RANK ORDER DIVERSITY DETECTORS FOR WIRELESS COMMUNICATIONS

Alexander Flaig and Gonzalo R. Arce

Department of Electrical and Computer Engineering
University of Delaware, Newark, DE 19716
flaig@ece.udel.edu

ABSTRACT

A novel class of robust detectors, the *rank order diversity (ROD) detectors*, is proposed. The ROD detectors exploit the diversity inherent in any repetition code by sorting the sampled channel output, weighting each of the sorted samples according to their rank, and summing the weighted order statistics to form the test statistic. The ROD detectors subsume the globally optimal detectors for the short-tailed uniform and the normal distributions, as well as the locally optimal detector for the heavy-tailed double-exponential distribution, suggesting a high efficiency of the ROD detectors over a vast range of possible noises statistics. It is shown that for large sample sizes and under mild conditions on the noise statistics, the ROD detector achieves a probability of error less or equal than that of the linear detector with equality only for the normal distribution. The performance of the ROD detectors is illustrated in a DS-CDMA network.

Index Terms: robust detection, wireless communications, rank order diversity, order statistics

1. INTRODUCTION

The design of detection schemes for spread spectrum multiple access (SSMA) networks in which security restrictions do not permit the distribution of all users signaling parameters is a formidable task, as the multiple access interference (MAI) has a non-Gaussian distribution with an exact shape that depends on the received power of each of the active users in the network. Variations in the users power due to roaming users, the advent or departure of users, and/or imperfect power control cause the statistics of the MAI to change rapidly and to assume vastly different characteristics ranging from near-Gaussian, over multi-modal, to heavy-tailed.

Under these conditions, detectors that are optimized for a specific distribution, including maximum-likelihood and multi-user detectors, are clearly not applicable as they suffer a drastic degradation in performance even for apparently small deviations from the nominal assumptions. While many of the classical robust detection schemes, including minimax detectors and non-parametric detectors, can offer an acceptable performance over a *limited* class of possible noise statistics [1], they become inefficient when the uncertainty about the noise distribution is large, as is the case in spread spectrum networks. Adaptive detectors who learn about the noise statistics and adjust their signal processing structure accordingly have the potential to maintain a good performance for *all* possible noise statistics.

Along this line, we introduce a novel adaptive detector for the robust detection of binary antipodal signals in non-Gaussian noise. The proposed *rank order diversity (ROD) detectors* employ a linear combination of rank order statistics (L-statistic) – samples rearranged in increasing order of their values – to test for the polarity of the transmitted signal. The weights of the test statistic are adjusted to the prevailing noise characteristics based on the second-order moments of the order statistics of the noise. The test statistic is drawn from the class of *robust* estimates in mathematical statistics where it was shown [2] that its variance when employed as a location estimate is strictly less than that of the sample mean for all symmetric distributions except the Gaussian. We show that this property translates, asymptotically as the number of samples goes to infinity and under mild assumptions on the noise distribution, in a probability of error of the ROD detector that is less than or equal to that of the linear detector with equality only for Gaussian noise.

We illustrate the potential of the ROD detectors in a DS-CDMA network in which users are allowed to depart and new users, whose parameters are not known, are permitted to join the system. The applications of the ROD detectors, however, are not limited to spread spectrum networks. In particular, it is shown in this paper that exploiting the rank order diversity can improve the capacity of *any* system in which linear detection schemes are employed despite prevailing non-Gaussian noise statistics.

2. OPTIMAL DETECTION AND RANK ORDER DIVERSITY DETECTION

Consider the coherent detection of a constant binary antipodal signal in additive white noise [1]. Denoting the observations from the noisy signal by X_i, this problem can be put in the framework of a binary hypothesis test:

$$\begin{array}{l} H_{+1}: X_i = +1 + Z_i \\ H_{-1}: X_i = -1 + Z_i, \quad i = 1, 2, \ldots, N, \end{array} \tag{1}$$

where the Z_i are assumed to be independent noise samples with a common univariate zero-mean probability density function (pdf) of the form $f(x) = f(x/\sigma)/\sigma$, where $\sigma > 0$ is a scale parameter. The optimum test for H_{+1} versus H_{-1} in the sense of minimizing the probability of an erroneous decision, i.e. choosing H_{+1} when $b = -1$ or H_{-1} when $b = +1$, can be put in the form

$$T_{opt} = \sum_{i=1}^{N} w_{opt}(X_i) X_i \left\{ \begin{array}{ll} > 0 & \Rightarrow H_{+1} \\ < 0 & \Rightarrow H_{-1}, \end{array} \right. , \tag{2}$$

[1]Note that restricting the signal to be constant does not result in a loss of generality due to the assumption of white noise.

This research was supported through collaborative participation in the Advanced Telecommunications/Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0002.

where $w_{opt}(x) = \ln[f(x-1)/f(x+1)]/x$, to stress the fact that each *sample* is weighted according to its *value* by a weight function that depends on the exact density of the noise in a critical way. The test statistic of the ROD detector achieves robustness by weighting each sample according to the *sample's rank*, or its value relative to those of the remaining samples, and sums the weighted samples.

Definition: Let $X_{(1)} \le X_{(2)} \le \ldots \le X_{(N)}$ be the order statistics (OS) obtained by rearranging the observations in (1) according to their values, then the ROD detector employs

$$T_{rod} = \sum_{i=1}^{N} w_i X_{(i)}, \tag{3}$$

to test for H_{+1} versus H_{-1}, where the w_i are real-valued weights. □

The ROD detector can be shown to be *globally optimal* for the *uniform* and the *Gaussian* distributions, for which T_{rod} reduces to the mid-range $T_{mr} = X_{(1)} + X_{(N)}$ and to the linear test statistic $T_{lin} = \sum_{i=1}^{N} X_i/N$, respectively. For the *double-exponential* noise density and a weak signal, T_{rod} reduces to the *sample median*, the test statistic of the *locally optimal detector*. Note, that in all three cases the ROD detector uses the *maximum-likelihood location estimator* as a test statistic.

The following Theorem, proven in [3], suggests that the minimization of the probability of error of the ROD detector for a symmetric but otherwise arbitrary distribution is equivalent to finding those weights that maximize the distribution of the L-statistic of the *noise* samples in the interval $(-1, +1)$ subject to a location invariance constraint.

Theorem: Denote the vectors of the OS of the noise and the corresponding weights by $\mathbf{z} = [Z_{(1)}, Z_{(2)}, \ldots, Z_{(N)}]^T$, and $\mathbf{w} = [w_1, w_2, \ldots, w_N]^T$, respectively, and let $Y(\mathbf{z};\mathbf{w}) = \mathbf{w}^T\mathbf{z}$, then the weight vector that minimizes the probability of error of the ROD detector is

$$\mathbf{w}_o = \arg\max_{\mathbf{w}} P\{|Y(\mathbf{z};\mathbf{w})| < 1\} \tag{4}$$

subject to $\mathbf{w}^T\mathbf{e} = 1$, where $\mathbf{e}$ is a N-long vector of ones.

Except when $w_i = 0$ for all i but one or $w_i = c$ for all i, a closed-form expression for $P\{|Y| < 1\}$ is difficult to obtain [3], obstructing the derivation of a closed-form solution for the optimal weights. Asymptotically, however, as $N \to \infty$ the distribution of Y converges to a normal distribution for sufficiently smooth weights [3]. Thus it suffices to minimize the variance of Y to find the *asymptotically optimal weights* $\mathbf{w}_o = \mathbf{e}^T\mathbf{Q}/\mathbf{e}^T\mathbf{Q}\mathbf{e}$, where $\mathbf{Q}$ is the inverse covariance matrix of $\mathbf{z}$ [3]. The L statistic with the weights $\mathbf{w}_o$ is known as the best linear unbiased estimator (BLUE) of location [2], i.e. the

$$Var(Y(\mathbf{z};\mathbf{w}_o)) \le \sigma^2, \tag{5}$$

with equality if and only if the noise distribution is Gaussian. The inequality (5) implies that the ROD detector will achieve a *probability of error that is asymptotically less or equal than that of the linear detector* whenever $Y(\mathbf{z};\mathbf{w}_o)$ is normally distributed. It can be shown [3] that this is the case if there exists some constant K such that $lim_{N\to\infty} N(N+1)\Delta_N = K$, where $\Delta_N = \max_{i=1,2,..,N-1} |w_o(i+1) - w_o(i)|$.

The following proposition, which follows directly from the definition of the statistical expectation, becomes useful in solving *adaptively* for the optimal weights for finite N.

Proposition: Maximizing the $P\{|\mathbf{w}^T\mathbf{z}| < 1\}$ with respect to $\mathbf{w}$ is equivalent to minimizing the risk $E\{g(\mathbf{w}^T\mathbf{z})\}$, where $g(y) = 0$ if $y < 1$ and $g(y) = 1$ otherwise.

Unfortunately, g is discontinuous, making the risk function difficult to minimize. Approximating g by $g'(y) = 1 - e^{-y^2}$, we can formulate the optimization problem as follows: The weights of the order statistic detector satisfy

$$\mathbf{w}_o = \arg\max_{\mathbf{w}} E\{1 - e^{-(\mathbf{w}^T\mathbf{x})^2}\} \tag{6}$$

subject to $\mathbf{w}^T\mathbf{e} = 1$. This constrained optimization problem has a unique solution due to the convexity of g', permitting a gradient-based solution. The problem of constraint adaptive optimization of L-filters has been solved previously for a mean-square error cost function in [4]. In this work, the authors developed an adaptation procedure that uses the location invariance constraint to eliminate the coefficient with the slowest convergence rate (the coefficient that corresponds to the median). This technique can be extended to the problem at hand. Define the reduced weight vectors

$$\mathbf{w}_1 = [w_1, \ldots, w_{\frac{N-1}{2}}]^T \text{ and } \mathbf{w}_2 = [w_{\frac{N+3}{2}} \ldots, w_N]^T, \tag{7}$$

and similarly

$$\mathbf{x}_1 = [X_{(1)}, \ldots, X_{(\frac{N-1}{2})}]^T \text{ and } \mathbf{x}_2 = [X_{(\frac{N+3}{2})}, \ldots, X_{(N)}]^T, \tag{8}$$

and note that

$$\mathbf{w} = [\mathbf{w}_1^T, 1 - \mathbf{e}^T\mathbf{w}_1 - \mathbf{e}^T\mathbf{w}_2, \mathbf{w}_2^T]^T, \tag{9}$$

and $\mathbf{x} = [\mathbf{x}_1, X_{(\frac{N+1}{2})}, \mathbf{x}_2]$. Using the above, we can rewrite

$$(\mathbf{w}^T\mathbf{x})^2 = \mathbf{w}^T\mathbf{R}\mathbf{w}, \tag{10}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{r}_1 & \mathbf{R}_2 \\ \mathbf{r}_1^T & r & \mathbf{r}_2^T \\ \mathbf{R}_3 & \mathbf{r}_2 & \mathbf{R}_4 \end{bmatrix}, \tag{11}$$

where $\mathbf{R}_i = \mathbf{x}_i\mathbf{x}_i^T$, $i = 1,2,3,4$, $\mathbf{r}_i = X_{((M+1)/2)}\mathbf{x}_i$, $i = 1, 2$, and $r = X_{((M+1)/2)}{}^2$. Using the short weight vectors (7), equation (10) can be simplified further:

$$(\mathbf{w}^T\mathbf{x})^2 = r - 2\mathbf{v}^T\mathbf{p} + \mathbf{v}^T\mathbf{R}'\mathbf{v}, \tag{12}$$

where $\mathbf{v} = [\mathbf{w}_1^T, \mathbf{w}_2^T]^T$, $\mathbf{p} = [r\mathbf{e}^T - \mathbf{r}_1^T, r\mathbf{e}^T - \mathbf{r}_2^T]^T$, and

$$\mathbf{R}' = \begin{bmatrix} \mathbf{R}_1 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_1\mathbf{e}^T & \mathbf{R}_2 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_1\mathbf{e}^T \\ \mathbf{R}_3 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_2\mathbf{e}^T & \mathbf{R}_4 + r\mathbf{e}\mathbf{e}^T - 2\mathbf{r}_2\mathbf{e}^T \end{bmatrix}. \tag{13}$$

Using (12), the optimization problem (6) can be stated in an *unconstrained* form in terms of the reduced weight vector:

$$\mathbf{v}_o = \arg\max_{\mathbf{v}} E\{1 - e^{-(r - 2\mathbf{v}^T\mathbf{p} + \mathbf{v}^T\mathbf{R}'\mathbf{v})}\}. \tag{14}$$

To solve for $\mathbf{v}_o$, we need the gradient of the risk function $J = E\{1 - e^{-(r-2\mathbf{v}^T\mathbf{p}+\mathbf{v}^T\mathbf{R}'\mathbf{v})}\}$ with respect to $\mathbf{v}$. Straightforward algebra yields

$$\frac{\partial J}{\partial \mathbf{v}} = -E(2\mathbf{p} - (\mathbf{R}' + \mathbf{R}'^T)\mathbf{v})e^{(-(r-2\mathbf{v}^T\mathbf{p}+\mathbf{v}^T\mathbf{R}'\mathbf{v}))}. \quad (15)$$

Using instantaneous estimates, we obtain the following iterative algorithm:

Definition: The reduced weight vector at iteration n is given by

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \mu(\mathbf{p} - \mathbf{R}_s\mathbf{v})e^{(-(r-2\mathbf{v}^T\mathbf{p}+\mathbf{v}^T\mathbf{R}'\mathbf{v}))}(n), \quad (16)$$

where $\mathbf{R}_s = (\mathbf{R}' + \mathbf{R}'^T)/2$ and μ is a positive step-size. □

Using (16), the full set of weights can be easily computed from (9) and (7). Note that the update term in (16) is a *nonlinear* function of the instantaneous estimates of the *second moments of the order statistics of the noise*. Necessary conditions for the convergence of (16) under assumptions similar to the *independence assumptions* made in the analysis of linear filters can be found in [3].

3. ASYMPTOTIC PERFORMANCE

A useful measure of comparison of two detectors under large-sample-size conditions is the *asymptotic efficiency*. The asymptotic efficiency of detector A when compared to detector B, denoted by $ARE_{A,B}$, is defined as the ratio of the number of samples needed by detectors A and B, respectively, to achieve some fixed probability of error as the sample size N goes to infinity. Under certain regularity conditions the $ARE_{A,B}$ can be expressed as the ratio

$$ARE_{A,B} = \frac{\varepsilon_A}{\varepsilon_B}, \quad (17)$$

where ε_T is the *efficacy* of a detector using test statistic T [5]. The following Theorem, proven in [3], gives the efficacy of the ROD detector with sufficiently smooth weights.

Theorem: Suppose the weights of the ROD detector can be put in the form $w_i = J(\frac{i}{N+1})/N$, where $J(u)$, $0 \leq u \leq 1$, is an associated weight function, then the efficacy of the ROD detector operating in noise with a density $f = F'$ is given by

$$\varepsilon_{rod} = 1/\sigma^2(J, F), \quad (18)$$

where $\sigma^2(J,F) = 2\int\int_{-\infty<x<y<+\infty} J(F(x))J((F(y))(F(x)(1-F(y)))dxdy$. In the special cases where the ROD detector reduces to the linear and the median detector, we have $\varepsilon_{lin} = 1/\sigma^2$ and $\varepsilon_{med} = 4f^2(0)$, respectively.

4. APPLICATION IN A DS-CDMA NETWORK

Consider a wireless DS-CDMA network in which K terminals are uniformly distributed within a circular cell around a centrally located receiver. Figure 1 shows a typical realization of such a system for $K = 8(9)$ users. Assuming antipodal modulation, random binary spreading codes consisting of N chips, and users that are

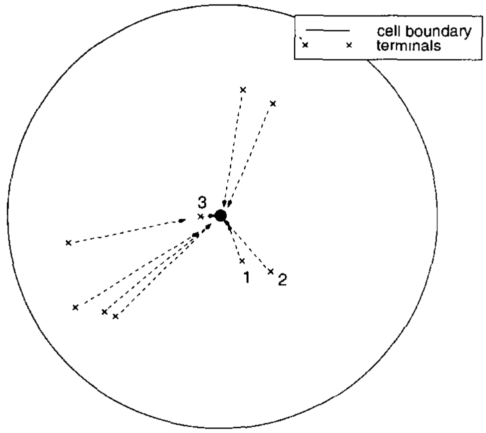

Figure 1: *Typical user distribution.*

synchronized at the chip level, the noise samples in (1) can be shown to be put in the form

$$Z_i = \xi_i + \eta_i, \quad (19)$$

where ξ_i is the multinomially distributed MAI component and η_i is a zero-mean Gaussian random variable with variance σ^2. Consequently, the density function of the compound noise is continuous with, in general, multiple modes.

Figures 2(a)-(c) compare the performance of the linear detector, the mid-range detector, the ROD detector, the optimal (single-user) detector, and the multiuser MMSE detector when employed by the two strongest users in the system (from the receivers point of view), labeled as 1 and 2 in Fig. 1. The optimal coefficients of the ROD detector are obtained by using algorithm (16), where, for simplicity, the optimization is performed for the hypothetical situation in which the background noise vanishes. The nonlinearity of the optimal single-user detector, g_{opt}, and the coefficients of the MMSE detector can be found in [3].

Consider user 1 (Fig. 2(a)) first. The linear detector, the single-user maximum-likelihood (ML) detector, and the multi-user MMSE detector achieve identical bit error rates due to the fact that the MAI seen by this user is a superposition of comparatively weak signals assuming a near-Gaussian distribution. Due to the mismatch between the prevailing noise distribution (MAI + background noise) and the one assumed during optimization, the ROD detector requires about $1dB$ more A/σ to achieve the same bit error rate than the previous three detectors. The noise seen by the second user (Fig. 2(b)) has a bimodal distribution due to the (stronger) signal of user 1, causing a mild near-far problem for the linear detector. The bit error rates of the MMSE, the single-user ML, and the ROD detectors are increased by approximately a factor of ten for $A_2/\sigma = 6dB$ compared to those of user 1 for the same SNR. Note, that the ROD detector is near-far resilient and achieves essentially the same bit error rate as the optimal single-user detector.

The good performance of the MMSE detector and the optimal single-user detector can, in general, not be retained if some of the system parameters change. To illustrate this, we add one more user (jammer), labeled by a 3 in Fig. 1, whose parameters are unknown to any one of the employed detectors. The coefficients of the above detectors are held fixed, i.e. a re-training time is not permitted which would enable the ROD detector and a blind

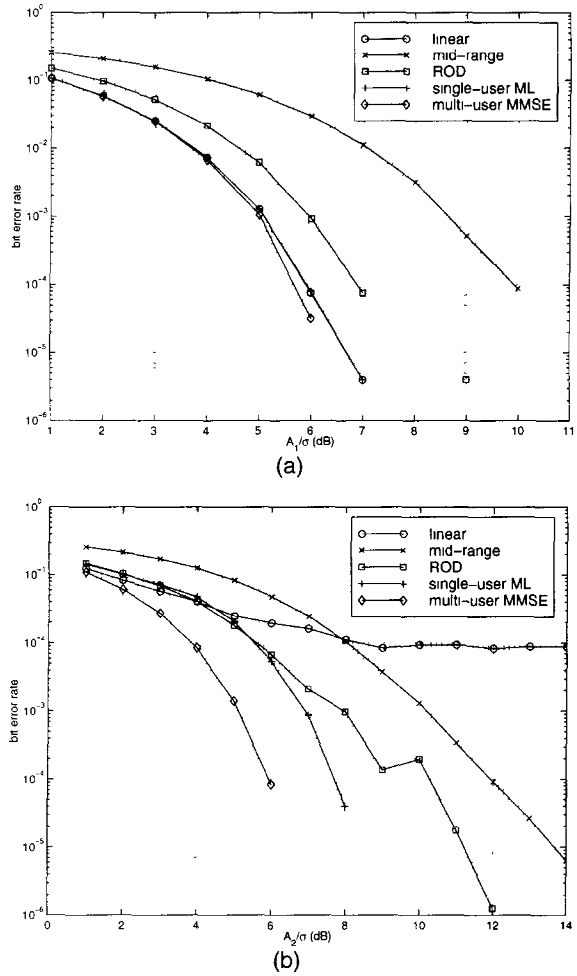

Figure 2: *Users* 1 *and* 2.

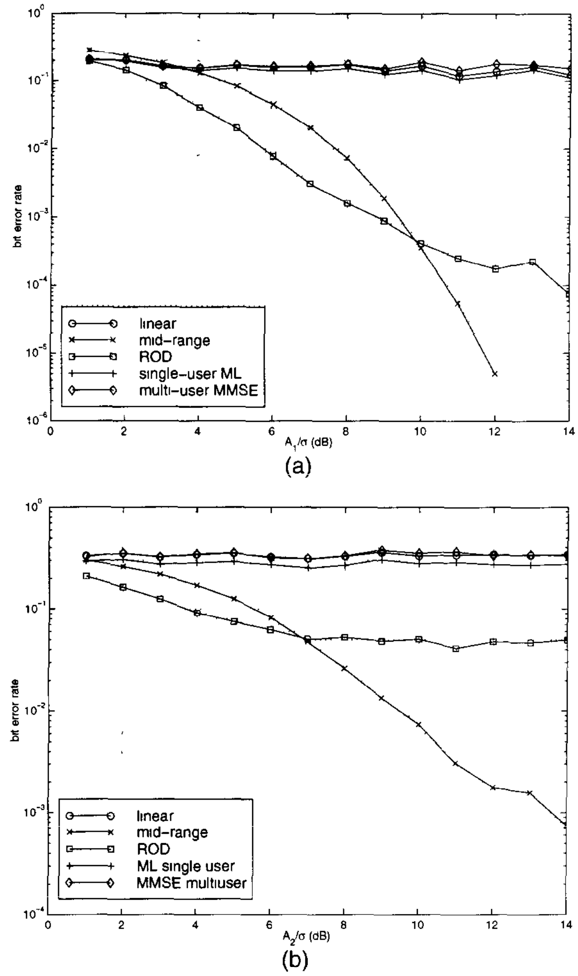

Figure 3: *User* 1 *and* 2 *(jammed).*

version of the MMSE detector, for instance, to learn the jammers signature. Figures 3(a)-(b) show the detector performance in the presence of the jammer for user 1 and 2, respectively. The results are averaged over 250 randomly chosen signature sequences for the jammer. As expected, neither the optimal single-user detector nor the MMSE detector are robust to the advent of the new user. In fact, the decision of the MMSE detector is governed by the jammer due to a relatively high average cross-correlation between the jammers signature sequence and the coefficients of the MMSE detector. Both, the ROD detector (optimized *prior* to the advent of the new user) as well as the mid-range detector proof themselves robust to the arrival of the new user. The optimized ROD detector looses some performance over the mid-range detector in high SNR as the distribution of the MAI has become multi-modal due to the advent of the jammer. Letting the weights adjust, it will, of course, regain a performance level that is higher than that of the mid-range detector.

Summarizing, it was found that (1) the ROD detector possesses the ability to *mitigate the near-far problem* – this crucial improvement over the conventional detector can be directly attributed to the fact that the order statistics are insensitive to scale, making the test statistic robust against power variations of the interfering users, and (2) the ROD detector is *less sensitive to changes of system parameters*, such as the arrival or departure of a user, than the MMSE multiuser detector. This robustness is achieved by exploiting only the second moments of the order statistics of the noise as opposed to relying on detailed information about the noise statistics.

5. REFERENCES

[1] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: a survey," *Proceedings of the IEEE*, vol. 73, pp. 433–481, Mar. 1985.

[2] L. Bondesson, "When is the sample mean BLUE," *Scand. J. Statist.*, vol. 3, pp. 116–120, 1976.

[3] A. Flaig and G. R. Arce, "Rank order diversity detectors for wireless communications," *IEEE Transactions on Information Theory*, 1998. to be submitted.

[4] C. Kotropoulos and I. Pitas, "Constrained adaptive LMS *L*-filters," *Signal Processing*, vol. 26, March 1992.

[5] S. A. Kassam, *Signal detection in non-Gaussian noise*. Springer, 1988.