*, 1268–1277.

28. Huang, L.; Long, T.; Wu, S. Source Enumeration for High-Resolution Array Processing Using Improved Gerschgorin Radii Without Eigendecomposition. *IEEE Trans. Signal Process.* **2008**, *56*, 5916–5925.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).