ROBUST MINIMUM DISTANCE NEYMAN-PEARSON DETECTION OF A WEAK SIGNAL IN NON-GAUSSIAN NOISE

Georgy Shevlyakov, Kyungmin Lee, Vladimir Shin and Kiseon Kim

School of Information and Mechatronics, GIST
1 Oryong-dong, Buk-gu, Gwangju 500-712, Korea
Email: {shev, kmlee, vishin, kskim}@gist.ac.kr

ABSTRACT

In practice, noise distributions usually are not Gaussian and may vary in a wide range from light-tailed to heavy-tailed forms. To provide robust detection of a weak signal, a maximin in the Huber sense Neyman-Pearson detector based on the minimum distance between the signal and observations is designed. Explicit formulas for the power of detection and the false-alarm probability are derived.The maximin detectors are written out for the classes of nondegenerate, with a bounded variance and contaminated Gaussian noise distributions along with some numerical results on their performance.

1. INTRODUCTION

Consider the problem of detection of a known signal $\{s_i\}_1^N$ in the additive i.i.d. noise $\{n_i\}_1^N$ with a symmetric pdf f from a class $\mathscr{F}$. Given $\{x_i\}_1^N$, it is necessary to decide whether the signal $\{s_i\}_1^N$ is observed. The problem of detection is set up as the problem of hypothesis testing

$$H_0: x_i = n_i \quad \text{versus} \quad H_1: x_i = \theta\, s_i + n_i, \quad i = 1,\ldots,N,$$

where the positive signal amplitude θ is assumed known. Given a pdf f, the classical theory of hypotheses testing yields various optimal (in the Bayesian, minimax, Neyman-Pearson senses) decision rules: all the optimal rules are based on the likelihood ratio (LR) statistic $T_N(\mathbf{x}) = \prod_i^N f(x_i - \theta\, s_i)/f(x_i)$ that should be compared with a certain threshold. The differences between the aforementioned approaches result only in the values of a threshold.

In many practical problems of signal processing, noise distributions are only partially known. For instance, it may be assumed that either noise is approximately Gaussian, or there is some information on its pdf behavior in the central zone and on the tails, on its moments and subranges, etc. In his seminal works on robust estimation and hypothesis testing, namely, in [4] and [5], Huber considers the classes $\mathscr{F}$ of allowable noise pdfs as the neighborhoods of nominal densities and applies minimax approach to design robust M-estimators of location and robust Neyman-Pearson tests using the maximum likelihood method for the least favorable noise distribution densities f^* in the aforementioned classes for the null and alternative hypotheses. In these cases, the optimal robust statistics have the structures of the bounded likelihood and of the bounded likelihood ratio, respectively. Further, both those results are used in robust detection. In [13], Huber's results on robust hypothesis testing are adapted for robust detection of a known signal in contaminated Gaussian noise. Next, Huber's minimax approach to robust estimation of location is used for asymptotically ($N \to \infty$) locally optimum robust detection of weak signals ($\theta \to 0$) maximizing either the slope of the detection power at $\theta = 0$ [13], or the Pitman efficacy of the test statistic [10], or the detection power [1] (see also [11], [12]). Later, some of these approaches have been extended to more complicated models of signals and noises [3], [15], [17]-[19].

In the cases of application of Huber's minimax approach to robust detection, optimal detection rules are designed for specially selected detection rules or test statistics, e.g., for robust detectors based on M-estimators in [1], for a generalized correlator statistic in [10] and [13], for a distance criterion in [15]. Here, we adapt the following robust minimum distance detection rule

$$\sum_{i=1}^N \rho(x_i) \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \sum_{i=1}^N \rho(x_i - s_i), \tag{1}$$

where $\rho(x)$ is a distance measure [3, 15, 17], to the Neyman-Pearson setting.

Further, we consider an asymptotic weak signal approach when the signal $\{s_i\}_1^N$ decreases with the sample size N as $s_i = s_{iN} = A_i/\sqrt{N}$ with finite constants A_i such that the signal energy is bounded. Within a weak signal approach, the false alarm probability converges as $N \to \infty$ to a nonzero limit [1], [3], and Huber's minimax theory can be used to analyze the detector [8]. Since weak signals are on the border of not be distinguishable, it is especially important to know the detector performance.

An outline of the remainder of the paper is as follows. In Section 2, the power and false alarm probability of the proposed asymptotically maximin decision rule are derived. In Section 3, the optimal maximin detection rules are written out for the nondegenerate, with a bounded variance, and contaminated Gaussian noise pdfs. In Section 4, the detector performance is studied on large samples in the Gaussian and contaminated Gaussian noise pdf models. In Section 5, some conclusions are drawn.

2. MAIN RESULTS

Consider the following generalization of minimum distance detection rule (1)

$$\sum_{i=1}^N \rho(x_i) - \sum_{i=1}^N \rho(x_i - s_i) \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \lambda_\alpha, \tag{2}$$

where λ_α is a threshold defined by the bound α upon the false alarm probability

$$P_F = Pr\left[\sum_{i=1}^N \rho(x_i) - \sum_{i=1}^N \rho(x_i - s_i) > \lambda_\alpha \,\middle|\, H_0\right] \le \alpha. \tag{3}$$

To formulate further results, we introduce the derivative of a loss function $\psi = \rho'$ called a score function, belonging to a certain class Ψ.

Assume the following conditions of regularity imposed on a signal $\{s_i\}_1^N$, densities f, and score functions ψ:

$(A1)$ The signal $\{s_i\}_1^N$ is weak in the sense that its amplitudes form the decreasing with N sequences $s_i = A_i/\sqrt{N}$ with the finite constants A_i, $i = 1, \ldots, N$ such that the signal energy is finite: $\lim_{N\to\infty} \sum_{i=1}^{N} s_i^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} A_i^2 = \mathscr{E} < \infty$.

$(A2)$ f is symmetric and unimodal.

$(A3)$ f and ψ are continuously differentiable on $(0, \infty)$.

$(A4)$ $0 < I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)\,dx < \infty$.

$(A5)$ $E_f\,\psi = \int_{-\infty}^{\infty} \psi(x) f(x)\,dx = 0$.

$(A6)$ $E_f\,\psi^2 = \int_{-\infty}^{\infty} \psi^2(x) f(x)\,dx < \infty$.

$(A7)$ $0 < E_f\,\psi' = \int_{-\infty}^{\infty} \psi'(x) f(x)\,dx < \infty$.

Now we briefly comment on these conditions. In the literature, the conditions imposed on pdfs and score functions take different forms depending on the pursued goals: in general, one may strengthen the conditions on pdfs and weaken those on score functions, and vice versa (various suggestions can be found in [2], [4] - [7]). In this paper, we use a balanced set of conditions partially following [7].

The first condition $(A1)$, as aforementioned, is a traditional requirement used in an asymptotic weak signal approach [3], [11]. The condition $(A2)$ is restrictive but necessary for Huber's minimax theory [4].

The condition $(A3)$ differs from common conditions of this kind (for example, see [2], pp. 125-127; [4], p. 78, where the smoothness of pdfs is required in $\mathbb{R}$) allowing a noise pdf to have a discontinuity of its derivative at the center of symmetry, e.g., like the Laplace pdf, and thus widening the class of admissible densities.

The conditions $(A4)$ - $(A7)$ requiring the existence of the Fisher information $I(f)$ and other integrals are commonly used for the proofs of consistency and asymptotic normality of M-estimators in robust statistics [2], [4].

The following result is basic for all further constructions.

Lemma 1: Given $\{s_i\}_1^N$, f and ψ satisfying conditions $(A1)$ - $(A3)$, $(A5)$ - $(A7)$, the detector power for the rule (2) takes the following form as $N \to \infty$:

$$P_D(\psi, f) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - [\mathscr{E}/V(\psi, f)]^{1/2}\right), \quad (4)$$

where $\Phi(z) = (2\pi)^{-1/2} \int_{-\infty}^{z} \exp(-t^2/2)\,dt$ is the standard Gaussian CDF and $N^{-1}V(\psi, f)$ is the asymptotic variance of Huber's M-estimators of location [4] with $V(\psi, f) = E_f\,\psi^2/(E_f\,\psi')^2$.

The threshold value $\lambda_\alpha = \lambda_\alpha(\psi, f)$ is given by

$$\lambda_\alpha(\psi, f) = \Phi^{-1}(1-\alpha)\left(\mathscr{E} E_f\,\psi^2\right)^{1/2} - 0.5\,\mathscr{E} E_f\,\psi'. \quad (5)$$

The sketch of proof: The derivation of formula (4) is based on the Taylor expansion of the left-hand part of (2) and it is similar to the techniques used for derivation of the asymptotic variance for M-estimators in [4]; some examples of the application of those techniques to detection problems can be found in [12]. Using the same techniques, it can be shown that the false alarm probability (3) is given by the following expression

$$P_F = 1 - \Phi\left((\lambda_\alpha + 0.5\mathscr{E} E_f\,\psi') \Big/ \left(\mathscr{E} E_f\,\psi^2\right)^{1/2}\right), \quad (6)$$

equating which to α, we get the threshold value (5).

The result of Lemma 1 will be sensible if the power P_D tends to unit with the increasing energy $\mathscr{E}$. From (4) it follows that this holds when $\Phi^{-1}(1-\alpha) - [\mathscr{E}/V(\psi, f)]^{1/2} < 0$ or when

$$\alpha > \underline{\alpha}(\psi, f) = 1 - \Phi\left(\sqrt{\mathscr{E}/V(\psi, f)}\right). \quad (7)$$

From the Taylor expansion of the left-hand part of (2) it immediately follows that inequality (7) implies the consistency of detection, i.e., $\lim_{N\to\infty} P_D = 1$.

The consistency condition (7) means that there should be a lower bound $\underline{\alpha}$ on the false alarm probability (similar bounds also arise in other settings [1], [10]). Evidently, the greater energy, the lower the required minimum of the false alarm probability.

The Neyman-Pearson setting requires maximizing the detection power P_D under the bounded false alarm probability $P_F \le \alpha$; apparently, it can be achieved by choosing the maximum likelihood loss function $\rho(x) = \rho_{ML}(x) = -\log f(x)$ with the corresponding score function $\psi(x) = \psi_{ML}(x) = -f'(x)/f(x)$ in detection rule (2).

Theorem 1: Given pdf f, the Neyman-Pearson detection rule is provided by (2) with $\rho_{ML}(x) = -\log f(x)$ and λ_α defined by (5) with $\psi_{ML}(x) = -f'(x)/f(x)$. The corresponding detection power is given by

$$P_D(f) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - [\mathscr{E} I(f)]^{1/2}\right).$$

Proof: Since the asymptotic variance $V(\psi, f)$ attains its lower Cramér-Rao boundary at the maximum likelihood score function $V_{min} = V(\psi_{ML}, f) = 1/I(f)$, the required result directly follows from (4).

Now we are in position to consider the minimax setting when pdf f is not known: we return to the initial assumption that it belongs to a certain convex class $\mathscr{F}$ of distribution densities. From (4) it follows that the maximin problem with respect to the detection power $P_D(\psi, f)$ is equivalent to the Huber minimax problem with respect to the asymptotic variance $V(\psi, f)$ of M-estimators:

$$\max_{\psi\in\Psi} \min_{f\in\mathscr{F}} P_D(\psi, f) \quad \Longleftrightarrow \quad \min_{\psi\in\Psi} \max_{f\in\mathscr{F}} V(\psi, f).$$

Theorem 2: Under the conditions $(A1)$ – $(A7)$, the maximin Neyman-Pearson detection rule is given by (2) with the maximum likelihood choice of the loss function ρ^* for the least favorable density f^* minimizing Fisher information for location

$$\rho^*(x) = -\log f^*(x), \quad f^* = \arg\min_{f\in\mathscr{F}} I(f).$$

The threshold is defined by (5) with $\psi = \psi^*$ and $f = f^*$. Further, the maximin solution (ψ^*, f^*) provides the guaranteed lower bound on the power

$$\begin{aligned} P_D(\psi^*, f) &\ge P_D(\psi^*, f^*) \\ &= 1 - \Phi\left(\Phi^{-1}(1-\alpha) - [\mathscr{E} I(f^*)]^{1/2}\right) \end{aligned} \quad (8)$$

under the bounded false alarm probability $P_F(\psi^*, f) \le \alpha$ for all $f \in \mathscr{F}$ and for α satisfying the consistency condition $\alpha > \alpha^* = 1 - \Phi\left(\sqrt{\mathscr{E}\, I(f^*)}\right)$.

Proof: The first assertion directly follows from the saddle-point property

$$V(\psi^*, f) \le V(\psi^*, f^*) \le V(\psi, f^*) \tag{9}$$

(see [4], Theorem 2). Since P_D depends on the score function ψ and pdf f only through the $V(\psi, f)$, we obtain formula (8) for the lower bound $P_D(\psi^*, f^*)$ on the power.

Next we check whether the inequality $P_F(\psi^*, f) \le \alpha$ holds for all pdfs f in class $\mathscr{F}$. Using formula (6) and substituting the threshold λ_α^* defined by (5) for λ_α into the both parts of this inequality, we rewrite the latter in the following form

$$E_{f^*}\, \psi^{*2} \left(\Phi^{-1}(1-\alpha) - \frac{1}{2} \sqrt{\frac{\mathscr{E}}{V(\psi^*, f^*)}} \right)$$

$$\ge E_f\, \psi^{*2} \left(\Phi^{-1}(1-\alpha) - \frac{1}{2} \sqrt{\frac{\mathscr{E}}{V(\psi^*, f)}} \right).$$

Now divide the both parts of this inequality by $(E_{f^*}\, \psi^{*\prime})^2$, and as $V(\psi^*, f^*) = 1/I(f^*)$, it takes the following form

$$\begin{aligned} &\frac{1}{I(f^*)} \left(\Phi^{-1}(1-\alpha) - \frac{1}{2} \sqrt{\mathscr{E}\, I(f^*)} \right) \\ &\ge \frac{E_f\, \psi^{*2}}{(E_{f^*}\, \psi^{*\prime})^2} \left(\Phi^{-1}(1-\alpha) - \frac{1}{2} \sqrt{\frac{\mathscr{E}}{V(\psi^*, f)}} \right). \end{aligned} \tag{10}$$

Further, the maximum of the ratio $E_f\, \psi^{*2}/(E_{f^*}\, \psi^{*\prime})^2$ is attained at $f = f^*$ being equal to $V(\psi^*, f^*)$; therefore, by the minimax property (9) $V(\psi^*, f) \le V(\psi^*, f^*)$, inequality (10) and the required inequality $P_F(\psi^*, f) \le \alpha$ hold for all $\alpha \in (0, 1)$.

3. MAXIMIN DETECTORS FOR VARIOUS NOISE DISTRIBUTION CLASSES

Within the minimax approach, the choice of a noise distribution class $\mathscr{F}$ entirely determines the structure of a maximin detector. Below we consider qualitatively different noise distribution classes with the corresponding least favorable densities and maximin detectors.

3.1 Nondegenerate Noise Distributions

In the class of nondegenerate pdfs (with a bounded density value at the center of symmetry)

$$\mathscr{F}_1 = \{ f:\ f(0) \ge 1/(2a) > 0 \},$$

the scale parameter a describes the distribution dispersion about the center of symmetry. The class $\mathscr{F}_1$ is one of the most wide classes: any unimodal distribution density with a nonzero value at the center of symmetry belongs to it. The least favorable density here is the Laplace [14]:

$$f_1^*(x) = (2a)^{-1} \exp\left(-|x|/a\right).$$

Thus, we have the maximin L_1-norm detector with $\rho^*(x) = |x|/a$

$$\sum_{i=1}^{N} |x_i| - \sum_{i=1}^{N} |x_i - s_i| \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \Phi^{-1}(1-\alpha)\sqrt{\mathscr{E}} - \mathscr{E}/(4a).$$

The lower bounds on power and on false alarm probability given by Theorem 2 are as follows

$$P_D(\psi^*, f) \ge 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{\mathscr{E}}/a \right),$$

$$\alpha > \alpha^* = 1 - \Phi\left(\sqrt{\mathscr{E}}/a \right).$$

3.2 Noise Distributions with a Bounded Variance

In the class of densities with a bounded variance

$$\mathscr{F}_2 = \left\{ f:\ \int_{-\infty}^{\infty} x^2 f(x)\, dx \le \overline{\sigma}^2 \right\},$$

the least favorable density is Gaussian [9]:

$$f_2^*(x) = N(x; 0, \overline{\sigma}) = (\sqrt{2\pi}\overline{\sigma})^{-1/2} \exp\left(-x^2/(2\overline{\sigma}^2)\right).$$

The maximin L_2-norm detector uses the quadratic distance $\rho^*(x) = x^2/(2\overline{\sigma}^2)$ and the detection rule (2) can be rewritten in *the correlation detector* form as

$$\sum_{i=1}^{N} x_i\, s_i \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \Phi^{-1}(1-\alpha)\,\overline{\sigma}\,\sqrt{\mathscr{E}} - \mathscr{E}.$$

The lower bounds on power and on false alarm probability are given by

$$P_D(\psi^*, f) \ge 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{\mathscr{E}}/\overline{\sigma} \right),$$

$$\alpha > \alpha^* = 1 - \Phi\left(\sqrt{\mathscr{E}}/\overline{\sigma} \right).$$

Remark: The minimax approach does not necessarily imply robustness, since the L_2-norm detector being maximin in the Huber sense in the class of distributions with a bounded variance, is not at all robust, nevertheless, being a detector of guaranteed power in class $\mathscr{F}_2$. Thus, if the upper-bound $\overline{\sigma}^2$ on variance is small, then the minimax approach yields a reasonable result and the L_2-norm detector can be successfully used with relatively light-tailed noise distributions, e.g., see [17]. On the contrary, if we deal with really heavy-tailed distributions (gross errors, impulse noises) when $\overline{\sigma}^2$ is large or even infinite like for the Cauchy-type distributions, then the maximin solution in class $\mathscr{F}_2$ is still trivially correct as $P_D(\psi^*, f) \ge \alpha$ and $\alpha > 1/2$, but practically senseless.

3.3 Contaminated Gaussian Noise Distributions

Consider the class of ε-contaminated Gaussian pdfs

$$\mathscr{F}_H = \{ f:\ f(x) = (1-\varepsilon) N(x; 0, \sigma) + \varepsilon h(x) \},$$

where $h(x)$ is an arbitrary pdf and ε ($0 \le \varepsilon < 1$) is a contamination parameter giving the fraction of contamination. The least favorable density consists of two parts: the Gaussian in the center and the exponential tails [4]. The maximin

Table 1: The Minimum SNR^* Values Providing Consistency of Detection

	$\varepsilon = 0.01$	$\varepsilon = 0.05$	$\varepsilon = 0.1$	$\varepsilon = 0.2$
$\alpha = 0.01$	5.77	6.80	8.07	9.47
$\alpha = 0.05$	2.88	3.40	4.03	5.54
$\alpha = 0.1$	1.75	2.06	2.45	3.36

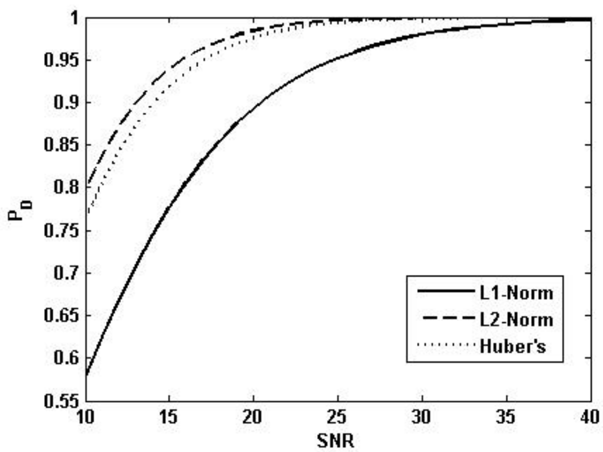

Figure 1: The asymptotic power of the L_1-norm, L_2-norm, and Huber's maximin detectors in the Gaussian noise with density $N(x;0,\sigma)$; $P_F = 0.01$.

Huber's detector uses the piece-wise linear-quadratic distance function $\rho^*(x) = x^2/(2\sigma^2)$ for $|x| \le k\sigma$ and $\rho^*(x) = k|x|/\sigma - k^2/2$ for $|x| > k$, where the dependence $k = k(\varepsilon)$ is tabulated in [4]. The threshold λ^*_α is given by (5)

$$\lambda^*_\alpha = \Phi^{-1}(1-\alpha)\left((1-\varepsilon)\left[2\Phi(k)-1\right]\frac{\mathscr{E}}{\sigma^2}\right)^{1/2} - \frac{(1-\varepsilon)\left[2\Phi(k)-1\right]}{2}\frac{\mathscr{E}}{\sigma^2}.$$

Here we present the numerical results for $\varepsilon = 0.1$: the lower bounds on power and on false alarm probability are as follows

$$P_D(\psi^*, f) \ge 1 - \Phi\left(\Phi^{-1}(1-\alpha) - 0.819\sqrt{\mathscr{E}}/\sigma\right),$$

$$\alpha > \alpha^* = 1 - \Phi\left(0.819\sqrt{\mathscr{E}}/\sigma\right).$$

4. PERFORMANCE EVALUATION

Now we compare the asymptotic performance of the maximin detectors in the classes $\mathscr{F}_1$, $\mathscr{F}_2$, and $\mathscr{F}_H$ with respect to their lower bounds on the false alarm probability and power. In order to use for comparisons the conventional signal-noise ratio term (SNR), we take the corresponding least favorable distributions of the same (unit) variance: set $a = 1/\sqrt{2}$, $\overline{\sigma}^2 = 1$ and $\varepsilon = 0.1$, $\sigma^2(f^*_H) = 1$ for the classes $\mathscr{F}_1$, $\mathscr{F}_2$ and $\mathscr{F}_H$, respectively. Next, choose the false alarm probability $\alpha = 0.01$. For the corresponding lower bounds on false alarm probability, we have $\alpha^*_1 = 1 - \Phi\left(\sqrt{2\,SNR}\right)$, $\alpha^*_2 = 1 - \Phi\left(\sqrt{SNR}\right)$, $\alpha^*_H = 1 - \Phi\left(0.798\sqrt{SNR}\right)$, where $SNR = \mathscr{E}/\sigma^2(f)$ is the customary signal-noise ratio. Then the lower bounds upon detection power are $P^*_{D1} = 1 - \Phi\left(2.3276 - \sqrt{2 \times SNR}\right)$, $P^*_{D2} = 1 - \Phi\left(2.3276 - \sqrt{SNR}\right)$, and $P^*_{DH} = 1 - \Phi\left(2.3276 - 0.798\sqrt{SNR}\right)$.

The minimal values SNR^* providing consistency of detection when $SNR > SNR^*$ are as follows:

$$SNR^*(\alpha) = \left[\Phi^{-1}(1-\alpha)\right]^2 / I(f^*).$$

As in practice the class of contaminated Gaussian noise distributions is mostly required, Table 1 exhibits the SNR^* values versus the false alarm probability α and the parameter of contamination ε.

Next we compare the asymptotic performance of the maximin detectors on the Gaussian distribution with density $N(x;0,1)$ and the heavy-tailed Cauchy-contaminated Gaussian with density $0.9N(x;0,1) + 0.1/(\pi(1+x^2))$. The results are exhibited in Figs. 1 – 2, where the power is computed for the SNR values sufficiently large to provide consistency of detection (see Table 1).

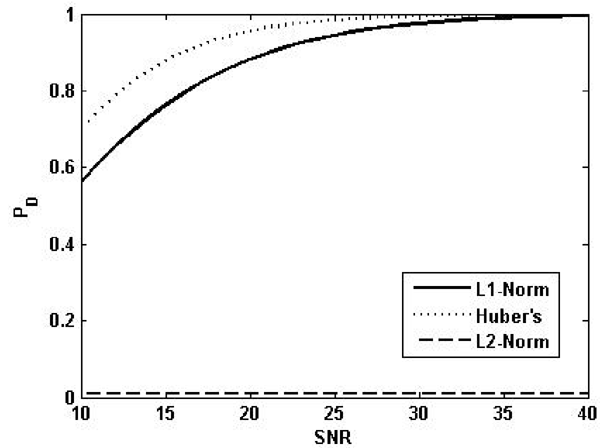

Figure 2: The asymptotic power of the L_1-norm, L_2-norm, and Huber's maximin detectors in the Cauchy contaminated Gaussian noise with density $0.9N(x;0,\sigma) + 0.1/(\pi(1+x^2))$; $P_F = 0.01$. In this case, $P_{D2} = 0.01$.

The L_1-norm and Huber's detectors confirm their robust properties in heavy-tailed noise, Huber's being better than the L_1-norm detector in the Gaussian noise. Naturally, the L_2-norm is optimal in the Gaussian noise and catastrophically bad in the contaminated Gaussian noise ($P_{D2} = \alpha$). Finally, Huber's detector can be regarded as a reasonable compromise between the L_1- and L_2-norm detectors.

The performance of the L_1-norm, L_2-norm and Huber's detectors was also studied by Monte Carlo technique on small samples ($N = 20$) and the obtained results were qualitatively similar to those on large samples.

5. CONCLUSION

Our main aim is to expose a new result on the application of Huber's minimax approach to robust detection in the particular case of the minimum distance Neyman-Pearson detectors.

First, we generalize the minimum distance detection rule (1) introducing the maximin Neyman-Pearson detector (2) with the guaranteed lower bound on the power under general regularity conditions (Theorem 2).

Second, the maximin detectors are designed for the classes of nondegenerate, with a bounded variance and ε-contaminated Gaussian noise distributions.

Finally, note that the obtained results can be extended on the classes with bounded distribution subranges and their various combinations [16], [17].

REFERENCES

[1] A. H. El-Sawy and V. David Vandelinde, "Robust detection of known signals," *IEEE Trans. Inform. Theory,* vol. 23, pp. 722-727, 1977.

[2] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, *Robust Statistics. The Approach Based on Influence Functions* , New York, Wiley, 1986.

[3] O. Hossjer and M. Mettiji, "Robust multiple classification of known signals in additive noise – an asymptotic weak signal approach," *IEEE Trans. Inform. Theory,* vol. 39, pp. 594-608, 1993.

[4] P. J. Huber, "Robust estimation of a location parameter", *Ann. Math. Statist.*, vol. 35, pp. 1-72, 1964.

[5] P. J. Huber, "A robust version of the probability ratio test", *Ann. Math. Statist.*, vol. 36, pp. 1753-1758, 1965.

[6] P. J. Huber, "The behavior of maximum likelihood estimates under nonstandard conditions", In: *Proc. 5th Berkeley Symp. on Math. Statist. Prob.*, vol. 1, Berkeley Univ. California Press, pp. 221-223, 1967.

[7] P. J. Huber, "Fisher information and spline interpolation", *Ann. Statist.*, vol. 2, pp. 1029-1033, 1974.

[8] P. J. Huber, "*Robust Statistics*," New York, Wiley, 1981.

[9] A. M. Kagan, Yu. V. Linnik, and S. R. Rao, *Characterization Problems in Mathematical Statistics*, New York, Wiley, 1973.

[10] S. A. Kassam and J. B. Thomas, "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 22, pp. 22-26, 1976.

[11] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: a survey," *Proc. IEEE*, vol. 73, pp. 433-481, 1985.

[12] S. A. Kassam, "*Signal Detection in Non-Gaussian Noise*," Berlin, Springer, 1988.

[13] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 17, pp. 50-56, 1971.

[14] B. T. Polyak and Ya. Z. Tsypkin, "Robust identification," *Automatica*, vol. 16, pp. 53-63, 1980.

[15] H. V. Poor, "Robust decision design using a distance criterion," *IEEE Trans. Inform. Theory*, vol. 26, pp. 575-587, 1980.

[16] G. L. Shevlyakov and N. O. Vilchevski, *Robustness in Data Analysis: criteria and methods* , Utrecht, VSP, 2002.

[17] G. L. Shevlyakov and K. S. Kim, "Robust Minimax Detection of a Weak Signal in Noise with a Bounded Variance and Density Value at the Center of Symmetry", *IEEE Trans. Inform. Theory*, vol. 52, pp. 1206-1211, 2006.

[18] A. Swami and B. Sadler, "On some detection and estimation problems in heavy-tailed noise", *Signal Processing*, vol. 82, pp. 1829-1846, 2002.

[19] X. Wang and H. V. Poor, "Robust multiuser detection in non-gaussian channels," *IEEE Trans. Signal Proc.*, vol. 47, pp. 289-305, 1999.