ON SIGNAL DETECTION USING THE BENJAMINI-HOCHBERG PROCEDURE

Pei-Jung Chung[1)], *Johann F. Böhme*[2)], *Christoph F. Mecklenbräuker*[3)], *Alfred O. Hero*[4)]

[1)]Dept. of Electronics Engineering, National Chiao Tung University, Taiwan
[2)]Dept. of Electrical Engineering and Information Science
Ruhr-Universität Bochum, Germany
[3)]ftw. Telecommunications Research Center Vienna, Austria
[4)]Dept. of Electrical Engineering and Computer Science
University of Michigan, USA

peijung_chung@yahoo.com, boehme@sth.rub.de, cfm@ftw.at, hero@eecs.umich.edu

ABSTRACT

We investigate a multiple hypothesis test designed 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 instead of the classical familywise error rate (FWE) criterion. In the previous study [3], the suggested procedure has shown promising results on simulated data. Here we carefully examine the independence condition required by the Benjamini Hochberg procedure to ensure the control of FDR. Based on the properties of beta distribution, we proved the applicability of the Benjamini Hochberg procedure to the underlying test. Further simulation results show that the false alarm rate is less than 0.02 for a choice of FDR as high as 0.1, implying the reliability of the test has not been affected by the increase in power.

1. INTRODUCTION

This work discusses signal detection using a multiple hypothesis test. Estimating the number of signals embedded in noisy observations is a key issue in array processing, harmonic retrieval, wireless communication and geophysical application. In [2] [7], a multiple testing procedure was suggested to determine the number of signals. Therein, a Bonferroni-Holm procedure [4] was used to control the familywise error-rate (FWE), the probability of erroneously rejecting any of the true hypotheses. As the control of FWE requires each test to be conducted at a significantly lower level, the Bonferroni-Holm procedure often leads to conservative results. To overcome this drawback, we adopted the false discovery rate (FDR) criterion suggested by Benjamini and Hochberg [1] to keep balance between type one error control and power [3]. Despite of the successful application of the Benjamini-Hochberg procedure in [3], a crucial question remains: *Does the Benjamini-Hochberg procedure control the FDR in our detection problem?* The goal of this work is to examine the test statistics carefully and show that the conditions required by the Benjamini Hochberg procedure are satisfied.

This paper is organized as follows. We give a brief description of the signal model in the next section. Then we present the multiple test procedure for signal detection. Section 4 introduces the idea of false discovery rate (FDR) and the Benjamini Hochberg procedure. In the subsequent section we show that the condition required by the Benjamini Hochberg procedure is satisfied in the proposed approach. Simulation results are presented and discussed in section 6. Our concluding remarks are given in section 7.

This work was supported in part by the National Science Council, Taiwan, under Contract NSC 94-2213-E-009-050 and partially supported by DARPA under the MOSAIC program, contract DAAD19-02-C-0055. The work of C.F. Mecklenbräuker was carried out with funding by Kplus, Infineon Technologies, and ARC Seibersdorf Research GmbH.

2. PROBLEM FORMULATION

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

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

where the ith column of the matrix

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

$\boldsymbol{d}(\theta_i) \in \mathbb{C}^{n\times 1}$ is the steering vector associated with the signal arriving from the direction θ_i. The signal waveform $\boldsymbol{s}(t) = [s_1(t),\ldots,s_m(t)]^T \in \mathbb{C}^{m\times 1}$ is considered as deterministic and unknown. 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 $\sigma^2 \boldsymbol{I}$, where σ^2 is the unknown noise spectral parameter and $\boldsymbol{I}$ is an identity matrix of corresponding dimension. Based on 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 sources. The following procedure detects one signal after another. More precisely, for $m = 1$,

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

For $m = 2, \ldots, M$

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

The steering matrix and signal vector are given by $\boldsymbol{H}_m(\boldsymbol{\theta}_m) = [\boldsymbol{d}(\theta_1), \ldots, \boldsymbol{d}(\theta_m)]$ and $\boldsymbol{s}_m(t) = [s_1(t), \ldots, s_m(t)]^T$, respectively.

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

$$T_m(\hat{\boldsymbol{\theta}}_m) = \log\left(\frac{\operatorname{tr}[(\boldsymbol{I} - \boldsymbol{P}_{m-1}(\hat{\boldsymbol{\theta}}_{m-1}))\hat{\boldsymbol{R}}]}{\operatorname{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}]}\right) \tag{5}$$

$$= \log\left(1 + \frac{n_1}{n_2} F_m(\hat{\boldsymbol{\theta}}_m)\right), \tag{6}$$

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

$$F_m(\hat{\boldsymbol{\theta}}_m) = \frac{n_2}{n_1} \frac{\operatorname{tr}[(\boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m) - \boldsymbol{P}_{m-1}(\hat{\boldsymbol{\theta}}_{m-1}))\hat{\boldsymbol{R}}]}{\operatorname{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}]} \tag{7}$$

is F_{n_1,n_2}–distributed where the degrees of freedom n_1, n_2 are given by [7]

$$n_1 = T(2 + r_m), \quad n_2 = T(2r_{\boldsymbol{x}} - 2m - r_{m-1}) \tag{8}$$

with $r_{\boldsymbol{x}} = \dim(\boldsymbol{x}(t)) = n$ and $r_m = \dim(\boldsymbol{\theta}_m) = m$.

From eq. (6) it is easy to see that in the narrow band case, the LR test is equivalent to the F-test proposed by Shumway [8]. 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.

4. 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 such as Bonferroni-type procedures [4] 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

$$k = \max\left\{m : p_{(m)} \leq \frac{m}{M} q\right\} \tag{9}$$

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

5. INDEPENDENCE OF TEST STATISTICS

In the previous work [3], we applied the Benjamini-Hochberg procedure based on the conjecture that the test statistics T_m $(m = 1, \ldots, M)$ are conditionally independent given $(\boldsymbol{I} -$

$\boldsymbol{P}_m(\hat{\boldsymbol{\theta}})$). In the following, we shall show that the test statistics under the null hypotheses H_m are independent. Consequently, the FDR of the multiple test (3) is controlled by the Benjamini Hochberg procedure. The following result from [5] [6] regarding properties of beta distribution plays a key role in our proof.

Result 1. Let $X_1^2, X_2^2, \ldots X_k^2$ be mutually independent with X_j^2 distributed as $\chi^2_{\nu_j}$ with ν_j degrees of freedom $(j = 1, 2, \ldots, k)$. Then

$$\begin{aligned} V_1^2 &= X_1^2/(X_1^2 + X_2^2) \\ V_2^2 &= (X_1^2 + X_2^2)/(X_1^2 + X_2^2 + X_3^2) \\ &\vdots \\ V_{(k-1)}^2 &= (X_1^2 + \ldots + X_{k-1}^2)/(X_1^2 + \ldots + X_k^2) \end{aligned} \tag{10}$$

are mutually independent random variables, each with a beta distribution with parameters (p, q). The values of p, q for V_j^2 are $\frac{1}{2}\sum_{i=1}^{j}\nu_i$, $\frac{1}{2}\nu_{j+1}$ respectively.

Theorem 1. The test statistics T_m, $(m = 1, \ldots, M)$ defined by (5) corresponding to the true null hypotheses are mutually independent.

Proof: We first consider the random variable

$$S_m = \left(\frac{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}]}{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_{m-1}(\hat{\boldsymbol{\theta}}_{m-1}))\hat{\boldsymbol{R}}]}\right),\ (m = 1, \ldots, M) \tag{11}$$

which is related to the test statistic through a monotone function

$$T_m(\hat{\boldsymbol{\theta}}_m) = -\log S_m\ . \tag{12}$$

Now we shall show that under null hypothesis, S_m $(m = 1, \ldots, M)$ are independent beta distributed random variables. By (12), this implies the independence of T_m $(m = 1, \ldots, M)$.

The term $\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}\,]$ appearing in (11) can be decomposed as

$$\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}\,] = \sigma^2(Y_M^2 + Y_{M-1}^2 + \ldots + Y_m^2). \tag{13}$$

where

$$\begin{aligned} Y_m^2 &= \frac{1}{\sigma^2}\mathrm{tr}[(\boldsymbol{P}_{m+1}(\hat{\boldsymbol{\theta}}_{m+1}) - \boldsymbol{P}_m(\hat{\boldsymbol{\theta}}_m))\hat{\boldsymbol{R}}\,], \\ &\qquad\qquad (m = 1, \ldots, (M-1)) \\ Y_M^2 &= \frac{1}{\sigma^2}\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_M(\hat{\boldsymbol{\theta}}_M))\hat{\boldsymbol{R}}\,] \end{aligned} \tag{14}$$

are asymptotically independent and $\chi^2_{\nu_m}$, $(\nu_m \in \mathbb{N})$ distributed under null hypotheses. By (13),

$$S_m = \frac{(Y_M^2 + Y_{M-1}^2 + \ldots + Y_m^2)}{(Y_M^2 + Y_{M-1}^2 + \ldots + Y_m^2 + Y_{m-1}^2)}\ . \tag{15}$$

According to *Result 1*, for independent Y_m^2 $(m = 1, \ldots, M)$ with central $\chi^2_{\nu_m}$ distribution, $S_1, S_2, \ldots, S_M$ are independent beta distributed random variables. The independence of T_m follows immediately. □

6. SIMULATION

In the previous study [3], we compared the FDR- and FWE-controlling procedure for various numbers of signals and different sample sizes. Here, we consider two important cases: 1. Signal sources are closely located. 2. Array output contains only noise. A uniformly linear array of 15 sensors with inter-element spacings of half a wavelength is used in all experiments. Each experiment performs 100 trials.

In the first experiment, the narrow band signals are generated by $m = 12$ sources of equal amplitudes. The number of snapshots is given by $T = 20$. The SNR varies from -8 dB to 8 dB in a 1 dB step. For comparison, the simulated data is applied to the Bonferroni-Holm procedure [4] as well. The sequentially rejective Bonferroni-Holm procedure keeps the FWE at the same level α as the classical Bonferroni test but is more powerful than it. The significance level of each test is given by $\alpha/(M + 1 - m)$. We use $q = 0.1$ and $\alpha = 0.1$ in the simulation. Fig. 1 presents results for 12 widely apart sources. The probability of correct detection $P(\hat{m} = 12)$ increases with increasing SNRs. The FDR-controlling procedure has a lower SNR threshold and a higher probability of detection in the threshold region. In Fig. 2, we show results for a more difficult scenario: two of the signal sources are closely located at $\theta_1 = 9^\circ$ and $\theta_2 = 12^\circ$ relative to broadside. Obviously, both procedures perform worse than in Fig. 1. The gap between FDR- and FWE-controlling procedures has widened in the threshold region. This implies that the FDR-controlling procedure is more useful in critical situations.

In the second experiment, the observation contains only noise. The number of snapshots is varied from $T = 10$ to 30 in a $\Delta T = 5$ step. The maximum number of signals is set to be $M = 4$. From Fig. 3 one can observe that both procedures have probability of correct decision $P(\hat{m} = 0)$ higher than 0.98 for all Ts. Although the FDR q and FWE α are chosen to be 0.1, the detection procedures provide very reliable results.

In summary, the FDR-controlling procedure leads to a higher probability of correct detection than the FWE-controlling procedure. In particular, for situations involving closely located signals, the difference between these two procedures become larger. In the noise only case, the proposed detection scheme has a false alarm rate less than 0.02 for a choice of FDR as high as 0.1.

7. CONCLUSION

This work discusses signal detection using a multiple hypothesis test under an FDR consideration of Benjamini and Hochberg. Compared to the classical FWE criterion, the FDR criterion leads to more powerful tests and controls the errors at a reasonable level. We proved that the conditions required by the Benjamini-Hochberg procedure are satisfied in the proposed detection procedure. Numerical experiments show that the FDR-controlling procedure has always a higher probability of detection than the FWE controlling procedure. More importantly, the reliability of the proposed test is not affected by the gain in power.

8. 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] Johann F. Böhme. Statistical array signal processing of measured sonar and seismic data. In *Proc. SPIE 2563 Advanced Signal Processing Algorithms*, pages 2–20, San Diego, Jul 1995.

[3] Pei-Jung Chung, Johann F. Böhme, Alfred O. Hero, and Christoph F. Mecklenbräuker. Signal detection using a multiple hypothesis test. In *Proc. Third IEEE Sensor Array and Multichannel Signal Processing Workshop*, Barcelona, Spain, July 18-21 2004.

[4] S. Holm. A simple sequentially rejective multiple test procedure. *Scand. J. Statist.*, 6:65–70, 1979.

[5] Norman L. Johnson and Samuel Kotz. *Continuous univariate distributions-2*. John Wiley & Sons, 1970.

[6] Eugene Lukacs. A characterization of the gamma distribution. *Ann. Statist.*, 26(2):319–324, Jun 1955.

[7] 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.

[8] 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.

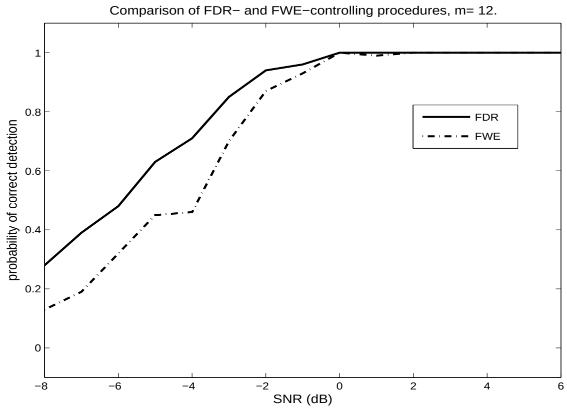

Fig. 1. Probability of correct detection. $M = 12$, SNR $= [-8 : 1 : 8]$ dB, number of snapshots $T = 20$. All sources are apart more than 7° .

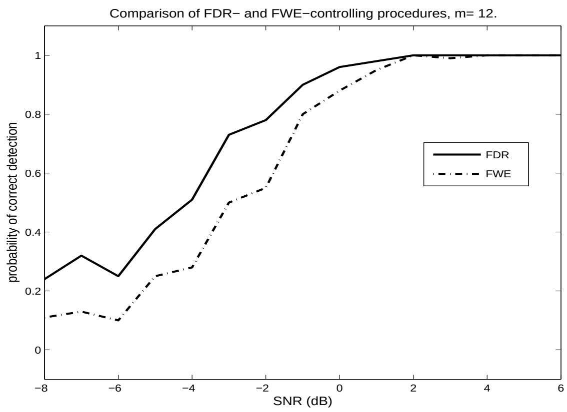

Fig. 2. Probability of correct detection. $M = 12$, SNR $= [-8 : 1 : 8]$ dB, number of snapshots $T = 20$. Two sources are closely located.

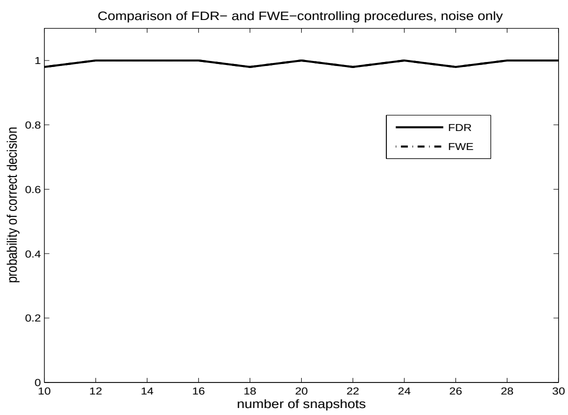

Fig. 3. Probability of correct decision. Noise only. $T = [10 : 5 : 30]$.