CONSISTENCY OF DETECTION OF THE NUMBER OF SIGNALS USING MULTIPLE HYPOTHESIS TESTS

Pei-Jung Chung

Institute for Digital Communications
Joint Research Institute for Signal & Image Processing
School of Engineering and Electronics, The University of Edinburgh, UK
peijung.chung@gmail.com

ABSTRACT

We study a multiple hypothesis test for detecting signals embedded in noisy observations of a sensor array. The global level of the multiple test is controlled by the false discovery rate (FDR) criterion recently suggested by Benjamini and Hochberg. In previous studies, the suggested procedure has shown promising results on both simulated and real data. Here we carefully examine the consistency property of the multiple test procedure. Applying the asymptotic properties of maximum likelihood (ML) estimation, we prove strong consistency under a mild condition on signal and noise eigenvalues. This condition enables us to find the minimum SNR to ensure consistency. Our analysis is further confirmed by numerical experiments conducted under low SNRs and closely located signal sources.

Keywords: array signal processing, signal detection, multiple test

1. INTRODUCTION

Determination of the number of signals embedded in noisy sensor outputs is a key issue in array processing and related applications [2]. Many high resolution methods, such as the maximum likelihood (ML) approach or MUSIC, assume a known number of signals. Performance of these techniques depends strongly on this knowledge [4]. In radar or geophysics, deciding how many incoming waves is as important as estimating the associated propagation parameters.

In [5] [6], we suggested a detection procedure based on multiple testing under false discovery rate (FDR) consideration. Compared to conventional methods based on information theoretic criteria, such as Akaike's information criterion (AIC) [11] or Rissanen's minimum description length (MDL) [12] [13], the multiple test procedure offers a higher probability of correct detection and a lower SNR threshold in the finite sample case. Furthermore, the proposed test can be applied to both narrow band and broadband signals.

In this work, we investigate the consistency property of the multiple test procedure. Applying asymptotic results of ML estimation under misspecified models, we prove that the multiple test procedure is a strongly consistent estimator for the number of signals under a mild condition on signal and noise eigenvalues. More precisely, the ratio between the smallest signal eigenvalue and the noise eigenvalue needs to exceed a threshold to ensure consistency. We derive an explicit expression for the threshold depending only on the number of signals and the number of sensors. With this condition, we can easily predict the region where consistency is guaranteed.

In the following section, we give a brief description of the signal model. Section 3 introduces the multiple test procedure and the FDR criterion. In section 4, we shall prove the consistency property of the multiple test procedure. Simulation results are presented and discussed in section 5. Our concluding remarks are given in section 6.

2. SIGNAL MODEL

Consider an array of n sensors receiving m narrow band signals emitted by far-field sources located at $\boldsymbol{\theta} = [\,\theta_1,\ldots,\theta_m]^T$. The array output $\boldsymbol{x}(t) \in \mathbb{C}^{n\times 1}$ can be expressed as

$$\boldsymbol{x}(t) = \boldsymbol{H}_m(\boldsymbol{\theta}_m)\boldsymbol{s}_m(t) + \boldsymbol{n}(t), \quad t = 1,\ldots,T \tag{1}$$

where the ith column of the matrix

$$\boldsymbol{H}_m(\boldsymbol{\theta}_m) = [\boldsymbol{d}(\theta_1)\cdots\boldsymbol{d}(\theta_i)\cdots\boldsymbol{d}(\theta_m)] \tag{2}$$

$\boldsymbol{d}(\theta_i) \in \mathbb{C}^{n\times 1}$ is the steering vector associated with the signal arriving from the direction θ_i. The unknown signal waveform $\boldsymbol{s}_m(t) = [s_1(t),\ldots,s_m(t)]^T \in \mathbb{C}^{m\times 1}$ is considered as a realization of a stationary random process. Furthermore, the noise vector $\boldsymbol{n}(t) \in \mathbb{C}^{n\times 1}$ is independent, identically complex normally distributed with zero mean and covariance matrix $\nu\boldsymbol{I}$, where ν is an unknown noise spectral parameter and $\boldsymbol{I}$ is an identity matrix of corresponding dimension. Given the set of observations $\{\boldsymbol{x}(t)\}_{t=1}^T$, the problem of central interest is to determine the number of signals m.

3. SIGNAL DETECTION USING A MULTIPLE HYPOTHESIS TEST

We formulate the problem of detecting the number of signals as a multiple hypothesis test. Let M denote the maximal number of signals. The following procedure detects one signal after another. More precisely, for $m = 1$,

$$\begin{aligned} H_1 &: \text{Data contains only noise.} \\ &\quad\ \boldsymbol{x}(t) = \boldsymbol{n}(t) \\ A_1 &: \text{Data contains at least 1 signals.} \\ &\quad\ \boldsymbol{x}(t) = \boldsymbol{H}_1(\boldsymbol{\theta}_1)\boldsymbol{s}_1(t) + \boldsymbol{n}(t) \end{aligned} \tag{3}$$

P.-J. Chung acknowledges support of her position from the Scottish Funding Council and their support of the Joint Research Institute with the Heriot-Watt University as a component part of the Edinburgh Research Partnership.

For $m = 2, \ldots, M$

$$
\begin{array}{lll}
H_m & : & \text{Data contains at most } (m-1) \text{ signals.} \\
 & & \boldsymbol{x}(t) = \boldsymbol{H}_{m-1}(\boldsymbol{\theta}_{m-1})\boldsymbol{s}_{m-1}(t) + \boldsymbol{n}(t) \\
A_m & : & \text{Data contains at least } m \text{ signals.} \\
 & & \boldsymbol{x}(t) = \boldsymbol{H}_m(\boldsymbol{\theta}_m)\boldsymbol{s}_m(t) + \boldsymbol{n}(t)
\end{array}
\tag{4}
$$

We use the subscripts $(m-1)$ and m to emphasize the dimension of the steering matrix and the signal vector under the null hypothesis H_m and the alternative A_m, respectively. Let $\{i_1, i_2, \ldots, i_r\}$ be an arbitrary subset of $\{1, 2, \ldots, M\}$ and suppose that among M hypotheses, r are rejected, namely $H_{i_1}, H_{i_2}, \ldots, H_{i_r}$. Then the number of signals is determined by the maximal index of $H_{i_1}, H_{i_2}, \ldots, H_{i_r}$. Namely,

$$\hat{m} := \max\{i_1, i_2, \ldots, i_r\}. \tag{5}$$

Which hypotheses are to be rejected depends on the adopted error criterion. In this work, we shall apply the Benjamini-Hochberg procedure to control the false discovery rate.

Based on the likelihood ratio (LR) principle, we obtain the test statistics $T_m(\hat{\boldsymbol{\theta}}_m)$, $(m = 1, \ldots, M)$ as follows.

$$
\begin{aligned}
T_m(\hat{\boldsymbol{\theta}}_m) &= \log\left(\frac{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_{m-1}(\hat{\boldsymbol{\theta}}_{m-1}))\hat{\boldsymbol{R}}]}{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}]}\right) \qquad (6)\\
&= \log\left(1 + \frac{n_1}{n_2}F_m(\hat{\boldsymbol{\theta}}_m)\right), \qquad (7)
\end{aligned}
$$

where $\hat{\boldsymbol{R}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}(t)\boldsymbol{x}(t)^H$ and $\boldsymbol{P}(\hat{\boldsymbol{\theta}}_m)$ is the projection matrix onto the subspace spanned by the columns of $\boldsymbol{H}_m(\hat{\boldsymbol{\theta}}_m)$. When $m = 1$, we define $\boldsymbol{P}_0(\cdot) = \boldsymbol{0}$. $\hat{\boldsymbol{\theta}}_m$ represents the ML estimate assuming that m signals are present in the observation.

Under hypothesis H_m, the statistic

$$F_m(\hat{\boldsymbol{\theta}}_m) = \frac{n_2}{n_1}\frac{\mathrm{tr}[(\boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m) - \boldsymbol{P}_{m-1}(\hat{\boldsymbol{\theta}}_{m-1}))\hat{\boldsymbol{R}}]}{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}]} \tag{8}$$

is F_{n_1,n_2}–distributed where the degrees of freedom n_1, n_2 are given by [8] [9]

$$n_1 = T(2 + r_m), \quad n_2 = T(2r_{\boldsymbol{x}} - 2m - r_m) \tag{9}$$

with $r_{\boldsymbol{x}} = \dim(\boldsymbol{x}(t)) = n$ being the dimension of $\boldsymbol{x}(t)$ and $r_m = \dim(\theta_m)$ denoting the dimension of the parameter vector associated with the mth signal. Since we only consider the direction of arrival (DOA) parameter, $r_m = 1$. More details about the F_{n_1,n_2}-distribution can be found in [7].

From eq. (7) it is easy to see that in the narrow band case, the LR test is equivalent to the F-test proposed by Shumway [10]. The F-test uses $F_m(\hat{\boldsymbol{\theta}}_m)$ in testing H_m against A_m. Given $(m-1)$ signals, whether a further signal exists is decided by whether the estimated increase in SNR is large enough.

Control of the False Discovery Rate

The control of type one error is an important issue in multiple inferences. A type one error occurs when the null hypothesis H_m is wrongly rejected. The traditional concern in multiple hypothesis problems has been about controlling the familywise error-rate (FWE). Given a certain significance level α, the control of FWE requires each of the M tests to be conducted at a lower level. When the number of tests increases, the power of the the FWE-controlling procedures is substantially reduced. The false discovery rate (FDR), suggested by Benjamini and Hochberg [1], is a completely different point of view for considering the errors in multiple testing. The FDR is defined as the expected proportion of errors among the rejected hypotheses. If all null hypotheses $\{H_1, H_2, \ldots, H_M\}$ are true, the FDR-controlling procedure controls the traditional FWE. But when many hypotheses are rejected, an erroneous rejection is not as crucial for drawing conclusion from the whole family of tests, the FDR is a desirable error rate to control.

Assume that among the M tested hypotheses $\{H_1, H_2, \ldots, H_M\}$, M_0 are true null hypotheses. Let $\{p_1, p_2, \ldots, p_M\}$ be the p-values (observed significance values) corresponding to the test statistics $\{T_1, T_2, \ldots, T_M\}$. By definition, $p_m = 1 - P_{H_m}(T_m)$ where P_{H_m} is the distribution function under H_m. Benjamini and Hochberg showed that when the test statistics *corresponding to the true null hypotheses* are independent, the following procedure controls the FDR at level $q \cdot M_0/M \leq q$[1].

The Benjamini Hochberg Procedure

Define

$$k = \max\left\{m : p_{(m)} \leq \frac{m}{M}q\right\} \tag{10}$$

and reject $H_{(1)} \ldots H_{(k)}$. If no such k exists, reject no hypothesis.

4. CONSISTENCY

In this section, we show that the multiple test procedure described previously yields a strongly consistent estimator under a mild condition. Let $\boldsymbol{R} = \boldsymbol{H}_0(\boldsymbol{\theta}_0)\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{H}_0(\boldsymbol{\theta}_0)^H + \nu\boldsymbol{I}$ denote the ensemble average of the sample covariance matrix $\hat{\boldsymbol{R}}$ where $\boldsymbol{H}_0$ and $\boldsymbol{C}_{\boldsymbol{s}} = E\boldsymbol{s}_{m_0}(t)\boldsymbol{s}_{m_0}(t)^H$ represent the steering matrix and signal covariance matrix associated with the true number of signals m_0. It is well known that the eigenvalues of $\boldsymbol{R}$ are characterized by $\lambda_1 \geq \ldots \geq \lambda_{m_0} > \lambda_{m_0+1} = \ldots = \lambda_n = \nu$. The largest m_0 eigenvalues and the remaining eigenvalues are referred to as the signal eigenvalues and noise eigenvalues, respectively.

Theorem (Consistency) Assuming that

$$\frac{\lambda_{m_0}}{\nu} > c_0, \qquad c_0 = \frac{1.5}{1 - \frac{1}{2(n-m_0)}}\,. \tag{11}$$

Then the multiple test procedure (3), (4) yields a strongly consistent estimator for the number of signals, i.e. $\hat{m} \to m_0$ with probability one as $T \to \infty$.

Proof As we determine the number of signals by (5), it suffices to show that hypotheses H_m's are retained for $m > m_0$ and H_{m_0} is rejected as T goes to infinity.

Assuming the signal model in section 2, the sample covariance matrix $\hat{\boldsymbol{R}}$ approaches the ensemble average $\boldsymbol{R}$ with probability one. From the asymptotic properties of ML estimation under misspecified models, we know that the ML estimate $\hat{\boldsymbol{\theta}}_m$ converges almost surely to the minimizing point of the following criterion [4]

$$\boldsymbol{\theta}_m^* = \arg\min_{\boldsymbol{\theta}_m} \mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\boldsymbol{\theta}_m))\boldsymbol{R}], \tag{12}$$

where m is the assumed number of signals. Furthermore,

$$\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\boldsymbol{\theta}_m^*))\boldsymbol{R}] = \lambda_{m+1} + \ldots + \lambda_n, \tag{13}$$
$$\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_{m-1}(\boldsymbol{\theta}_{m-1}^*))\boldsymbol{R}] = \lambda_m + \lambda_{m+1} + \ldots + \lambda_n. \tag{14}$$

Consequently, as $T \to \infty$, the statistic $T_m(\hat{\boldsymbol{\theta}}_m)$ in (6) approaches

$$T_m^* = \log(1 + \frac{\lambda_m}{\lambda_{m+1} + \ldots + \lambda_n}) \tag{15}$$

and the statistic $F_m(\hat{\boldsymbol{\theta}}_m)$ in (8) approaches

$$F_m^* = \frac{\lambda_m/n_1}{(\lambda_{m+1} + \ldots + \lambda_n)/n_2} = \frac{\frac{\lambda_m}{(1+1/2)}}{\frac{\lambda_{m+1}+\ldots+\lambda_n}{(n-m-1/2)}} \tag{16}$$

with probability one. The degrees of freedom n_1 and n_2 are given by (9) with $r_m = 1$.

For $m > m_0$, we have $\lambda_m = \lambda_{m+1} = \ldots = \lambda_n = \nu$. Therefore,

$$F_m^* = \frac{1 - \frac{1}{2(n-m)}}{1 + \frac{1}{2}} < 1. \tag{17}$$

For $m = m_0$,

$$F_{m_0}^* = \frac{\frac{\lambda_{m_0}}{(1+1/2)}}{\frac{\nu(n-m_0)}{(n-m_0-1/2)}} = \frac{\lambda_{m_0}}{\nu} \cdot \frac{1 - \frac{1}{2(n-m_0)}}{1.5}. \tag{18}$$

Recall that at each test stage, the p-value is obtained under the assumption that $F_m(\hat{\boldsymbol{\theta}}_m)$ is F_{n_1,n_2}-distributed where n_1 and n_2 grows with increasing T. Applying the central limit theorem, the asymptotic F_{n_1,n_2}-distribution can be approximated by a normal distribution with mean 1 and variance $2/n_1$, denoted by $\mathcal{N}(1, 2/n_1)$ [3].

For $m > m_0$, from (17) we know that $F_m^* < 1$. The asymptotic distribution $\mathcal{N}(1, 2/n_1)$ leads to an observed significance level p_m larger than 0.5. Given a reasonably chosen FDR level, for example $q = 0.1$, p_m is compared with $q' \leq q$. As a result,

$$p_m > 0.5 > q \geq q'. \tag{19}$$

The hypotheses H_ms for $m > m_0$ are all retained as $T \to \infty$.

For $m = m_0$, to guarantee H_{m_0} to be rejected, the p-value must satisfy

$$p_{m_0} = 1 - \Phi(\frac{F_{m_0}^* - 1}{\sqrt{2/n_1}}) \leq q/M, \tag{20}$$

where $\Phi(\cdot)$ is the distribution function of standard normal distribution $\mathcal{N}(0, 1)$. Thus,

$$F_{m_0}^* \geq 1 + \sqrt{\frac{2}{n_1}}\Phi^{-1}(1 - \frac{q}{M}) = 1 + O(n_1^{-1/2}). \tag{21}$$

This can be achieved if

$$F_{m_0}^* = \frac{\lambda_{m_0}}{\nu} \cdot \frac{1 - \frac{1}{2(n-m_0)}}{1.5} > 1. \tag{22}$$

Thus, assuming the condition (11), H_{m_0} is rejected with probability one as $T \to \infty$. □

The above proof shows that the consistency property is governed by the condition (11). Eq. (11) implies that the ratio between the smallest signal eigenvalue and the noise eigenvalue must be at least as large as $c_0 = 1.5/(1 - \frac{1}{2(n-m_0)})$. The threshold c_0 is determined by the number of sensors n and the true number of signals m_0. The ratio λ_{m_0}/ν is closely related to SNR associated with the weakest source signal. To ensure consistency, SNR needs to exceed a threshold.

5. SIMULATION

In this section, we examine the impact of the ratio λ_{m_0}/ν on the consistency property of the multiple test procedure. In particular, we shall investigate whether eq. (11) is a sufficient condition for consistency.

In the first experiment, a uniform linear array of $n = 15$ sensors with inter-element spacings of half a wavelength $\lambda/2$ is employed. The narrow band signals are generated by $m_0 = 3$ uncorrelated signals located at $[-30° \; 20° \; 24°]$ of various strengths. The difference of signal strengths is $[-1 \; 1 \; 0]$ dB where 0 dB corresponds to the reference signal. The SNR, defined as $10\log\left(|s_i(t)|^2/\nu\right)$ for the ith signal, varies from -12 to -6 dB in a 1 dB step. To simulate large sample scenarios, we use $T = 5000$ snapshots for each of the 100 trials performed. The FDR is controlled at level $q = 0.1$.

Fig. 1 presents the probability of correct detection vs. SNR in the upper part and the empirical mean of the ratio λ_{m_0}/ν in the lower part. By correct detection, we mean $\hat{m} = m_0$. The dashed line corresponds to the c_0 level. Given $m_0 = 3$, $n = 15$, by (11), the lower bound on the ratio λ_{m_0}/ν is $c_0 = 1.565$. For SNR less than -10 dB, the probability of correct detection is zero. In the same SNR region, one can observe that the estimated λ_{m_0}/ν is less than the required minimum value c_0. At SNR $= -10$ dB, the probability of correct detection is approximately 0.5. Interestingly, the average value of λ_{m_0}/ν is 1.58 which is very close to $c_0 = 1.565$. For SNR ≥ -9 dB, the ratio $\lambda_{m_0}/\nu > c_0$ and the probability of correct detection reaches 100%. The confirms the importance of the condition (11) to consistency.

The second experiment uses a similar scenario as that in the previous experiment. The number of sensors of the array is reduced to $n = 7$. By (11), a smaller n leads to a larger $c_0 = 1.714$. In addition, the ratio λ_{m_0}/ν becomes lower. Consequently, one needs higher SNRs to ensure consistency. As presented in fig. 2, the probability of correct detection achieves 100% for SNR ≥ 2 dB where the condition (11) is satisfied. At SNR $= 1$ dB, λ_{m_0}/ν is very close to c_0 and the probability of correct detection is about 50%. For SNR ≤ 0 dB, $\lambda_{m_0}/\nu < c_0$ and the probability of correct detection is zero.

In summary, the asymptotic behavior of the multiple hypothesis test for determining the number of signals is governed by the ratio between the smallest signal eigenvalue and noise eigenvalue. Numerical results show that the lower bound on this ratio provides an accurate estimate for the minimum SNR to ensure consistency.

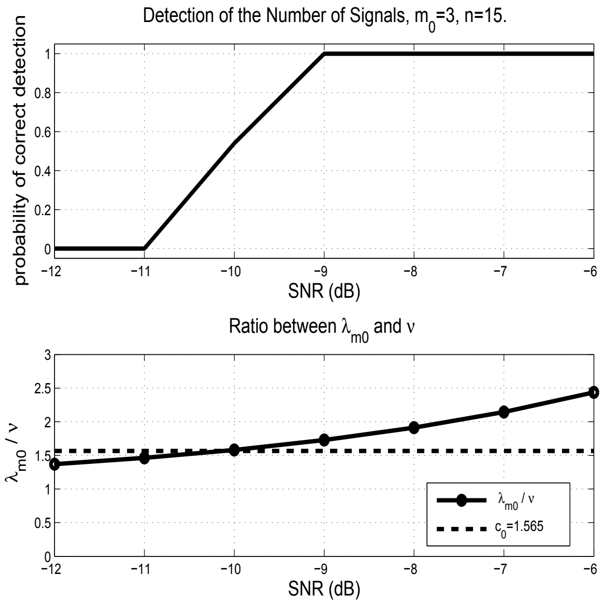

Fig. 1. Upper panel: empirical probability of correct detection. Lower panel: estimated ratio between λ_{m_0} and ν. $m_0 = 3$, $n = 15$. SNR $= [-12 : 1 : -6]$ dB.

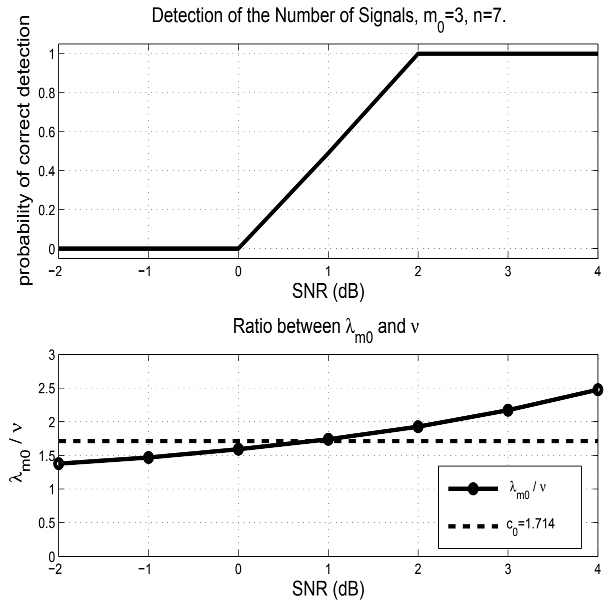

Fig. 2. Upper panel: empirical probability of correct detection. Lower panel: estimated ratio between λ_{m_0} and ν. $m_0 = 3$, $n = 7$. SNR $= [-2 : 1 : 4]$ dB.

6. CONCLUSION

This work discusses the consistency property of a multiple hypothesis test for estimating the number of signals. To increase the power of the test, the global significance level is controlled by the FDR criterion. Applying asymptotic properties of ML estimation, we proved that this procedure is a strongly consistent estimator when the ratio between the smallest signal eigenvalue and the noise eigenvalue exceeds a certain threshold. We derived an explicit expression for the threshold that depends only on the number of signals and the number of sensors. With this expression, one can easily find the minimum SNR to ensure consistency. Our analysis is further validated by numerical experiments carried out with closely located signal sources and low SNRs.

7. REFERENCES

[1] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57:289–300, 1995.

[2] J. F. Böhme. Array processing. In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, pages 1–63. Prentice Hall, Englewood Cliffs, N.J., 1991.

[3] Johann F. Böhme. *Stochastische Signale*. Teubner Verlag, 2nd edition, 1998.

[4] P.-J. Chung. Stochastic ML estimation under misspecified number of signals. In *Proc. EUSIPCO*, Florence, Italy, September 2006.

[5] P.-J. Chung, J. F. Böhme, C. F. Mecklenbräuker, and A. O. Hero. On signal detection using the Benjamini-Hochberg procedure. In *Proc. IEEE Workshop on Statistical and Signal Processing*, Bordeaux, France, July 2005.

[6] P.-J. Chung, N. Czink, and C. F. Mecklenbräuker. Model order selection for multipath MIMO channels using the Benjamini-Hochberg procedure. In *ITG-/IEEE-Workshop on Smart Antennas*, Ulm, Germany, March 2006.

[7] N.L. Johnson and S. Kotz. *Continuous Univariate Distributions — 1*. Applied Probability and Statistics. John Wiley & Sons, New York, 1970.

[8] D. Maiwald. *Breitbandverfahren zur Signalentdeckung und –ortung mit Sensorgruppen in Seismik– und Sonaranwendungen*. Dr.–Ing. Dissertation, Dept. of Electrical Engineering, Ruhr–Universität Bochum, Shaker Verlag, Aachen, 1995.

[9] D. Maiwald and J. F. Böhme. Multiple testing for seismic data using bootstrap. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, volume VI, pages 89–92, Adelaide, 1994.

[10] R.H. Shumway. Replicated time-series regression: An approach to signal estimation and detection. In D.R. Brillinger and P.R. Krishnaiah, editors, *Handbook of Statistics, Vol. 3*, pages 383–408. Elsevier Science Publishers B.V., 1983.

[11] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE ASSP*, 33(2):387–392, April 1983.

[12] M. Wax and I. Ziskind. Detection of the number of coherent signals by the MDL principle. *IEEE ASSP*, 37(8):1190–1196, August 1989.

[13] M. Wong, Q. T. Zou, and J. P. Reilly. On information theoretic criterion for determining the number of signals in high resolution array processing. *IEEE Trans. Acoust., Speech, Signal Processing*, 38(11):1959–1971, November 1990.