Robust Detection of Known Signals in Asymmetric Noise

SALEEM A. KASSAM, MEMBER, IEEE, GEORGE MOUSTAKIDES, STUDENT MEMBER, IEEE, AND JUNG GIL SHIN, MEMBER, IEEE

Abstract—**The detection of signals in noise with possibly asymmetric probability density functions is considered. The noise density model allows a symmetric contaminated-nominal central part and an arbitrary tail behavior. For detection of known signals, the robust nonlinear-correlator (NC) detector is obtained based on detector efficacy as performance criterion. The robust M-detector structure for constant-signal detection is also explicitly obtained.**

I. Introduction

Following the fundamental works of Huber on robust estimation [1] and robust hypothesis testing [2], many further developments and applications of robustness theory have been formulated by researchers in the communication sciences. Concepts of robustness in signal processing applications were certainly in existence prior to Huber's results (e.g., [3], [4]). However it is generally accepted that the techniques and results in [1], [2] formed an important basis for much of the considerable subsequent research activity on robust schemes for signal estimation, detection, and filtering applications. Two recent survey papers [5], [6] list a large number of references on robust techniques.

In [7] Huber's ideas were applied to obtain the structures of asymptotically robust signal detectors. This resulted specifically in the canonical *limiter-correlator* detector for a weak deterministic signal in nominally Gaussian noise modeled as having a mixture or contaminated probability density. In [8] this result was extended to apply to other nominal noise densities. Both [7] and [8] considered detection structures of the type where the sum of memoryless transformations of each discrete-time input observation (the test statistic) is compared to a fixed threshold. For example, the limiter-correlator robust detector for a signal vector $(s_1, s_2, \cdots, s_n)$ in an observation vector $(X_1, X_2, \cdots, X_n)$ with independent and identically distributed additive noise computes the test statistic $T_n = \sum_{i=1}^{n} L_i(X_i)$, where $L_i(X_i) = s_i \ell(X_i)$ and ℓ is a limiter characteristic. In general for arbitrary ℓ, we will call such detector structures *NC-detectors*, the test statistic being an instantaneous nonlinear transformation of the observation correlated with the signal. Note that this is the structure of a likelihood-ratio test on the X_i. More recently Huber's results were used in [9], [10] to obtain directly the robust *M-detectors* for both the fixed sample and sequential cases. An M-detector structure is obtained when the detection test statistic Q_n is obtained as that function of the observations minimizing $\sum_{i=1}^{n} M(X_i - s_i Q_n)$, where M is some appropriately chosen function. Note that Q_n may be used as an estimator for the signal amplitude, and such an estimator is called an M-estimator because of its similarity to maximum likelihood estimators in general.

Two major factors limit the applicability of such results for signal processing schemes. One of these is the requirement of independence for the sequence of discrete-time input data to the detectors. This requirement of independence has recently been addressed in [11], where it was shown that robust detector structures can be derived for operation under conditions of weak dependence in the input sequence. The results in [11] were developed for detection applications following similar considerations which had earlier been applied in [12], [13] for robust estimation. The second main limitation of many previous results on robust detection has been the assumption that the allowable noise density functions are symmetric. We will be concerned with this latter problem, and will develop the structures of the robust NC- and M-detectors for robust detection of weak deterministic signals with a noise model allowing asymmetry in the univariate noise density functions. Our study was largely motivated by some recent work on robust estimation with asymmetrically distributed noise [14], [15]; in particular we will adapt and draw upon the techniques and results in [15] for this work.

To introduce the asymmetric noise density class, let us recall some of the pertinent results on robust detection of weak deterministic signals. Consider the ϵ-contamination class $\mathcal{F}_{g,\epsilon}$ for noise densities f on the real line defined by

$$\mathcal{F}_{g,\epsilon} = \{f \mid f = (1-\epsilon)g + \epsilon h, \quad h \in \mathcal{H}\} \tag{1}$$

where g is a strongly unimodal symmetric nominal density function, and ϵ in $[0, 1)$ is a given maximum degree of contamination by an arbitrary density h in the class of all

Manuscript received October 6, 1980. This research was supported by the Naval Research Laboratory under contract N00014-78-C-0798 and by the Office of Naval Research under contract N00014-80-K-0945. This paper was presented at the 18th Annual Allerton Conference on Communication, Control, and Computing, Allerton House, Monticello, IL, Oct. 8–10, 1980.

S. A. Kassam is with the Department of Systems Engineering, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104.

G. Moustakides is with the Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ 08544.

J. G. Shin is with the Department of Electrical Engineering, Stevens Institute of Technology, Hoboken, NJ 07030.

0018-9448/82/0100-0084$00.75

bounded symmetric densities $\mathcal{H}$. The results in [1], [7], [8] show that a limiter characteristic $\ell = \ell_R$ exists which results in a robust NC-detector; specifically,

$$\ell_R(x) = \begin{cases} \dfrac{-g'(x)}{g(x)}, & |x| < a \\ \dfrac{-g'(a)}{g(a)} \operatorname{sgn}(x), & |x| > a, \end{cases} \tag{2}$$

where a is a positive constant depending on ϵ and g. The robustness of ℓ_R may be characterized by its property of being the optimum NC-detection characteristic for a least-favorable density $f_R \in \mathcal{F}_{g,\epsilon}$ in terms of performance measured by detection *efficacy* [16]. In addition the resulting detector can perform within a maximum false-alarm rate constraint depending on g and ϵ [8].

The efficacy $\mathcal{E}(f, \ell_R)$ for the above robust NC-detector with $L_i(x) = s_i \ell_R(x)$ and unit average signal power (a normalized efficacy) becomes

$$\mathcal{E}(f, \ell_R) = \frac{\left[\int_{-\infty}^{\infty} \ell_R'(x) f(x)\, dx\right]^2}{\int_{-\infty}^{\infty} \ell_R^2(x) f(x)\, dx}. \tag{3}$$

In (3), the numerator arises from the derivative with respect to amplitude of the mean of $\ell_R(X_i)$ and the denominator is the variance of $\ell_R(X_i)$. Even if h and thus f were not symmetric, the condition $\int_{-\infty}^{\infty} \ell_R(x) f(x)\, dx = 0$ on allowable noise densities f, instead of symmetry, would also lead to the conclusion that ℓ_R is robust. Since g is a nominal symmetric density, this means that the class $\mathcal{H}$ in (1) can be enlarged. A simpler extension of the class $\mathcal{F}_{\epsilon, g}$ for which robustness of ℓ_R also holds gives the class

$$\tilde{\mathcal{F}}_{g,\epsilon} = \{ f \mid f = (1-\epsilon) g + \epsilon h, \quad h \in \tilde{\mathcal{H}} \} \tag{4}$$

with $\tilde{\mathcal{H}}$ the class of all bounded densities h which are symmetric on $[-d, d]$ and have equal tail probabilities on $(-\infty, d)$ and (d, ∞), with $d \geq a$. Note that a is a positive constant which depends on g and ϵ. This extension is also applicable for the results on M-detectors in [9], [10].

Although the class $\tilde{\mathcal{F}}_{g,\epsilon}$ is a class of densities which are not necessarily symmetric, they are nevertheless symmetric in the middle. This is usually a satisfactory assumption. The degree of tail asymmetry is controlled in two ways; there is still an underlying nominal component $(1-\epsilon)g$, and the contaminations h are zero median. The noise model in [15] removes these last assumptions and allows f to be essentially arbitrary outside some interval $[-d, d]$. We will obtain results on robust detection using such a model. Specifically we consider the class $\hat{\mathcal{F}}_{g,\epsilon,d}$ of noise densities f given by

$$\hat{\mathcal{F}}_{g,\epsilon,d} = \left\{ f \mid f = \begin{cases} (1-\epsilon) g + \epsilon h, \text{ on } [-d, d], & h \in \mathcal{H} \\ \text{arbitrary, outside } [-d, d]. \end{cases} \right. \tag{5}$$

Here $\epsilon \in (0, 1)$ is the maximum degree of contamination of a nominal density function g. The density g is assumed to be strongly unimodal (i.e., $-\log g$ is convex) and symmetric, and in addition we assume it to be sufficiently regular so that g' is absolutely continuous. The parameter d is a positive parameter specifying the interval around the origin in which the noise density f is a bounded, symmetric, contaminated version of g. The class $\mathcal{H}$ is the class of all bounded, symmetric density functions. Thus on $[-d, d]$ all $f \in \hat{\mathcal{F}}_{g,\epsilon,d}$ are bounded and symmetric. Note that a valid $f \in \hat{\mathcal{F}}_{g,\epsilon,d}$ could be zero on $(-\infty, -d)$, and place a probability of $2(1-\epsilon)[1 - G(d)] + \epsilon$ on (d, ∞), where G is the distribution function corresponding to g. If g is the zero-mean Gaussian density with variance σ^2, a reasonable specification of d may be a number between 2σ and 4σ, and ϵ is typically between 0.001 and 0.1.

In the next sections we will consider the robust NC- and M-detectors for the noise model of (5). In the next section, the results in [15] are applied to obtain the robust NC-detector nonlinearity with performance characterized by detection efficacy. In Section III results on robust M-detection are obtained. These latter results are more significant, in spite of the limitation to constant signal detection, because they yield a stronger statement about performance of the robust M-detector. Specifically the performance index used (asymptotic variance) implies that both detection probability and false-alarm probability characteristics are taken into account [9]. The robust NC-detector performance is characterized by efficacy alone so that if the false-alarm probability of the detector cannot be maintained at the design value, then relative efficiencies or detection probability comparisons cannot follow directly.

We are concerned with the asymptotic theory of robust detection. The results are applicable in practice to situations where sample sizes are large and, for NC-detectors, under the additional constraint of low signal strength. Obviously when the sample size is small (of the order of five or ten) actual detection performance may not be reasonably predicted from such asymptotic results. On the other hand, previous studies [9], [17] on the type of detector nonlinearities we consider show that in many cases for moderate sample sizes (of the order of 50) asymptotic performance is a good indication of actual performance. It should be noted, however, that in some cases, convergence to asymptotic characteristics may be quite slow.

II. Robust NC-Detector for Asymmetric Noise Densities

For a vector of observations $(X_1, X_2, \cdots, X_n)$ of length n described by

$$X_i = N_i + \theta s_i, \qquad i = 1, 2, \cdots, n \tag{6}$$

where $(s_1, s_2, \cdots, s_n)$ is a deterministic signal vector and the N_i are independent and identically distributed noise components, we want to test the null hypothesis H_0 that $\theta = 0$ versus the alternative H_1 that $\theta > 0$. For an NC-detector using test statistic

$$T_n = \sum_{i=1}^{n} s_i \ell(X_i), \tag{7}$$

we want to obtain the characteristic ℓ which results in a robust detector for allowable univariate noise density functions f in the class $\hat{\mathfrak{F}}_{g,\epsilon,d}$. As a criterion of performance, we will use the detector efficacy $\mathcal{E}(f,\ell)$ which is dependent on f and ℓ, and defined as

$$\mathcal{E}(f,\ell)=\lim_{n\to\infty}\frac{\left[\frac{d}{d\theta}E_\theta\{T_n\}\Big|_{\theta=0}\right]^2}{\operatorname{var}_\theta\{T_n\}\,|_{\theta=0}}. \tag{8}$$

This is an asymptotic measure of detector performance applicable in cases where the sample size is large and signal amplitude is small. We will not here impose a constraint on the false-alarm or type 1 error probability, which would lead to a stronger robustness property as in [7], [8]. Thus it is implicitly assumed that the detector threshold can be adjusted to always obtain the desired size for the detector for any $f\in\hat{\mathfrak{F}}_{g,\epsilon,d}$. Under this condition, the detector efficacy is directly related to the slope of the detector power function at $\theta=0$. We will comment further on the false alarm probability constraint for the robust NC-detector at the end of this section.

It is clear that for an NC-detector to be consistent for all $f\in\hat{\mathfrak{F}}_{g,\epsilon,d}$, the characteristic ℓ has to vanish outside $[-d,d]$. Since $\hat{\mathfrak{F}}_{g,\epsilon,d}$ consists of densities symmetric on $[-d,d]$, we additionally require that the allowable ℓ are symmetric. Let L_c denote the class of all NC-detector characteristics ℓ satisfying

1) $\ell(x)=0,\ |x|\geq c$,
2) $\ell(x)=-\ell(-x)$,
3) ℓ is absolutely continuous,

the parameter c being a nonzero cutoff value $c\leq d$; the value of c is set by consistency requirements, as we will discuss soon.

A. Solution for Efficacy-Robust Detector

We want to find a least-favorable density f_R in $\hat{\mathfrak{F}}_{g,\epsilon,d}$ and a corresponding optimum characteristic ℓ_R in L_c such that

$$\inf_{f\in\hat{\mathfrak{F}}_{g,\epsilon,d}}\mathcal{E}(f,\ell_R)=\mathcal{E}(f_R,\ell_R). \tag{9}$$

Note that we will then have

$$\mathcal{E}(f_R,\ell_R)=\sup_{\ell\in L_c}\mathcal{E}(f_R,\ell) \tag{10}$$

since we require ℓ_R to be optimum for f_R.

The following theorem establishes the condition under which a pair (f_R,ℓ_R) can be found in $\hat{\mathfrak{F}}_{g,\epsilon,d}\times L_c$, satisfying (9) and (10) with a finite nonzero value $\mathcal{E}(f_R,\ell_R)$. It is directly related to Theorem 3.1 in [15].

Theorem 1: If the condition

$$\epsilon<(1-\epsilon)\{2cg(0)-[2G(c)-1]\} \tag{11}$$

is satisfied where G is the distribution function corresponding to g, the density function

$$f_R(x)=\begin{cases}(1-\epsilon)g(x), & |x|\leq a_0\\ \dfrac{(1-\epsilon)g(a_0)}{\cosh^2\left[\frac{1}{2}a_1(c-a_0)\right]}\cosh^2\left[\frac{1}{2}a_1(c-|x|)\right], & a_0<|x|\leq c\\ \text{arbitrary}, & |x|>c\end{cases} \tag{12}$$

and the corresponding optimum characteristic in L_c

$$\ell_R(x)=\begin{cases}\dfrac{-f_R'(x)}{f_R(x)}, & |x|\leq c\\ 0, & |x|>c\end{cases} \tag{13}$$

satisfy (9) and (10), with $0<a_0<c$ and $(-g'(a_0))/(g(a_0))<a_1$ being the unique solutions of

$$\epsilon=\int_{-c}^{c}f_R(x)\,dx-(1-\epsilon)\int_{-c}^{c}g(x)\,dx \tag{14}$$

and

$$\frac{-g'(a_0)}{a_1g(a_0)}=\tanh\left[\frac{1}{2}a_1(c-a_0)\right]. \tag{15}$$

Proof of Theorem 1: The existence of the density function $f_R\in\hat{\mathfrak{F}}_{g,\epsilon,d}$ is established in Appendix I.

The efficacy $\mathcal{E}(f,\ell)$ can be written as

$$\mathcal{E}(f,\ell)=\frac{\left[\int_{-c}^{c}\ell'(x)f(x)\,dx\right]^2}{\int_{-c}^{c}\ell^2(x)f(x)\,dx}, \tag{16}$$

and also

$$\mathcal{E}(f_R,\ell)=\frac{\left[\int_{c}^{-c}\ell(x)f_R'(x)\,dx\right]^2}{\int_{-c}^{c}\ell^2(x)f_R(x)\,dx}, \tag{17}$$

since ℓ and f_R are absolutely continuous; we have assumed that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}s_i^2=1.$$

From (17) the optimality of ℓ_R in (10) follows from the Schwarz inequality.

Now $\mathcal{E}(f,\ell_R)$ of (16) is a convex function of f ([1], Lemma 6), and $\hat{\mathfrak{F}}_{g,\epsilon,d}$ is a convex set of density functions. Let $\mathcal{E}_R(\gamma)\triangleq\mathcal{E}([(1-\gamma)f_R+\gamma f],\ell_R)$. Then to show (9), it is sufficient to show that [18]

$$\mathcal{E}_R'(\gamma)\,|_{\gamma=0}\geq 0,\qquad f\in\hat{\mathfrak{F}}_{g,\epsilon,d} \tag{18}$$

which is equivalent to the condition

$$\int_0^c[2\ell_R'(x)-\ell_R^2(x)][f(x)-f_R(x)]\,dx\geq 0,\qquad f\in\hat{\mathfrak{F}}_{g,\epsilon,d}. \tag{19}$$

TABLE I
UPPER BOUND ON ϵ AS FUNCTION OF c FOR WHICH (11) HOLDS, FOR UNIT-VARIANCE GAUSSIAN DENSITY g

c	ε (c)
1.0	0.103
1.5	0.248
2.0	0.391
2.5	0.502
3.0	0.583

Now

$$a_1^2 + 2\ell_R'(x) - \ell_R^2(x) \geq 0, \qquad 0 \leq x \leq a_0,$$

because $a_1 = \ell_R(a_0) \geq \ell_R(x)$ and $\ell_R'(x) \geq 0$ for $x \in [0, a_0]$. Also by direct evaluation, we have

$$a_1^2 + 2\ell_R'(x) - \ell_R^2(x) = 0, \qquad a_0 < x \leq c.$$

In addition, $f(x) - f_R(x) \geq 0$ for $x \in [0, a_0]$ and $\int_0^c \{f(x) - f_R(x)]\, dx \leq 0$ because $\int_0^c f(x)\, dx \leq (1-\epsilon)\int_0^c g(x)\, dx + \epsilon/2 = \int_0^c f_R(x)\, dx$, from (14). From these, (19) follows directly.

Comments on Theorem 1: If condition (11) is not satisfied, then f_R can be picked to be a constant on $[-c, c]$ so that $\mathcal{E}(f_R, \ell) = 0$. In ([15], Table I) and in Table I, some numerical values are given for the upper bounds $\epsilon(c)$ on ϵ, for different c, for which (11) is satisfied with g the unit variance Gaussian density. The value of $\epsilon(c)$ increases with c, and $\epsilon(c)$ is larger than 0.1 if c is larger than unity. Our proof of Theorem 1 is a direct extension of the proof of Theorem 3.1 in [15], which placed more restrictive conditions on the allowable density functions and estimator characteristics. Our criterion of detection efficacy is used as an estimator criterion in [15]; however its interpretation as an estimation variance for the statistics in [15] requires further assumptions which are not simply characterized as conditions on the allowable density functions. We will elaborate on this in the next section.

Fig. 1 is a sketch of a typical ℓ_R function for the Gaussian nominal density g.

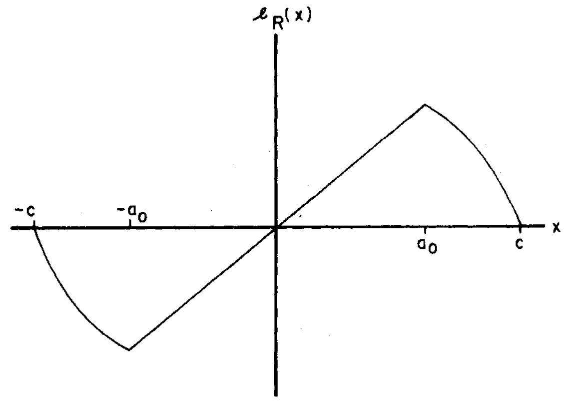

Fig. 1. Typical robust detector characteristic for Gaussian nominal noise.

B. Consistency of Robust Detector and Performance Evaluation

The restriction that $\ell(x)$ be zero outside $[-d, d]$ is required in order that with a fixed structure, the type I error probability of the detector approaches zero as $n \to \infty$. The solution for the robust detector nonlinearity ℓ_R has been obtained for $c \leq d$. It is clear that the worst-case performance of the detector in terms of efficacy will improve as the value of c gets closer to d. However, while detector performance for vanishing signal strengths may be important, thus justifying the use of efficacies, we should have at the very least a *consistent* test so that the type II error probability approaches zero as $n \to \infty$ for all values of signal strength θ in an interval of interest.

For $c \leq d$, it can be easily shown that ℓ_R of (13) results in a test which is consistent for θ in a positive neighborhood of the origin. To see this, it is sufficient to show that the slope at $\theta = 0$ of the mean function of ℓ_R given by

$$m'(0) = \int_{-c}^{c} \ell_R'(x) f(x)\, dx,$$

is positive for all $f \in \mathfrak{F}_{g,\epsilon,d}$. Now under the condition of Theorem 1, $\mathcal{E}(f_R, \ell_R)$ is positive, and $\int_{-c}^{c} \ell_R'(x) f_R(x)\, dx$ is positive. Since $\hat{\mathfrak{F}}_{g,\epsilon,d}$ is a convex set, it follows that $m'(0) > 0$ for all $f \in \hat{\mathfrak{F}}_{g,\epsilon,d}$.

Even if $m'(0)$ is positive, the mean function $m(\theta) = \int_{-c}^{c} \ell_R(x) f(x-\theta)\, dx$ will become zero for some positive value $\theta_{\max}$ of θ. In Table II we show the computed values of $\theta_{\max}$ as a function of c and ϵ, for the case $c = d$, with g again the unit-variance Gaussian density. These values were obtained by considering for each θ the noise density minimizing $m(\theta)$. These results indicate that for given ϵ, increasing d with $c = d$ leads to a consistent test for an increasing range of values of θ; in the limit we get consistency for all θ and the solution degenerates to that in [7], [8].

If the limiter-correlator detector [7], [8] is used when asymmetry is present, it is clear that detector performance can degrade rapidly with increasing n because of the bias introduced. The robust *NC*-detector derived here is clearly unbiased when signal amplitude θ is bounded by $d - c$, and in fact has been shown above to remain unbiased for θ bounded by a positive $\theta_{\max}$ even when $c = d$ for the Gaussian example considered.

In this section we have been able to establish robustness of the *NC*-detector using ℓ_R with performance being defined by detector efficacy. It has not been possible to prove a stronger result which simultaneously bounds the worst-case asymptotic local slope of the power function and the false-alarm probability as in [7], [8], where the fixed threshold of the robust *NC*-detector could also be determined. If it is assumed that the threshold can always be adjusted (adaptively) to maintain a fixed false-alarm probability for all $f \in \hat{F}_{g,\epsilon,d}$, then the efficacy of the detector is directly related to the slope of its power function at $\theta = 0$.

TABLE II
UPPER BOUND $\theta_{\max}$ ON θ AS FUNCTION OF c, ϵ (WITH $c = d$) FOR WHICH ROBUST NC-DETECTOR IS CONSISTENT (UNIT-VARIANCE GAUSSIAN g)

ϵ \ $c=d$	1.0	1.5	2.0	2.5	3.0
0.05	0.0	1.7	2.7	3.5	4.0
0.10	0.0	0.9	2.4	3.0	3.5
0.15	---	0.0	2.0	2.7	3.2
0.20	---	0.0	1.5	2.3	2.8
0.25	---	---	1.2	1.9	2.4
0.30	---	---	0.9	1.6	2.1
0.35	---	---	0.5	1.3	1.8
0.40	---	---	---	0.8	1.5

The reason why the stronger result is not possible here is that ℓ_R is not a monotone limiter characteristic, so that the numerator and denominator in (16) are not separately minimized and maximized, respectively, by f_R when $\ell = \ell_R$. Nevertheless the results of this section are of interest because detection efficacy is a generally accepted measure of differential signal-to-noise ratio and has been extensively applied in robustness studies. In the next section we consider an M-detector robust structure which allows a stronger robustness property to be derived.

III. ROBUST M-DETECTOR FOR ASYMMETRIC NOISE DENSITIES

The test statistic Q_n for an M-detector is a maximum-likelihood-type estimator for the signal parameter θ. In general if M is any reasonable function, e.g., $M = -\log f$ for some density function f, an M-detector test statistic can be defined as the function Q_n minimizing $\sum_{i=1}^{n} M(X_i - s_i Q_n)$. If M is sufficiently regular Q_n can also be defined as the statistic satisfying

$$\sum_{i=1}^{n} s_i \psi(X_i - s_i Q_n) = 0, \tag{20}$$

the function ψ being the derivative of M. In this section (20) will be taken to define the test statistic of an M-detector, which will therefore be characterized by the function ψ. We will also restrict attention to the special case of *constant* signals, and without further loss of generality we will take $s_i = 1$, $i = 1, 2, \cdots, n$. The general case of nonconstant signals requires further considerations, as we will indicate later. We will therefore be focusing on the special case of (6), giving X_i as

$$X_i = N_i + \theta, \qquad i = 1, 2, \cdots, n, \tag{21}$$

and our M-detectors will be based on statistics Q_n satisfying

$$\sum_{i=1}^{n} \psi(X_i - Q_n) = 0. \tag{22}$$

We are interested in the class $\hat{\mathfrak{F}}_{g,\epsilon,d}$ (or a useful subset of the class $\hat{F}_{g,\epsilon,d}$) of densities, and we impose the following reasonable constraints in defining the class Ψ_c of allowable M-detector functions ψ we will consider:

1) $\psi(x) = 0$, $|x| \geq c$,
2) $\psi(x) = -\psi(-x)$,
3) $\psi'(x)$ is bounded and piecewise-continuous on $[-c, c]$.

An additional constraint will be added soon. The value of the parameter $c < d$, which defines the size of Ψ_c, is set by requirements we consider next.

To complete the specification of our class of M-detectors a solution scheme has to be specified for obtaining Q_n satisfying (22). This is necessary because a solution to (22) may not be unique. The scheme we specify is a simpler iterative procedure than that considered in [15]. Although its numerical convergence rate to the solution may be somewhat smaller, it allows us to obtain more explicit robustness results than were obtained in [15] (specifically note [15, Remark 3.5]).

We define the test-statistic Q_n of an M-detector based on $\psi \in \Psi_c$ in terms of the sequence $\{Q_n^j\}$ given by

$$Q_n^{j+1} = Q_n^j + \frac{\lambda_n(Q_n^j)}{D}, \qquad j = 0, 1, 2, \cdots, \tag{23}$$

where

$$\lambda_n(q) = \int_{-\infty}^{\infty} \psi(x - q)\, dF_n(x), \tag{24}$$

F_n being the empirical distribution function of the n observations, with D being a positive constant. The solution Q_n is defined as $Q_n = \lim_{j\to\infty} Q_n^j$, with Q_n^0 a sample median (or any consistent estimator of the median of the X_i), provided the limit exists. Otherwise Q_n is taken to be Q_n^0.

In terms of the quantities g, ϵ, and d defining $\hat{\mathfrak{F}}_{g,\epsilon,d}$, we now define a new parameter k_0 by

$$k_0 = G^{-1}\left[\frac{1}{2} + \frac{\tau}{2(1-\epsilon)}\right] \tag{25}$$

with

$$\tau = 2(1-\epsilon)[1 - G(d)] + \epsilon, \tag{26}$$

or directly as

$$k_0 = G^{-1}\left[1 - G(d) + \frac{1}{2(1-\epsilon)}\right], \tag{27}$$

G being the distribution function corresponding to the density g. Note that as indicated in the Section I, τ is the maximum value of the total probability which may be arbitrarily distributed outside $[-d, d]$. We will assume that our model is such as to make $\tau < 0.5$, so that the median of any $f \in \hat{\mathfrak{F}}_{g,\epsilon,d}$ lies in $(-d, d)$. This implies the restriction $\epsilon < 0.5$. From (26) we have

$$d = G^{-1}\left[\frac{1}{2} + \frac{1-\tau}{2(1-\epsilon)}\right],$$

so that $k_0 < d$. Under this assumption ($\tau < 0.5$) the maximum value of $|m|$, where m is the median of $f \in \hat{\mathfrak{F}}_{g,\epsilon,d}$, satisfies

$$(1-\epsilon)[G(|m|) - G(-d)] = \frac{1}{2}$$

so that

$$G(|m|) = 1 - G(d) + \frac{1}{2(1-\epsilon)},$$

and thus the median is always in $[-k_0, k_0]$, from (27).

Our objective now is to establish conditions for the asymptotic normality of Q_n based on $\psi \in \Psi_c$ for $f \in \hat{\mathfrak{F}}_{g,\epsilon,d}$. We will finally be able to define subsets of Ψ_c and $\hat{\mathfrak{F}}_{g,\epsilon,d}$ over which asymptotic normality holds, and we will obtain the saddlepoint solution for the asymptotic variance over these classes.

We consider first the consistency of Q_n.

Lemma 1: Let $c \le d - k_0$ and ψ be any characteristic in Ψ_c. Suppose that for a given $f \in \hat{\mathfrak{F}}_{g,\epsilon,d}$ the median lies in the open interval $(-k_0, k_0)$ and the function

$$\lambda(q) = \int_{-c+q}^{c+q} \psi(x-q) f(x)\, dx \tag{28}$$

is strictly decreasing on $[-k_0, k_0]$. Then for $D > (1/2)\max_{[-c,c]}|\psi'(x)|$ in (23), the M-detector based on ψ has for this f a test statistic Q_n which converges in probability to θ; in addition,

$$\lim_{n\to\infty} P\left\{\sum_{i=1}^{n} \psi(X_i - Q_n) = 0\right\} = 1. \tag{29}$$

The proof of Lemma 1 is given in Appendix II. Note that this lemma implies that the iterations for Q_n in (23) will converge (and therefore Q_n is not defined as Q_n^0) with a probability approaching unity as $n \to \infty$.

The following lemma is concerned with asymptotic normality of Q_n.

Lemma 2: Let (ψ, f) be a pair in $\Psi_c \times \hat{\mathfrak{F}}_{g,\epsilon,d}$, with $c \le d - k_0$, for which the conditions of Lemma 1 are satisfied. If in addition we have $\lambda'(q) < 0$ in a closed neighborhood of the origin, then $\sqrt{n}(Q_n - \theta)$ is asymptotically normally distributed with variance $V(f,\psi)$ given by

$$V(f,\psi) = \frac{1}{\mathcal{E}(f,\psi)} = \frac{\int_{-c}^{c} \psi^2(x) f(x)\, dx}{\left[\int_{-c}^{c} \psi'(x) f(x)\, dx\right]^2}. \tag{30}$$

Lemma 2 follows from results in ([19, Section IV]) where general conditions are given ensuring asymptotic normality of M-estimators. That these conditions are met under Lemma 2 can be easily demonstrated.

We are now ready to obtain a least favorable density and corresponding optimum M-detector characteristic which together form a saddlepoint for performance in terms of asymptotic variance of a consistent and asymptotically normal detection test statistic. First we define $\psi_R \in \Psi_c$ by $\psi_R(x) = \ell_R(x)$ of (13) where $f_R(x)$ was defined in (12). Then we have the following.

Lemma 3: Let g, ϵ, d be such that $\tau < 1/2$ in (26) and k_0 of (27) is less than $a_0/2$, for $c \le d - k_0$ satisfying (11). [The parameter a_0 is present in the definition of $\psi_R = \ell_R$ in (13).] Then for $f = f_R$ and $\psi = \psi_R$, the test statistic Q_n obtained from (23) with $D > (1/2)\max_{[-c,c]}|\psi_R'(x)|$ is consistent for θ, and $\sqrt{n}(Q_n - \theta)$ is asymptotically normal with variance

$$V(f_R, \psi_R) = \frac{\int_{c}^{-c} \psi_R^2(x) f_R(x)\, dx}{\left[\int_{-c}^{c} \psi_R(x) f_R'(x)\, dx\right]^2}. \tag{31}$$

Proof of Lemma 3: Under condition (11) f_R of (12) is strictly unimodal on $[-d, d]$, by which we mean that the symmetric function f_R is decreasing on $(0, d]$. Now $\psi_R = \ell_R$ of (13) is positive on $(0, c)$ and strictly increasing on $(-a_0, a_0)$. From this it follows that

$$\lambda_R(q) = \int_{-c+q}^{c+q} \psi_R(x-q) f_R(x)\, dx = \int_{-c}^{c} \psi_R(x) f_R(x+q)\, dx$$

is a decreasing function of q on $(-a_0/2, a_0/2)$, and therefore on $[-k_0, k_0]$. Also $\lambda_R'(q) = -\int_{-c}^{c}\psi_R'(x) f_R(x+q)\, dx$ is negative and continuous at $q = 0$. We also have $\lambda_R'(0) = \int_{-c}^{c}\psi_R(x) f_R'(x)\, dx$. The result then follows from Lemmas 1 and 2.

Let $\mathfrak{F}^*_{g,\epsilon,d} \subset \hat{\mathfrak{F}}_{g,\epsilon,d}$ be the subset of densities in $\hat{\mathfrak{F}}_{g,\epsilon,d}$ which are strictly unimodal on $[-d, d]$. In addition the parameters ϵ and d are restricted to satisfy the condition $\tau < 0.5$, with τ defined in (26), and the medians are assumed to lie in $(-k_0, k_0)$, with k_0 defined in (27). This last condition requires positive probabilities to be distributed both on $(-\infty, -d)$ and on (d, ∞), and is an insignificant restriction. Finally with $c = d - k_0$, the parameter a_0 defining f_R in (12) is assumed to be larger than $2k_0$ and also (11) is assumed to be satisfied. It is easy to show that these conditions are satisfied for reasonable choices of g, ϵ, and d. For example, let g be the zero-mean, unit-variance Gaussian density, let $\epsilon = 0.05$ and $d = 2$. Then we have $\tau = 0.093$ and $k_0 = 0.123$. With $c = d - k_0 = 1.877$ (11) is satisfied, and the value of $a_0/2$ is 0.502.

We finally restrict consideration to the subset $\psi_c^* \subset \Psi_c$, containing M-detector characteristics which are nonnegative on $(0, c)$ in addition to satisfying the three conditions defining ψ_c. This is a reasonable restriction in view of the strict unimodality restriction used in defining $\mathfrak{F}^*_{g,\epsilon,d}$.

Theorem 2: a) Let $\psi_R \in \Psi_c^*$ be defined by $\psi_R = \ell_R$ of (13), and let Q_n be the test statistic arising from the M-detector based on ψ_R, with $D > (1/2)\max_{[-c,c]}|\psi_R'(x)|$. Then for any $f \in \mathfrak{F}^*_{g,\epsilon,d}$, Q_n is a consistent estimator for θ; $\sqrt{n}(Q_n - \theta)$ is asymptotically normal with variance

$V(f, \psi_R)$ satisfying

$$\max_{f \in \mathfrak{F}^*_{g,\epsilon,d}} V(f, \psi_R) = V(f_R, \psi_R), \tag{32}$$

where $V(f_R, \psi_R)$ was defined in (31).

b) For any $\psi \in \Psi^*_c$ which gives a consistent statistic for all $f \in \mathfrak{F}^*_{g,\epsilon,d}$, we have

$$V(f_R, \psi_R) \le V(f_R, \psi) \tag{33}$$

where $V(f_R, \psi)$ is the asymptotic variance of the normalized statistic based on ψ.

Proof of Theorem 2: a) This follows by extending the proof of Lemma 3 for any strictly unimodal density, which shows that $1/V(f, \psi_R)$ is the efficacy $E(f, \psi_R)$ minimized in Theorem 1.

b) To prove (33), we only have to show that $V(f_R, \psi)$ is obtained by replacing ψ_R with ψ in (31). This follows from the fact that for any $\psi \in \Psi^*_c$ which is not identically zero we have $\lambda'(q) < 0$ in a closed neighborhood of the origin (see Proof of Lemma 3). The other conditions (including consistency) in ([19, Section IV]) being true, asymptotic normality and the formula for the asymptotic variance follow.

Theorem 2 provides the main result of this section. We started from classes $\hat{F}_{g,\epsilon,d}$ and Ψ_c of densities and detector characteristics, respectively, but our least-favorable density and robust characteristic were obtained from amongst functions in $\mathfrak{F}^*_{g,\epsilon,d}$ and Ψ^*_c, respectively. The essential restriction added was that the densities considered be strictly unimodal on $[-d, d]$; this is not unreasonable because the nominal density g is assumed to have this property. In applying the robust statistic Q_n arising from ψ_R in signal detection, one has to set a threshold based on the maximum variance $V(f_R, \psi_R)$. Thus the false-alarm probability constraint is automatically satisfied. In addition for any $f \in \mathfrak{F}^*_{g,\epsilon,d}$, the asymptotic power function or the slope of the power function at the origin can be lower bounded by the corresponding values for $f = f_R$, depending on specific conditions on the signal strength parameter and detection threshold values. The details can be found in [9]; the main condition has been proved in Theorem 2 (specifically the condition in [9, Lemma 2]).

The major reason why we confined attention to the constant-signal case in this section is that we need a reasonable initial value (e.g., the median Q_n^0) in starting the iterations in (23), to guarantee a consistent statistic Q_n. Extension of our results to the general case seems possible, and would appear to require an initial estimate of the median based on some nonparametric or other simple regression procedure [20], [21]. The consistency proofs and conditions would also have to be extended.

IV. Conclusion

We have derived the structures of robust NC- and M-detectors for known-signal detection in noise for which the probability density has a symmetric, contaminated central part and arbitrary tail behavior. This model has been used previously for robust estimation studies.

The robust NC-detector was derived for performance defined by detection efficacy, a weak-signal large-sample-size asymptotic performance measure. Although the detection efficacy is directly related to the slope of the detector power function, it was not possible to obtain simultaneous control on the false-alarm probability.

For constant-signal detection, the robust M-detector characteristic was obtained for performance characterized by the asymptotic variance of the test statistic. This result for the asymptotic variance together with earlier results on M-detectors, allow more interesting robust detection solutions which can maintain the false-alarm probability within desired upper bounds. Our robustness results differ from previous results on robust estimation which also examined asymptotic variance in that we considered a simpler solution strategy and obtained the explicit saddlepoint solution for well-defined classes of noise and detector characteristics.

Appendix I
Existence of Solution in Theorem 1

The existence of solutions for (14) and (15) has been essentially proved in Theorem 3.1 of [15]. For any $a_0 \in (0, c)$, the left side of (15) is positive and decreasing from unity to zero as a_1 increases from $-g'(a_0)/g(a_0)$ to ∞. The right side increases at the same time from $\tanh[-1/2(c - a_0)g'(a_0)/g(a_0)]$ to ∞. Note that the tanh function is less than unity for finite arguments. For given $a_0 \in (0, c)$, defining f_R as in (12) with $a_1 > -g'(a_0)/g(a_0)$ the unique solution of (15), the right side of (14) is continuous in a_0 and increases from zero to the right side of (11) as a_0 decreases from c to zero.

To show that $f_R \in \hat{\mathfrak{F}}_{g,\epsilon,d}$, it is sufficient to show that $h_R \triangleq [f_R - (1-\epsilon)g]/\epsilon$ is nonnegative on (a_0, c). Since g is strongly unimodal, so that $g(x) \le g(a_0)e^{-k(x-a_0)}$, $x \ge a_0$, with $k = -g'(a_0)/g(a_0)$, we consider the difference

$$\Delta(x) = \frac{\cosh^2\left[\frac{1}{2}a_1(c-x)\right]}{\cosh^2\left[\frac{1}{2}a_1(c-a_0)\right]} - e^{-k(x-a_0)}.$$

To show that $\Delta(x) \ge 0$, we prove that for $x \ge a_0$,

$$e^{(k/2)(x-a_0)} \frac{\cosh\left[\frac{a_1}{2}(c-x)\right]}{\cosh\left[\frac{a_1}{2}(c-a_0)\right]} \ge 1.$$

Now the left side of the above inequality is unity when $x = a_0$, and its derivative with respect to x has the same sign as that of $\{k - a_1 \tanh[a_1/2(c-x)]\}$, which is positive for $x > a_0$. Thus f_R of (12) is in $\hat{\mathfrak{F}}_{g,\epsilon,d}$.

Appendix II
Proof of Lemma 1

We assume $\theta = 0$ without loss of generality. Note that $\lambda(0) = 0$, and $\lambda(q) = -\lambda(-q)$ for $|q| \le k_0$. The equation $\lambda(q) = 0$ has only one root, $q = 0$, in $[-k_0, k_0]$. Consider the iterative scheme

$$q^{j+1} = q^j + \lambda(q^j)/D,$$

with $q^0 \in [-k_0, k_0]$. If $\lambda'(q) > -2D$ for $0 \le q \le k_0$ then $q^j \to 0$ as $j \to \infty$. This is because with $h(q)$ defined as

$$h(q) = q + \lambda(q)/D,$$

we then have $|h(q)| < |q|$ for $0 < |q| \le k_0$ from which the result follows. The condition $\lambda'(q) > -2D$ for $0 \le q \le k_0$ follows if $D > 1/2 \max_{[-c,c]} |\psi'(x)|$.

It can be shown—by extending the proof in [15] because here ψ' is only piecewise continuous—that

$$\sup_{|q| \le k_0} |\lambda_n(q) - \lambda(q)| \to 0 \quad \text{in probability,} \tag{B1}$$

and

$$\sup_{|q| \le k_0} |\lambda'_n(q) - \lambda'(q)| \to 0 \quad \text{in probability.} \tag{B2}$$

Let A_1 be the event $\{\lambda_n(k_0) < 0, \lambda_n(-k_0) > 0$, and $\lambda'_n(q) \le 0$ for all $|q| \le k_0\}$. Let random variable Z_n be defined by

$$Z_n = \begin{cases} s \in \lambda_n^{-1}(0) \cap (-k_0, k_0) & \text{for which} \\ \quad |s| \text{ is maximum,} & \text{if } A_1 \text{ true} \\ Q_n^0, \quad \text{the sample median,} & \text{if } A_1 \text{ not true.} \end{cases}$$

Define the event A_2 to be $\{|Z_n| < \epsilon\}$, the event A_3 to be $\{|Q_n^0 - Z_n| < k_0 - |Z_n|\}$, and the event A_4 to be $\{|\lambda'_n(q)| < 2D$, for $|q| \le k_0\}$. Then $P(A_2) \to 1$ because $P(A_1)$ and $P(A_2 | A_1)$ both converge to unity as $n \to \infty$. This follows from (B1) and (B2). Also $P(A_3) \to 1$ because $P(A_3) \ge P(\{|Q_n^0| < k_0 - 2|Z_n|\})$ which converges to 1. Finally (B2) implies $P(A_4) \to 1$. Thus $P(\cap_{i=1}^4 A_i) \to 1$.

Now suppose $q \in [-k_0, k_0]$ is not in $\lambda_n^{-1}(0) \cap (-k_0, k_0)$, assuming $A_1 - A_4$ are true. Then for any $z \in \lambda_n^{-1}(0) \cap (-k_0, k_0)$

$$|q + \lambda_n(q)/D - z| < |q - z|,$$

because for $q > z$ we have $\lambda_n(q) < 0$ and $-2(q - z) < \lambda_n(q)/D$ from the mean-value theorem, and similarly for $q < z$. With d_j defined by

$$d_j = \sup_{z_n \in \lambda_n^{-1}(0) \cap (-k_0, k_0)} |Q_n^j - z_n|,$$

we have $d_{j+1} < d_j$ if $d_j > 0$. Thus $\lim_{j \to \infty} d_j$ exists and is nonnegative. Let ρ be the length of the interval set $\lambda_n^{-1}(0) \cap (-k_0, k_0)$. If $\rho = 0$ it follows directly that $Q_n^j \to \lambda_n^{-1}(0)$. If $\rho > 0$ and for j arbitrarily large Q_n^j and Q_n^{j+1} can be on opposite sides of this interval set, then $d_j - d_{j+1} > \rho$; this is a contradiction since $d_j - d_{j+1} \to 0$. Since Q_n^j is thus eventually ($j \to \infty$) on the same side of the set $\lambda_n^{-1}(0) \cap (-k_0, k_0)$, it converges to a point in this set.

This completes the proof of Lemma 1.

REFERENCES

[1] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Stat.*, vol. 35, pp. 73–101, 1964.

[2] ——, "A robust version of the probability ratio test," *Ann. Math. Stat.*, vol. 36, pp. 1753–1758, 1965.

[3] M. C. Yovits and J. L. Jackson, "Linear filter optimization with game theory considerations," *IRE National Conv. Rec.*, Part 4, pp. 193–199, 1955.

[4] W. L. Root, "Communications through unspecified additive noise," *Inform. Contr.*, vol. 4, pp. 15–29, March 1961.

[5] A. A. Ershov, "Stable methods of estimating parameters," *Automat. Remote Contr.*, vol. 39, no. 7, Part 2, pp. 1152–1181, July 1978.

[6] V. M. Krasnenker, "Stable (robust) detection methods for signal against a noise background (survey)," *Automat. Remote Contr.*, vol. 41, no. 5, Part 1, pp. 640–659, May 1980.

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

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

[9] A. H. El-Sawy and V. D. VandeLinde, "Robust detection of known signals," *IEEE Trans. Inform Theory*, vol. IT-23, pp. 722–727, Nov. 1977.

[10] ——, "Robust sequential detection of signals in noise," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 346–353, May 1979.

[11] H. V. Poor, "On detection in weakly dependent noise," *Proc. 1979 Conf. on Inform. Sci. and Syst.*, Johns Hopkins, pp. 499–504, Mar. 1979.

[12] S. L. Portnoy, "Robust estimation in dependent situations," *Ann. Stat.*, vol. 5, pp. 22–43, 1977.

[13] ——, "Further remarks on robust estimation in dependent situations," *Ann. Stat.*, vol. 7, pp. 224–231, 1979.

[14] L. A. Jaeckel, "Robust estimates of location: symmetry and asymmetric contamination," *Ann. Math. Stat.*, vol. 42, pp. 1020–1034, 1971.

[15] J. R. Collins, "Robust estimation of a location parameter in the presence of asymmetry," *Ann. Stat.*, vol. 4, pp. 68–85, 1976.

[16] J. Capon, "On the asymptotic efficiency of locally optimum detectors," *IRE Trans. Inform. Theory*, vol. IT-7, pp. 67–71, April 1961.

[17] D. F. Andrews, *et al.*, *Robust Estimates of Location, Survey and Advances*. Princeton, NJ: Princeton University Press, 1972.

[18] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[19] P. J. Huber, "The behavior of maximum likelihood estimates under nonstandard conditions," *Proc. Fifth Berkeley Symp. On Math. Stat. Prob.*, vol. 1, pp. 221–233, 1967.

[20] R. V. Hogg, "Statistical robustness: one view of its use in applications today," *Amer. Statistician*, vol. 33, pp. 108–115, Aug. 1979.

[21] D. F. Andrews, "A robust method for multiple linear regression," *Technometrics*, vol. 16, pp. 523–531, Nov. 1974.