[12] G. D. Forney Jr., R. Gallager, G. Lang, F. Longstaff, and S. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 5, pp. 632–647, Sep. 1984.
[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: Wiley, 1968.
[14] D. E. Knuth, *The Art of Computer Programming Volume 2: Seminumerical Algorithms*. Reading, MA: Addison-Wesley, 1969.
[15] S. D. Casey and B. M. Sadler, "Modifications of the Euclidean algorithm for isolating periodicites from a sparse set of noisy measurements," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 2260–2272, Sep. 1996.

Robust Minimax Detection of a Weak Signal in Noise With a Bounded Variance and Density Value at the Center of Symmetry

Georgy Shevlyakov, *Member, IEEE,* and Kiseon Kim, *Senior Member, IEEE*

***Abstract*—In practical communication environments, it is frequently observed that the underlying noise distribution is not Gaussian and may vary in a wide range from short-tailed to heavy-tailed forms. To describe partially known noise distribution densities, a distribution class characterized by the upper-bounds upon a noise variance and a density dispersion in the central part is used. The results on the minimax variance estimation in the Huber sense are applied to the problem of asymptotically minimax detection of a weak signal. The least favorable density minimizing Fisher information over this class is called the Weber–Hermite density and it has the Gaussian and Laplace densities as limiting cases. The subsequent minimax detector has the following form: i) with relatively small variances, it is the minimum L_2-norm distance rule; ii) with relatively large variances, it is the L_1-norm distance rule; iii) it is a compromise between these extremes with relatively moderate variances. It is shown that the proposed minimax detector is robust and close to Huber's for heavy-tailed distributions and more efficient than Huber's for short-tailed ones both in asymptotics and on finite samples.**

***Index Terms*—Huber's M-estimators, least favorable distributions, non-Gaussian noise, robust minimum distance detection.**

I. Introduction

Consider the problem of detection of a known signal θ in the additive independent and identically distributed (i.i.d.) noise $\{n_i\}_1^N$ with pdf f from a certain class $\mathcal{F}$. Given $\{x_i\}_1^N$, it is necessary to decide whether the signal θ is observed. This problem of binary detection is set up as the problem of hypotheses testing: $H_0 : x_i = n_i$ versus $H_1 : x_i = \theta + n_i,\ i = 1, \ldots, N$. Given a pdf f, the classical theory of hypotheses testing yields various optimal (in the Bayesian, minimax, Neyman-Pearson senses) algorithms for the solution of this problem: all the optimal algorithms are based on the value of the likelihood ratio (LR) statistic $T_N(\mathbf{x}) = \prod_i^N f(x_i - \theta)/f(x_i)$ that must be compared with a certain threshold. The differences between the aforementioned approaches result only in the values of a threshold.

Manuscript received June 10, 2004; revised November 10, 2005. The material in this correspondence was presented in part at Nordic Radio Symposium 2004, Oulu, Finland, August 2004.

The authors are with the Department of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea (e-mail: shev@gist.ac.kr; kskim@gist.ac.kr).

Communicated by X. Wang, Associate Editor for Detection and Estimation.
Digital Object Identifier 10.1109/TIT.2005.864462

In this correspondence, we consider the asymptotic weak signal approach when the useful signal θ decreases with sample size as $\theta = \theta_N = A/\sqrt{N}$ given some constant $A > 0$. For reasonable decision rules, the error probability then converges as $N \to \infty$ to a nonzero limit [6]. Moreover, within this approach, the error probability is closely related to the Pitman efficacy of the detector test statistic, and therefore, Huber's minimax theory can be used to analyze the detector [10]–[12]. Finally, since weak signals are on the border of not be distinguishable, and therefore, it is especially important to know the error probabilities.

In what follows, we deal with the following minimum distance detection rule [6]:

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

where $\rho(z)$ is a loss function characterizing the assumed form of a distance. This choice of a detection rule is mainly determined by the fact that it allows for the direct and simple use of Huber's minimax theory on M-estimators of location [7], [8]. Further, it can be seen that the choice $\rho(z) = -\log f(z)$ makes the optimal LR test statistic minimizing the Bayesian risk with equal costs and prior probabilities of hypotheses. Note, that in this case, it is necessary to know exactly the shape of pdf f to figure out the distance function, and the LR-statistics usually behave poorly under the departures from the assumed pdf model.

In many practical problems of radio-location, acoustics, and communications, noise distributions are only partially known. For instance, it may be known that either the underlying pdf is approximately Gaussian, or there is some information on its behavior in the central zone and on the tails, or an impulsive noise may distort the observed signal, etc. For these detection problems, some robust alternatives to the classical methods have been proposed in [8], [5], [10]–[12], [6], [4]. Recently, some of these approaches have been extended to more complicated static models of signals under the assumptions of the approximately Gaussian character of noise distributions [3], [18]. Heavy-tailed non-Gaussian noise models with finite and infinite variances both for static and dynamic systems are considered in many works, for example, in [2], [13], [16]. However, we are interested in a static model containing short-tailed noise pdfs with small variances as well as the heavy-tailed ones with large or even with infinite variances.

Within the minimax approach, the choice of a distribution class $\mathcal{F}$ determines all the subsequent stages and the qualitative character of the corresponding robust procedure. In its turn, the choice of a distribution class depends either on the available prior information about data distributions, or on the possibilities of getting this information from the data sample.

Being historically the first [7], various ε-neighborhoods of the Gaussian distribution are not the only models of interest. In practice there often exists a prior information about the distribution dispersion in its central part and/or its tails, about the moments and/or subranges of a distribution. The empirical distribution function and relative estimators of a distribution shape (quantile functions and their approximations, histograms, kernel estimators) along with their confidence boundaries give other examples. In order to enhance efficiency of robust minimax procedures, it is reasonable to use such information in minimax settings by introducing the corresponding distribution classes. In Section II, we describe such a class.

We now dwell on the contributions of this paper. In [15], the distribution class with a bounded variance and density value at the center of symmetry, as well as some other classes with bounded distribution variances and subranges, were effectively used for robust estimation of location, regression and autogression parameters in heavy-tailed distribution models being close in performance to Huber's conventional approach. In this paper, first, we apply the aforementioned class to the problems of detection under heavy-tailed noise distributions, and since the problems of estimation and detection are very close, we may expect the same effect also in this case. Second, and this was unfortunately overlooked in [15], under short-tailed noise distributions, minimax algorithms in this class work much better than Huber's conventional algorithms. Third, we consider not a neighborhood of the Gaussian pdf but a class in which the Gaussian pdf with the corresponding least squares method is not explicitly assumed in the model but arises as the least favorable distribution in the classes with bounded noise variances, and just this extends the possibilities of Huber's minimax approach and provides the aforementioned effect.

This paper is organized as follows. In Section II, a brief survey of the classical results within Huber's minimax approach is given and the least favorable density in the class of densities with an upper-bounded variance and density value at the center of symmetry is described. In Section III, we show that the error probability for the minimum distance detector is asymptotically minimax in the Huber sense and study the detector performance both in asymptotics and on finite samples for the Gaussian, heavy-tailed ε-contaminated Gaussian, Cauchy, and short-tailed generalized Gaussian noise pdfs. In Section IV, concluding remarks are made.

II. The Least Favorable Density in the Class With a Bounded Variance and Density Value at the Center of Symmetry

Since our results are essentially based on Huber's minimax approach to robust estimation of location, we briefly recall its basic stages. In general, the minimax principle aims at the least favorable situation for which it suggests the best solution. Thus, in some sense, this approach provides a guaranteed result, possibly too pessimistic.

Let $x_1, \ldots, x_N$ be i.i.d. random variables with common density $f(x-\theta)$ in a convex class $\mathcal{F}$. Then the M-estimator $\hat{\theta}_N$ of a location parameter θ is defined as a zero of $\sum_1^N \psi(x_i - \cdot)$ with a suitable score function ψ [7]. The minimax approach implies the determination of the least favorable density f^* minimizing Fisher information $I(f) = \int (f'/f)^2 f\,dx$ over the class $\mathcal{F}: f^* = \arg\min_{f\in\mathcal{F}} I(f)$, followed by designing the maximum-likelihood estimator (MLE) with the score function $\psi^* = -f^{*\prime}/f^*$. Under rather general conditions of regularity, $\sqrt{N}(\hat{\theta}_N - \theta)$ is asymptotically normal with variance $V(\psi, f) = \int \psi^2 f dx / [\int \psi' f dx]^2$ satisfying the minimax property

$$V(\psi^*, f) \le V(\psi^*, f^*) \le V(\psi, f^*).$$

The saddle-point pair (ψ^*, f^*) provides the guaranteed boundary of estimation accuracy $V(\psi^*, f) \le V(\psi^*, f^*) = 1/I(f^*)$ for all $f \in \mathcal{F}$.

In other words, Huber proposed to use the supremum of the asymptotic variance $V(\psi^*, f^*) \ge V(\psi^*, f)$ as a measure of robustness of the optimal M-estimator with the score function ψ^*: the less the range of the optimal estimator variance $V(\psi^*, f)$ over the class $\mathcal{F}$, the more robust is this estimator, and vice versa.

The shape of the least favorable density f^* and the corresponding score function ψ^* is wholly determined by the structure of class $\mathcal{F}$. We now enlist several results on the least favorable distributions in the distribution classes qualitatively different from the conventional ε-contaminated Gaussian models. The symmetry and unimodality of distribution densities are assumed.

In the class $\mathcal{F}_1$ of nondegenerate pdfs (with a bounded density value at the center of symmetry), the least favorable density is known to be the Laplace [14], [4]

$$\mathcal{F}_1 = \{f\colon f(0) \ge 1/(2a) > 0\}$$
$$f_1^*(x) = L(x;0,a) = (2a)^{-1}\exp(-|x|/a),$$

here the scale parameter a characterizes the pdf dispersion about the center of symmetry. In this case, we have the sign score function $\psi_1^*(z) = \operatorname{sgn}(z)/a$ and the sample median as the optimal L_1-norm estimator. It is one of the most wide classes: any unimodal distribution density with a nonzero value at the center of symmetry belongs to it. The condition of belonging to this class is very close to the complete lack of information about an underlying distribution.

In the class $\mathcal{F}_2$ of pdfs with an upper-bounded variance, the least favorable density is the Gaussian [9]

$$\mathcal{F}_2 = \left\{f\colon \sigma^2(f) = \int x^2 f(x)dx \le \bar{\sigma}^2\right\}$$
$$f_2^*(x) = N(x;0,\bar{\sigma}) = \frac{1}{\bar{\sigma}\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\bar{\sigma}^2}\right)$$

with the corresponding linear score function $\psi_2^*(z) = z/\overline{\sigma}^2$ and the sample mean as the optimal L_2-norm estimator.

Remark: Note that minimax approach does not necessarily imply robustness like in the class $\mathcal{F}_2$. But the lack of stability of this solution in heavy-tailed models may be compensated by its higher efficiency under short-tailed distributions.

As the optimal solutions for two aforementioned classes have compensable characteristics of robustness and efficiency, we propose another class of pdfs containing the restrictions of the both classes $\mathcal{F}_1$ and $\mathcal{F}_2$

$$\mathcal{F}_{12} = \{f : f(0) \ge 1/(2a) > 0,\ \sigma^2(f) \le \bar{\sigma}^2\}. \tag{2}$$

Note that the introduced class of densities comprises qualitatively different densities, for example, the Gaussian, the heavy-tailed ε-contaminated Gaussian, Laplace, Cauchy-type (with $\bar{\sigma}^2 = \infty$), short-tailed pdfs close to the uniform, etc.

For the class $\mathcal{F}_{12}$, the least favorable pdf simultaneously depends on the two parameters a and $\bar{\sigma}$ through their ratio $\bar{\sigma}/a$ naturally having the Gaussian and Laplace densities as the particular cases and is of the form [17], [15]:

$$f_{12}^*(x) = \begin{cases} N(x;0,\bar{\sigma}), & \text{for } \bar{\sigma}^2/a^2 < 2/\pi \\ \mathrm{WH}(x;0,\nu,\bar{\sigma}), & \text{for } 2/\pi \le \bar{\sigma}^2/a^2 \le 2 \\ L(x;0,a), & \text{for } \overline{\sigma}^2/a^2 > 2. \end{cases} \tag{3}$$

Here $N(x;0,\bar{\sigma})$ and $L(x;0,a)$ are the Gaussian and Laplace pdfs, respectively, and $WH(x;0,\nu,\bar{\sigma})$ being called the Weber–Hermite pdf is given by

$$\mathrm{WH}(x;0,\nu,\overline{\sigma}) = \frac{\Gamma(-\nu)\sqrt{2\nu+1+1/S(\nu)}}{\sqrt{2\pi}\,\overline{\sigma}\,S(\nu)}\mathcal{D}_\nu^2\left(\frac{|x|}{\overline{\sigma}}\sqrt{2\nu+1+1/S(\nu)}\right) \tag{4}$$

with the real-valued shape parameter ν that takes its values in the interval $(-\infty, 0]$ and also depends on the ratio $\bar{\sigma}/a$ (for details, see [15, p. 68]). Further, $\mathcal{D}_\nu(\cdot)$ are the Weber-Hermite functions or the functions of the parabolic cylinder [1], $S(\nu) = [\psi(1/2 - \nu/2) - \psi(-\nu/2)]/2$, and in this context, $\psi(x) = d\ln\Gamma(x)/dx$ is the digamma function.

The Weber–Hermite pdfs (4) arise as the solution to the Euler–Lagrange equation for the variational problem of minimizing Fisher information. The Gaussian and Laplace pdfs are the particular cases of (4) when $\nu = 0$ and $\nu \to -\infty$, respectively. For $\nu < 0$, the shape of the Weber–Hermite pdf takes an intermediate unimodal form between the Gaussian and Laplace pdf forms, having a discontinuity of the pdf derivative at the center of symmetry similar to the Laplace pdf.

To understand the branch structure of the least favorable density $f_{12}^*(x)$, we have to consider the two aspects of the question. First, the value of the ratio $\bar{\sigma}^2/a^2$ reflects the relative weight of tails in the total distribution dispersion, since the pdf value at the center of symmetry and the variance along with their corresponding bounds a and $\bar{\sigma}$ are especially sensitive to the form of a distribution central part and tails, respectively. Thus, for $\overline{\sigma}^2/a^2 < 2/\pi$ or with relatively short tails, we have the Gaussian branch of the $f_{12}^*(x)$, and so on. Second, different branches of the $f_{12}^*(x)$ appear due to the degree in which the constraints are taken into account: for $\bar{\sigma}^2/a^2 < 2/\pi$ or with relatively short tails, only the restriction upon a variance does matter (it becomes the equality $\sigma^2(f_{12}^*) = \bar{\sigma}^2$), and the restriction upon the pdf value at the center of symmetry takes the form of the strict inequality $f_{12}^*(0) > 1/(2a)$, and therefore it is removed, thus class $\mathcal{F}_{12}$ is reduced to class $\mathcal{F}_2$ with the corresponding least favorable Gaussian density. Similarly, for $\bar{\sigma}^2/a^2 > 2$ or with relatively heavy tails, only the restriction upon the pdf value at the center of symmetry is essential and the the restriction upon a variance is removed. Therefore, class $\mathcal{F}_{12}$ is reduced to class $\mathcal{F}_1$ with the corresponding least favorable Laplace density. Finally, for $2/\pi \le \bar{\sigma}^2/a^2 < 2$, both restrictions become the equalities and both must be taken into account, so the least favorable density takes the most general Weber-Hermite form.

The corresponding minimax estimator of location can be described as follows: i) with $\bar{\sigma}^2(f) \le 2a^2/\pi$ or with relatively small variances (relatively short tails), it is the sample mean or the L_2-norm estimator; ii) with $\bar{\sigma}^2 > 2a^2$ or with relatively large variances (relatively heavy tailes), it is the sample median or the L_1-norm estimator; iii) with relatively moderate variances, it is a compromise between the L_1-norm and the L_2-norm estimators. In the latter case, the loss function is $\rho(z) = -\log f_{12}^*(z)$ that can be rather effectively (with at most 2.5% relative error) approximated by the low-complexity L_{p^*}-norm loss function with the power $p^* \in (1,2)$ given by (for details, see [15, p. 76–79])

$$p^* = \begin{cases} 5.33 - 7.61x + 3.73x^2, & 2/\pi < x \le 1.35 \\ 2.66 - 1.65x + 0.41x^2, & 1.35 < x < 2 \end{cases} \tag{5}$$

where $x = \bar{\sigma}^2/a^2$.

III. Detection Performance in Asymptotics and on Finite Samples

Under rather general conditions of regularity imposed on the classes Ψ and $\mathcal{F}$ (see, say, [5, pp. 125–127]), the error probability of detection for the minimum distance rule (1) takes the following form as $N \to \infty$:

$$P_E = \mathcal{Q}\left(\frac{A}{2\sqrt{V(\psi, f)}}\right) \tag{6}$$

where

$$\mathcal{Q}(x) = (2\pi)^{-1/2}\int_x^\infty e^{-t^2/2}dt$$

the parameter A determines the amplitude of a weak signal as a decreasing sequence $\theta = \theta_N = A/\sqrt{N}$, and $V(\psi, f)$ is the aforementioned asymptotic variance of Huber's M-estimators of location. Since the signal energy E is equal to $\theta^2 N$, we have $A = \sqrt{E}$, and moreover, in the particular case of the unit noise variance, it can be written as $A = \sqrt{\mathrm{SNR}}$. This result can be derived using the standard asymptotic technique as in [7], [6], [12].

From (6) it follows that the minimax problem with respect to the error probability $\min_{\psi\in\Psi}\max_{f\in\mathcal{F}} P_E(\psi, f)$ is equivalent to Huber's minimax problem $\min_{\psi\in\Psi}\max_{f\in\mathcal{F}} V(\psi, f)$. Thus, all the results on the minimax estimation of location are also applicable in this case: the optimal loss function ρ^* in the minimum distance detector (1) is defined by the maximum likelihood choice for the least favorable (informative) pdf f^* minimizing Fisher information: $\rho^*(z) = -\log f^*(z)$. Moreover, the error probability is upper-bounded in class $\mathcal{F}: P_E(\psi^*, f) \le P_E(\psi^*, f^*)$.

Further, we compute the error probabilities for the Gaussian noise with the pdf $f(x) = N(x; 0, 1)$, the heavy-tailed Cauchy noise with the pdf $C(x; 0, 1) = 1/[\pi(1 + x^2)]$, the heavy-tailed ε-contaminated Gaussian noise with the pdf $f_{\mathrm{CN}}(x) = (1-\varepsilon)N(x; 0, 1) + \varepsilon C(x; 0, 1)$ where ε is the contamination parameter $(0 \le \varepsilon < 1)$, and for the generalized Gaussian noises with the pdf

$$f_{\mathrm{GG}}(x; \beta, q) = \frac{q}{2\beta\Gamma(1/q)}\exp\left(-\frac{|x|^q}{\beta^q}\right)$$

where β and q are the parameters of scale and shape, respectively.

Also we set the unit distribution variance with the following exceptions—for the Cauchy and ε-contaminated Gaussian pdfs.

Further, we compare the performance of the low-complexity minimax L_{p^*}-norm detector

$$\sum_{i=1}^N |x_i|^{p^*} \lessgtr \sum_{i=1}^N |x_i - \theta|^{p^*} \tag{7}$$

in which the power p^* is chosen from (5) with the L_1-, L_2-norm, and Huber's detectors. In the latter case, we consider the minimum distance detection rule (1) with the conventional Huber loss function $\rho_H^*(z) = -\log f_H^*(z)$ where the least favorable distribution is described by the Gaussian central part and exponential tails (for details, see [8]) and the contamination parameter ε is fixed to 0.1. The chosen value of ε seems to be a reasonable upper bound upon the contamination parameter in applications [5].

To clarify the procedure of computing the error probability, let us consider the Gaussian case in details. To simplify the analysis, we rewrite (6) as follows:

$$P_E = \mathcal{Q}\left(A\frac{I_1}{2\sqrt{I_2}}\right) \tag{8}$$

where

$$I_1 = \int_{-\infty}^{\infty} \psi'(x)f(x)dx \quad \text{and} \quad I_2 = \int_{-\infty}^{\infty} \psi^2(x)f(x)dx.$$

Example 1: Consider the error probability for the minimax L_{p^*}-norm and L_2-norm detectors in the Gaussian noise. Then the choice of the optimal structure is defined by the ratio σ^2/a^2. Subsequently, we have $f(0) = 1/\sqrt{2\pi}, \sigma^2(f) = 1, a = \sqrt{\pi/2}, \sigma^2/a^2 = 2/\pi$. Hence, the minimax detector is based on the L_2-norm distance with $p^* = 2$, the score function is linear $\psi^*(x) = x$, the integrals are $I_1 = 1$ and $I_2 = 1$, and thus, its error probability is given by $P_E = \mathcal{Q}(0.5A)$.

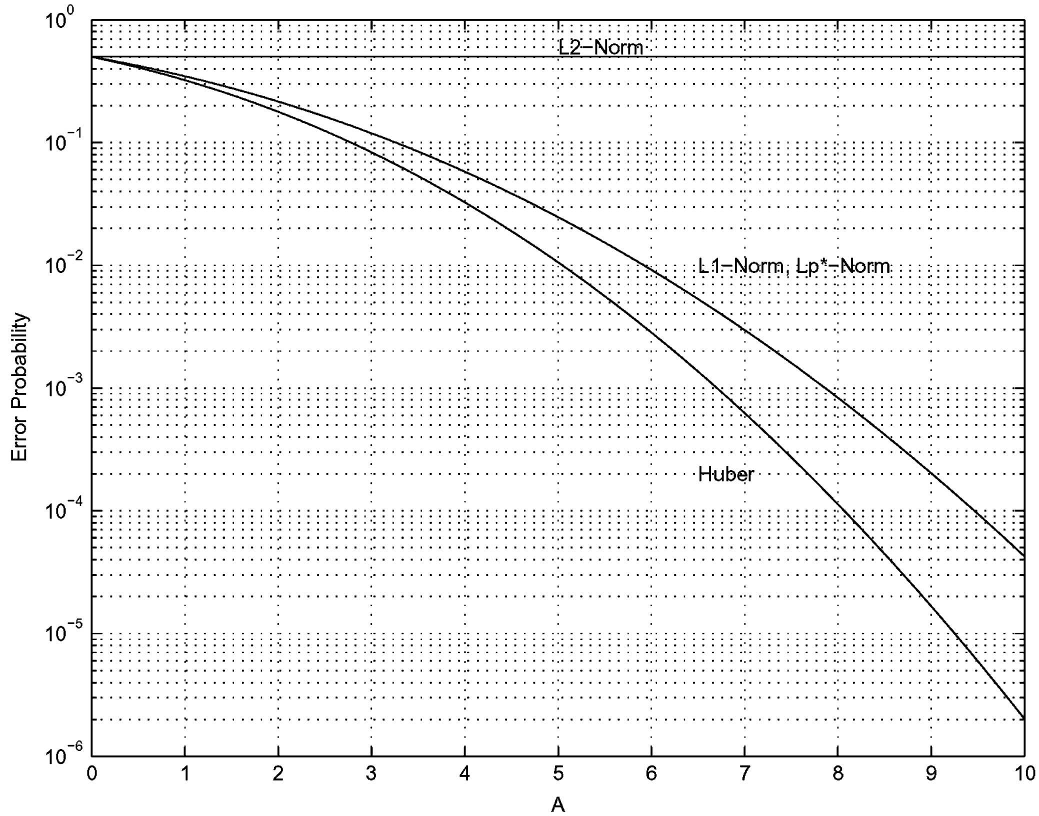

Fig. 1. Detection in the ε-contaminated Gaussian noise: asymptotics, $\varepsilon = 0.1, A = \sqrt{E}$. (To avoid ambiguity, note that here and henceforth, the parameter A unit is not dB).

For the L_2-norm detector, apparently the choice is the same: $\psi(x) = x$ and $P_E = \mathcal{Q}(0.5A)$.

Example 2: Consider the error probability for Huber's and L_1-norm detectors in the Gaussian noise. For $\varepsilon = 0.1$, the score function is $\psi(x) = \psi_H^*(x) = \max[-1.14, \min(x, 1.14)]$, the required integrals cannot be evaluated in a closed form, hence they were computed numerically, and $P_E = \mathcal{Q}(0.461A)$.

For the L_1-norm detector with $p^* = 1$, the score function is the sign function $\psi(x) = \mathrm{sgn}(x)$, the integrals are $I_1 = 2f(0) = \sqrt{2/\pi}$ and $I_2 = 1$, and thus, we get $P_E = \mathcal{Q}(A/\sqrt{2\pi}) = \mathcal{Q}(0.399A)$.

The results of computing for the Gaussian, extremely heavy-tailed Cauchy, heavy-tailed ε-contaminated Gaussian with $\varepsilon = 0.1$ and the close to the uniform pdf exponential-power ($q = 100$) density are exhibited in Table I and Figs. 1 and 2.

TABLE I
THE FACTOR $\frac{I_1}{2\sqrt{I_2}}$ IN (8) FOR VARIOUS NOISES

	L_1	L_2	Huber's: $\varepsilon = 0.1$	L_{p^*}
Gaussian: $N(x;0,1)$	*Ex.2* 0.399	*Ex.1* 0.5	*Ex.2* 0.461	*Ex.1* 0.5 $p^* = 2$
Cauchy: $C(x;0,1)$	0.318	0	0.274	0.318 $p^* = 1$
Contaminated Gaussian: $0.9\,N(x;0,1)$ $+0.1\,C(x;0,1)$	0.393	0	0.461	0.393 $p^* = 1$
Generalized Gaussian: $q = 1.5$	0.476	0.5	0.554	0.649 $p^* = 1.5$
Generalized Gaussian: $q = 4$	0.320	0.5	0.428	0.5 $p^* = 2$
Generalized Gaussian: $q = 100$	0.289	0.5	0.386	0.5 $p^* = 2$

The structure of the minimax L_{p^*}-norm detector is determined by the ratio $\bar{\sigma}^2/a^2$, and contrary to Huber's detector, the parameters $\bar{\sigma}^2$ and a^2 of class $\mathcal{F}_{12}$ can be directly estimated from the sample. In general, for estimating $\bar{\sigma}^2$ we can use, for example, the upper confidence limit for the estimate of variance, and for $1/(2a)$ the lower confidence limit for the nonparametric estimate of a distribution density at the center of symmetry: $\underline{f}(0) \le \hat{f}(0)$. Taking into account the L_{p^*}-norm form of the minimax detector with the parameter $1 \le p^* \le 2$ when both restrictions of class $\mathcal{F}_{12}$ hold as equalities, we choose the estimates of variance and density as the characteristics of this class. For a variance, this is the customary sample variance

$$\hat{\bar{\sigma}}^2 = N^{-1} \sum_{i=1}^{N} (x_i - \overline{x})^2.$$

To avoid the difficulties of nonparametric estimation of a distribution density, we estimate the value of the underlying density at its center of symmetry and the related parameter a using the following simple formula based on the central order statistics $x_{(k)}$ and $x_{(k+1)}$ ($N = 2k$ or $N = 2k+1$) [15]

$$\hat{a} = 1/[2\hat{f}(0)] = [(N+1)(x_{(k+1)} - x_{(k)})]/2.$$

On finite samples when $N = 20$ and $N = 100$, the performance of the L_{p^*}-norm, Huber's, L_1- and L_2-norm detectors under the Gaussian, Cauchy, ε-contaminated Gaussian and uniform noises was

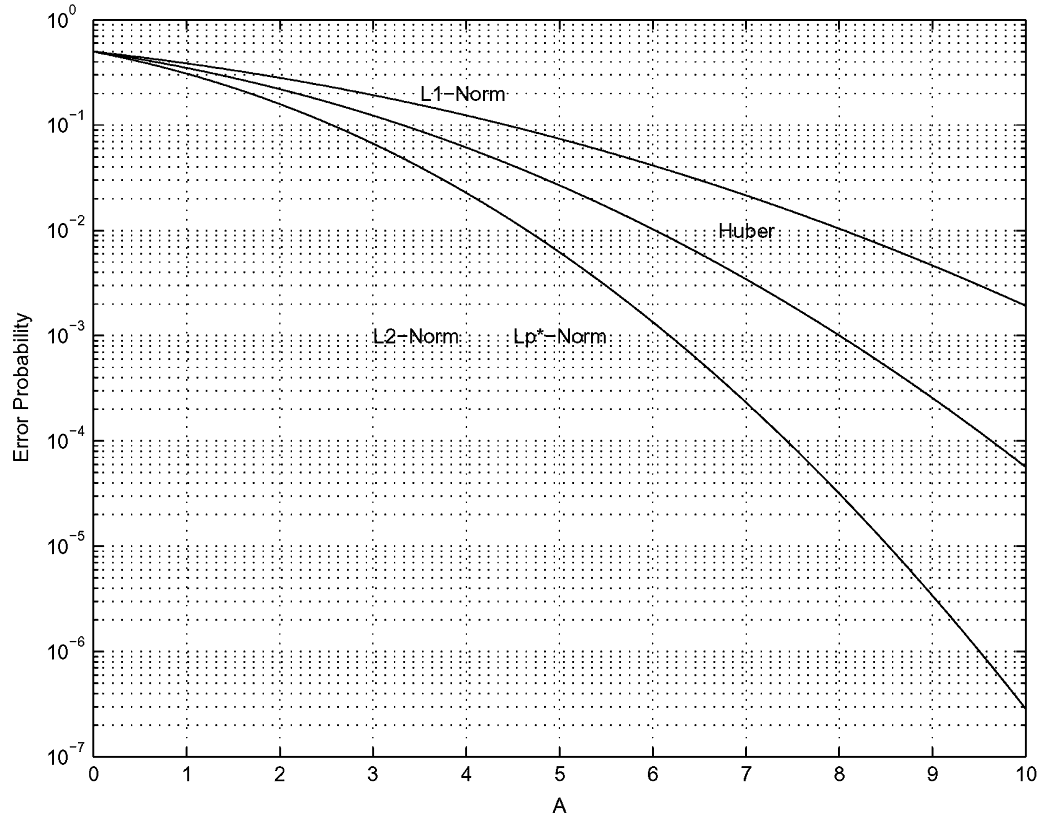

Fig. 2. Detection in the generalized Gaussian noise close to uniform: asymptotics, $q = 100, A = \sqrt{\mathrm{SNR}}$.

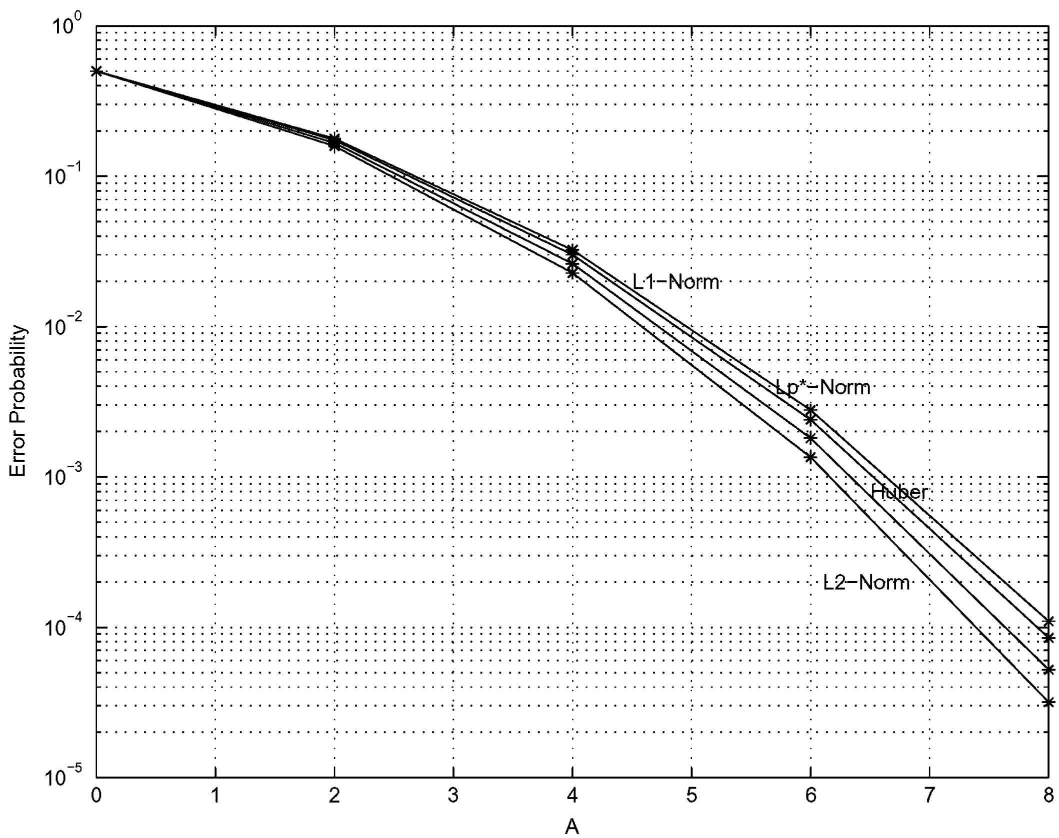

Fig. 3. Detection in the Gaussian noise: $N = 20, A = \sqrt{\mathrm{SNR}}$.

studied by Monte Carlo technique. The detection model was chosen consistently with the initial assumption of detection of a weak signal: $H_0 : x_i = n_i$ versus $H_1 : x_i = \theta + n_i, \; i = 1, \ldots, N$, where the useful signal $\theta = \theta_N = A/\sqrt{N}$.

On small samples with $N = 20$, the results of modeling in the Gaussian noise are displayed in Fig. 3; the other results are discussed below. On large samples with $N = 100$, the results of modeling are close to the asymptotic results given by (8).

IV. Concluding Remarks

The Gaussian Noise: Large Samples and Asymptotics. The minimax L_{p^*}-norm detector coincides with the optimal L_2-norm detector both being better than Huber's and the L_1-norm detector.

The Gaussian Noise: Small Samples. On the contrary, Fig. 3. shows that the minimax L_{p^*}-norm detector is close in performance to the robust L_1-norm detector being slightly inferior to Huber's on small samples.

The Sample Size Effects. There are some peculiarities of the dependence of the minimax detector on the sample size N, this deserves a separate attention. Since the estimators $\hat{\bar{\sigma}}^2$ and $\hat{a}$ of the parameters of class $\tilde{\mathcal{F}}$ are consistent, the sample minimax detection performance tends in probability to the exact minimax detection performance as $N \to \infty$. This is confirmed by Monte Carlo modeling on samples of size $N = 100$.

The results exhibited in Fig. 3. can be explained by the bias of the sample distribution of the threshold statistic $\hat{\bar{\sigma}}^2/\hat{a}^2$ with small N: its values determine the choice of the appropriate branch of the algorithm. For Gaussian samples of size $N = 20$, the choice of the L_2-norm occurs approximately for the 10% of cases ($P[\hat{\bar{\sigma}}^2/\hat{a}^2 < 2/\pi] \approx 0.1$), the choice of the L_1-norm detector—for the 20% of cases ($P[\hat{\bar{\sigma}}^2/\hat{a}^2 > 2] \approx 0.2$), and the L_{p^*}-norm detectors with $1 < p^* < 2$ are chosen for the rest 70% of cases with the average value of the power parameter $\bar{p^*} \approx 1.25$. For samples of size $N = 100$, we have the opposite situation: approximately for the 45% of cases, the L_2-norm branch of the minimax detector is realized, whereas the L_1-norm branch occurs only for the 5% with the average value of the power as $\overline{p^*} \approx 1.85$.

The Cauchy and ε-Contaminated Gaussian Noise. Though these models describe extremely heavy-tailed noises, nevertheless they deserve attention as, say, the Cauchy pdf may arise through the distribution of the ratio of Gaussian random variables. From Fig. 1 and Table I it can be seen that the L_2-norm detector with the linear test statistic naturally has the extremely poor performance both in asymptotics and on finite samples; the minimax, L_1-norm and Huber's detectors exhibit their good robust properties with the latter being evidently better than the former in the mixture models, and vice versa for the Cauchy noise. Note that we also examined the detection performance in other heavy-tailed distribution models, namely, the Laplace, the mixture models with the Gaussian contamination not so heavy as the Cauchy, and the generalized Gaussian pdfs with $1 < q < 2$, but the obtained results were qualitatively the same.

The Short-Tailed Noise. The detection performance again reverts not once in short-tailed noise models described by the generalized Gaussian pdf with $q = 100$ close to the uniform (see Fig. 2). In asymptotics, the L_2-norm and minimax detectors proved their superiority over Huber's and the L_1-norm detectors, but on small samples, the aforementioned small size sample effect reveals itself: the minimax detector is close in performance to the L_1-norm detector, and thus it is slightly inferior to Huber's. The qualitatively similar effects were also observed in some other examined short-tailed and finite pdf models, for example, for the generalized Gaussian pdf with $q = 4$.

Final Remark. Our main aim was to show some new possibilities of Huber's minimax approach to robust detection associated with the usage of a new class of densities with an upper-bounded variance. The proposed low-complexity minimax power detector exhibits both high robustness in heavy-tailed noise and good efficiency in short-tailed noise on small and large samples. The similar approach can be applied to much more complicated models of signal detection than the simple binary detection model analyzed in this paper.

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.

[2] R. J. Adler, R. E. Feldman, and M. S. Taqqu, *A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions*. Basel, Switzerland: Birkhauser, 1997.

[3] H. Delic and A. Hocanin, "Robust detection in DS-CDMA," *IEEE Trans. Veh. Technol.*, vol. 51, no. 1, pp. 155–170, Jan. 2002.

[4] H. Delic, P. Papantoni-Kazakos, and D. Kazakos, "Fundamental structures and asymptotic performance criteria in decentralized binary hypothesis testing," *IEEE Trans. Commun.*, vol. 43, no. 1, pp. 32–43, Jan. 1995.

[5] 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.

[6] O. Hössjer and M. Mettiji, "Robust multiple classification of known signals in additive noise—An asymptotic weak signal approach," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 594–608, Mar. 1993.

[7] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, pp. 73–101.

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

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

[10] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. Berlin, Germany: Springer-Verlag, 1988.

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

[12] D. Kazakos and P. Papantoni-Kazakos, *Detection and Estimation*. Rockville, MD: Computer Science, 1990.

[13] J. H. McCulloch, *Numerical Approximation of the Symmetric Stable Distribution and Density. A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Cambridge, MA: Birkhauser, 1998.

[14] B. T. Polyak and Ya. Z. Tsypkin, "Robust identification," in *Identif. Syst. Parameter Estim., Part 1, Proc. 4th IFAC Symp.*, Tbilisi, U.S.S.R., 1976, pp. 203–224.

[15] G. L. Shevlyakov and N. O. Vilchevski, *Robustness in Data Analysis: Criteria and Methods*. Utrecht, The Netherlands: VSP, 2002.

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

[17] N. O. Vilchevski and G. L. Shevlyakov, "Robust estimation of the shift parameter under a bounded noise variance," *Autom. Remote Control*, vol. 45, pp. 1048–1053, 1984.

[18] X. Wang and H. V. Poor, "Robust multiuser detection in nongaussian channels," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 289–305, Feb. 1999.