ON THE PERFORMANCE OF AN ADAPTATION OF ADICHIE'S RANK TESTS FOR SIGNAL DETECTION: AND ITS RELATIONSHIP TO THE MATCHED FILTER.

Christopher L. Brown, Abdelhak M. Zoubir and Boualem Boashash

Signal Processing Research Centre,
Queensland University of Technology
GPO Box 2434, Brisbane QLD 4001, Australia
Email: c.brown@qut.edu.au

ABSTRACT

The Adichie rank test and signed rank test are adapted for signal detection. We establish a relationship with the correlation between a function of the signal to be detected and the ranks of the observed data. A comparison between the power of these tests and the constant false alarm rate matched filter (CFAR MF) shows that the rank tests perform better when longer observations are available and for the symmetric alpha stable distributions encountered in applications with impulsive interference.

1. INTRODUCTION

In this paper, we consider the problem of choosing the best signal detector in applications where distributions are non-Gaussian. In particular for certain applications involving, for example impulsive interference, we investigate the use of two rank tests formulated by Adichie [1], show their relationship and compare their performance to the constant false alarm rate matched filter (CFAR MF) [5]. A description of these techniques is given in the following sections.

Rank tests have been used when the distribution of the interference is unknown or difficult to identify because they make very weak assumptions [3]. For this reason, their performance for a wide range of non-Gaussian distributions is of significant interest and needs to be evaluated.

We include in this study a number of symmetric α stable distributions. These distributions have been used to characterise impulsive interference, such as those in underwater acoustic signals, low-frequency atmospheric noise and some man-made noise [4]. Except for the Gaussian distribution, which is a limiting case of α stable distributions, they have infinite variance.

1.1. Observation Model

The observation model to be considered may be expressed as

$$X_n = \theta s_n + W_n, \qquad n = 0, \pm 1, \ldots, \tag{1}$$

where θ is the signal strength parameter, s_n is the known, deterministic signal to be detected, and W_n is a stationary random independent and identically distributed (iid) interference process with an unknown, but continuous, distribution and unknown power. When considering a finite number of measurements, $\mathbf{x} = [x_1, \ldots, x_N]$ this model can be re-expressed with N-vector variates,

$$\mathbf{x} = \theta \mathbf{s} + \mathbf{w}.$$

When determining the presence of the known signal, $\mathbf{s}$, in the observed signal, $\mathbf{x}$, we test the hypotheses

$$\mathsf{H} : \theta = 0$$

against

$$\mathsf{K} : \theta > 0 \, .$$

As θ is a measure of the signal strength, it is unnecessary to test for $\theta < 0$.

1.2. CFAR matched filter

A classical solution to detection in the model in equation (1) is provided by the constant false alarm rate matched filter (CFAR MF) [5], expressed as

$$T_N(\mathbf{X}) = \frac{\mathbf{s}\mathbf{P_s}\mathbf{X}^T / \sqrt{\mathbf{s}\mathbf{s}^T}}{\sqrt{\mathbf{X}(\mathbf{I} - \mathbf{P_s})\mathbf{X}^T / (N-1)}}$$

where $\mathbf{P_s} = \mathbf{s}^T\mathbf{s}/\mathbf{s}\mathbf{s}^T$ and $\mathbf{I}$ is the identity matrix. This statistic has a t_{N-1} distribution under H. It can be used to construct a uniformly most powerful (UMP) invariant test for testing H against K in equation (1) when the interference, W_n, is Gaussian distributed with unknown variance. Consequently, we can expect that when non-Gaussian distributions are encountered, more powerful tests will exist.

2. RANK BASED TESTS

2.1. Adichie rank test

Rank based tests have been studied over a long period of time and have shown to be powerful alternatives to classical techniques in circumstances where little is known of the distribution of interference [2, 6, 8]. Adichie [1] proposed a number of techniques applicable when considering hypotheses that test parameters of a linear model. In this paper, we adapt Adichie's test to the signal detection problem.

For the case when we are considering a single signal, s, and iid interference, let us define

$$\begin{aligned}
\mathbf{z} &= [z_1, \ldots, z_N] = [s_1 - \bar{s}, \ldots, s_N - \bar{s}] \\
\mathbf{r} &= [r_1, \ldots, r_N], \text{ where } r_i \text{ is rank of } x_i \\
v &= \sum_{i=1}^{n} z_i \, \psi(r_i) \\
A^2(\psi) &= \int \psi^2(u) du - \left(\int \psi(u) du \right)^2 \\
M &= v^2 / \left(\left(\mathbf{z}\, \mathbf{z}^T \right) A^2(\psi) \right).
\end{aligned}$$

Where $\mathbf{z}$ is signal to be detected, standardised by its mean, $\mathbf{r}$ is the ranks of the observed signal, $\psi(\cdot)$ is a function of the ranks, usually the Wilcoxon scores or Normal scores and M is the test statistic.

Although the procedure is tedious, the exact distribution for M can be calculated for a given $\mathbf{z}$ and N. However, under certain regularity conditions (see [1] for details) it has been shown that, as N increases, asymptotically M has a χ_1^2 distribution. In this paper we shall use this approximation.

If we use the Wilcoxon scores function for ψ, that is $\psi(i) = i/(n+1)$, then $A^2(\psi) = \frac{1}{12}$. It can also be seen that

$$\mathbf{z}\, \mathbf{z}^T = \sum_{i=1}^{N} z_i^2 = \sum_{i=1}^{N} (s_i - \bar{s})^2,$$

and as $\mathbf{s}$ is known, so is $\mathbf{z}\, \mathbf{z}^T$. Therefore

$$M \propto \left(\sum_{i=1}^{N} z_i \, r_i \right)^2 = \left(\mathbf{z}\, \mathbf{r}^T \right)^2. \tag{2}$$

This is the square of the correlation between a function of the signal, $\mathbf{z}$, and the ranks of the data, $\mathbf{r}$. This highlights similarities between this rank based method and other tests based on a correlation function, such as the CFAR MF. In rank based methods, the observed data is ranked, since its distribution is unknown or unspecified. The signal to be detected is not ranked, as it is known and deterministic.

Under H, $\mathbf{s}$ is not present in $\mathbf{x}$, hence s_i and x_i are uncorrelated, as are z_i and r_i. Thus $\mathsf{E}[z_i \, r_i] = \mathsf{E}[z_i]\mathsf{E}[r_i] = 0$, since $\mathsf{E}[z_i] = 0$. As N increases, $\left(\sum_{i=1}^{N} z_i \, r_i \right)$ asymptotically approaches a Gaussian random variable with zero mean, through the Central Limit Theorem, and hence, M approaches a central χ_1^2 variable (the inclusion of the constant terms in the full equation ensures M is appropriately scaled to be a χ_1^2 random variable).

However, under K, s_i and x_i are correlated, and generally, $\mathsf{E}[z_i \, r_i] \neq 0$. Thus M approaches a non-central χ_1^2 variable.

In equation (2) we have seen the mathematical motivation for using this rank test and its similarity, at a basic level, to other correlation based detectors. Using the asymptotic distribution of M we can design a detector based on this test.

2.2. Adichie signed rank test

Adichie also proposed a *signed rank test* for the case when the density distribution of x is symmetric. Let

$$\begin{aligned}
\mathbf{r}^+ &= [r_1^+, \ldots, r_N^+], \text{where } r_i^+ \text{is the rank of } |x_i| \\
\mathsf{sgn}(x) &= \begin{cases} +1, & x > 0 \\ -1, & x < 0 \end{cases} \\
v^+ &= \sum_{i=1}^{N} s_i \, \phi(r_i^+) \, \mathsf{sgn}(x_i) \\
M^+ &= (v^+)^2 / \left(\left(\mathbf{s}\, \mathbf{s}^T \right) A^2(\bar{\phi}) \right).
\end{aligned}$$

Where $\mathbf{r}^+$ is the ranks of the magnitude of the observed signal, and $\phi(\cdot)$ is a function similar to the scoring function $\psi(\cdot)$.

Again, we note that the statistic M^+ is proportional to the square of the correlation between the signal, $\mathbf{s}$, and a function of the ranks of the observations, in this case, $\phi(r_i^+) \, \mathsf{sgn}(x_i)$. Under similar conditions mentioned in section 2.1, the asymptotic distribution of M^+ is χ_1^2.

3. RESULTS AND DISCUSSION

In order to provide guidelines for the use of the proposed method, we have performed computer simulations to estimate the power of the tests being considered; using a number of different distributions. The signal used in all cases was a sinusoid. All results were generated by running 1000 simulations for each test at a significance level of $\alpha = 5\%$.

In order to determine the power of the tests based on the CFAR MF, Adichie rank test (Adi. I) and signed rank test (Adi. II) under a range of scenarios, four different operating conditions have been simulated and are shown in Tables 1 – 6. Note that for the two shorter sequence lengths of $N = 10$ and $N = 50$, results for a detection scheme based on the bootstrap [7] have also been included.

- $N = 10$ observations, signal to noise ratio (SNR)=0dB
- $N = 50$ observations, SNR=-9dB
- $N = 100$ observations, SNR=-11dB.

The SNRs have been chosen so the tests considered achieve power of around 80% to 95%. It would be unrealistic to compare estimated power levels that are all near 100%. Similarly, no information is obtained by comparing tests at power levels where all tests perform poorly.

The most significant observation to be made from inspecting the tabulated data is that the rank tests appear to perform better than the CFAR MF for the symmetric α stable (SαS) distributions tested – i.e. Cauchy, SαS($\alpha = 1.5$), SαS($\alpha = 0.75$) and SαS($\alpha = 0.25$). The only exception is the Gaussian distribution which is a limiting case of α stable distributions. Rank tests, by their nature, are unaffected by scaling, and this may have increased their power, compared to other methods, for the heavy-tailed α stable distributions.

The CFAR MF outperforms the rank tests for short sequence lengths, even for non-Gaussian distributions. As stated in section 2.1, Adichie's test statistic *asymptotically* has a χ_1^2 distribution. This approximation becomes more accurate for longer sequence lengths and therefore the power of the rank tests increases, relative to the CFAR MF.

This problem may be reduced by calculating the exact distribution of the statistic – as mentioned previously. It may also explain the fact that the significance level appears to be underestimated for the shorter sequences. Additionally, it is often the case with rank tests that longer sequences assist them overcoming problems caused by the "discreteness" of ranks.

Generally speaking, when few observations are available, techniques such as the bootstrap have been shown [7] to have comparable or superior performance to the CFAR MF. For longer sequences, such as $N = 100$ the computational burden of the bootstrap makes it impractical.

The significance level of all tests was maintained, except for Adichie's signed rank test for longer sequences in the case of a number of distributions, primarily the sum of Gaussian and α stable distributions.

The effect of varying the sequence length, while maintaining the SNR at -9dB, when using Gaussian interference, is shown in Figure 1. It should be evident that the difference in the number of samples required to achieve a certain detection power, as a proportion of the total number of samples, decreases for larger N.

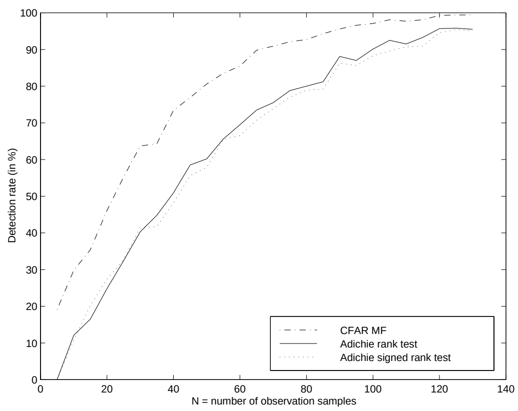

Figure 1: Detection rates of tests at SNR$= -9$dB and Gaussian interference

Figure 2 shows the degradation in performance of the rank tests and the CFAR MF as the SNR is reduced, while maintaining a constant number of data points of $N = 50$ (the false alarm rate was maintained by all tests). The interference was Gaussian distributed.

These results show that the rank based methods' results at around 2dB higher SNR are comparable to the CFAR MF's, remembering that the CFAR MF is the optimal detector for additive Gaussian interference of unknown power.

4. CONCLUSIONS

Adichie's rank tests can be viewed as a correlator detector utilising rank information when exact distributional properties of the observed signal are unavailable. It more powerful than the classical CFAR MF against the symmetric α stable distributions tested and against a number of non-Gaussian distributions when long observations are available. This appears to be a significant result. As the number of observations is decreased, the rank tests' performance deteriorates so that the CFAR MF detector becomes more powerful.

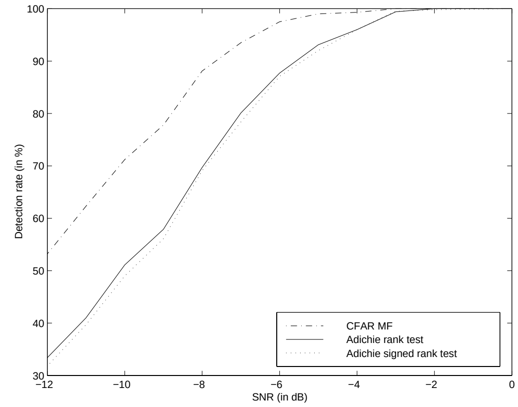

Figure 2: Detection rates of tests for $N = 50$ and a Gaussian distribution.

5. ACKNOWLEDGEMENTS

The authors would like to thank Mr Hwa-Tung Ong for his assistance in running bootstrap simulations.

6. REFERENCES

[1] J. N. Adichie. Rank tests in linear models. In P. R. Krishnaiah and P. K. Sen, editors, *Nonparametric Methods*, volume 4 of *Handbook of Statistics*, chapter 11, pages 229–257. Elsevier Science Publishers, Amsterdam, The Netherlands, 1984.

[2] C. L. Brown and A. M. Zoubir. Rank tests in signal detection. (under review).

[3] E. L. Lehmann. *Nonparametrics: Statistical Methods Based on Ranks.* Holden-Day, San Francisco, 1975.

[4] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications.* John Wiley & Sons, 1995.

[5] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* Addison-Wesley, 1991.

[6] A. M. Zoubir. Statistical signal processing for application to over-the-horizon radar. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 94*, volume VI, pages 113–116, Adelaide, Australia, 1994.

[7] A. M. Zoubir. Signal detection using the bootstrap. *Signal Processing*, (under review). (Invited).

[8] A. M. Zoubir, T. Grewing, and J. F. Böhme. Tests for optimizing sensor positions using linear regression. In *Signal Processing VI: Theories and Applications. Proceedings of EUSIPCO-92, Sixth European Signal Processing Conference*, volume 2, pages 641–644, Brussels, Belgium, August 1992.

Dist[1].	CFAR MF	Adi. I	Adi. II	Boot.
$N(0,1)$	4.1	4	2.8	4.5
$U(0,1)$	5.7	3.9	4.3	5.6
Laplace	4.1	3.3	3.3	4.3
Log Normal	3.9	3	2.4	4.3
t_2	4.5	2.6	2.1	3.8
t_3	5	4.6	4	5.8
t_8	5.1	3	2.5	5
$\sum_G^I$	3.2	4.1	4.1	3.6
$\sum_G^{II}$	5.8	3.1	4.3	4.3
Cauchy	4.5	3.9	4	4.7
SαS($\alpha = 1.5$)	6	5	3	5
SαS($\alpha = 0.25$)	0	4	0	2
SαS($\alpha = 0.75$)	5.4	3.4	2.5	4.1

Table 1: False alarm rates (in %) for $N = 10$, SNR=0dB.

Dist.	CFAR MF	Adi. I	Adi. II	Boot.
$N(0,1)$	87.9	61.1	55.5	84.3
$U(0,1)$	91	51.1	52.6	86.9
Laplace	86.2	66.3	60.8	84.6
Log Normal	92.6	86.4	88	91.9
t_2	93.1	91	87.3	92.8
t_3	87.8	75.4	68.7	86.8
t_8	84.8	60.8	56.7	82.6
$\sum_G^I$	88.5	70.8	73.7	86.4
$\sum_G^{II}$	88.2	64.8	58.8	81.6
Cauchy	99.8	99.8	99.7	99.5
SαS($\alpha = 1.5$)	95	94	91	95
SαS($\alpha = 0.25$)	95	100	100	100
SαS($\alpha = 0.75$)	99.4	99.8	99.4	99.6

Table 2: Detection rates (in %) for $N = 10$, SNR=0dB.

Dist.	CFAR MF	Adi. I	Adi. II	Boot.
$N(0,1)$	4.8	4.8	4.9	5
$U(0,1)$	5.3	6	4.7	4.9
Laplace	5.4	5.2	5.7	5.5
Log Normal	3.3	5.4	8.9	4.4
t_2	5.6	6	5.8	6.3
t_3	4.5	5.1	4.9	4
t_8	5	3.3	4.7	4.9
$\sum_G^I$	4.5	5.9	7.1	4.4
$\sum_G^{II}$	4.3	5.9	6.2	3.5
Cauchy	4	4.8	13.7	3.7
SαS($\alpha = 1.5$)	4	5.1	6	5
SαS($\alpha = 0.25$)	2.5	4	41.8	2.2
SαS($\alpha = 0.75$)	3.2	4.7	25.5	4.1

Table 3: False alarm rates (in %) for $N = 50$, SNR= -9dB.

[1] The abreviations used for the distributions tested were: $N(0,1)$ – Gaussian, $U(0,1)$ – Uniform, SαS – symmetric α stable

Dist.	CFAR MF	Adi. I	Adi. II	Boot.
$N(0,1)$	80.6	60.2	57.9	77.6
$U(0,1)$	80.7	55.7	41.3	77.9
Laplace	79.8	76.5	81.9	78.3
Log Normal	86.2	98.7	78.5	87.3
t_2	93	99.9	99.9	92.3
t_3	80.5	83.1	85.7	78.2
t_8	81.7	69.2	68.2	80.1
$\sum_G^I$	81.6	87.8	73.5	80.2
$\sum_G^{II}$	78.6	76.7	78.3	75.1
Cauchy	96.5	100	99.5	95.7
SαS($\alpha = 1.5$)	93.7	100	98.8	92.6
SαS($\alpha = 0.25$)	96.9	100	98.7	96.5
SαS($\alpha = 0.75$)	97.3	100	98.8	95.6

Table 4: Detection rates (in %) for $N = 50$, SNR= -9dB.

Dist.	CFAR MF	Adi. I	Adi. II
$N(0,1)$	5.3	5.1	4.3
$U(0,1)$	4.6	4.4	3.9
Laplace	4.6	4.8	4.3
Log Normal	3.9	5.4	15.1
t_2	5	4.4	6.4
t_3	4.3	5.7	5.2
t_8	3.9	4	3.5
$\sum_G^I$	4.1	4.5	10.3
$\sum_G^{II}$	5.7	4.8	7
Cauchy	4.4	4.4	33.9
SαS($\alpha = 1.5$)	3.1	4.3	9.5
SαS($\alpha = 0.25$)	3.3	4.4	98.1
SαS($\alpha = 0.75$)	4.1	4.7	59.4

Table 5: False alarm rates (in %) for $N = 100$, SNR=-11dB.

Dist.	CFAR MF	Adi. I	Adi. II
$N(0,1)$	89.6	74	70.4
$U(0,1)$	88.7	70.6	55.2
Laplace	88.2	87	91.4
Log Normal	89.8	99.8	78.8
t_2	95.4	100	99.7
t_3	90	93.5	95.3
t_8	89.7	79.1	77.9
$\sum_G^I$	88.9	96.5	77.3
$\sum_G^{II}$	87.7	88.3	92.3
Cauchy	98.4	100	99.4
SαS($\alpha = 1.5$)	94	100	98.7
SαS($\alpha = 0.25$)	98.1	100	98.8
SαS($\alpha = 0.75$)	95.9	100	98.8

Table 6: Detection rates (in %) for $N = 100$, SNR= -11dB.

$\sum_G^I = 0.5N(\frac{-1}{10}, \frac{1}{16}) + 0.1N(\frac{1}{5}, \frac{4}{25}) + 0.2N(\frac{1}{2}, \frac{1}{25}) + 0.2N(1, \frac{1}{25})$,
$\sum_G^{II} = 0.2N(\frac{-3}{10}, \frac{1}{4}) + 0.5N(\frac{7}{10}, \frac{1}{16}) + 0.3N(\frac{4}{5}, \frac{4}{25})$