Threshold Detection in Generalized Non-Additive Signals and Noise

David Middleton
Consultant, LLNL

December 22, 1997

This is an informal report intended primarily for internal or limited external distribution. The opinions and conclusions stated are those of the author and may or may not be those of the Laboratory. Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.
Threshold Detection in Generalized Non-Additive Signals and Noise*

by

David Middleton
127 E. 91 St., New York, NY 10128

December 22, 1997

Table of Contents

Abstract .................................................. p. 1

1. Introduction ........................................... p. 1

2. The Classical Binary On-Off Case: Threshold Optimum Detection for Additive Signals and Noise
   2.1 LOBD's: Independent Samples ................... p. 2

   3.1 Coherent and Incoherent Threshold Detection
   3.2 The Decision Process and Performance ....... p. 6

4. A Radar/Sonar Example: Class A Scatter Models [12], [13]: ............................................ p.11

5. Concluding Remarks .................................. p.13

6. References ............................................ p.14
Threshold Detection in Generalized Non-Additive Signals and Noise*

David Middleton
127 E. 91 St., New York, NY 10128

December 22, 1997

Abstract:

The "classical" theory of optimum (binary-"on-off") threshold detection for additive signals and generalized (i.e. nongaussian) noise is extended to the canonical nonadditive threshold situation. In the important (and usual) applications where the noise is sampled independently, a canonical threshold optimum theory is outlined here, which is found formally to parallel the earlier additive theory, including the critical properties of locally optimum Bayes detection algorithms, which are asymptotically normal and optimum as well.

The important Class A clutter model provides an explicit example of optimal threshold envelope detection, for the non-additive cases of signal and noise. Various extensions are noted in the concluding section, as are selected references.


1. Introduction:

Although the majority of physical applications involving signals in noise are additive, i.e. the desired signal, as well as possible undesired signals, appear additively with the accompanying noise, there are important cases where signal and noise are not additively combined. Examples where the noise is originally gaussian are signal and noise envelopes in narrow band reception [1], and similarly, phase reception in such situations [2]. Similar non-additive combinations of input signals and noise occur in broad-band operations [3], and after nonlinear operations and filtering [4]. Analogous cases arise when the original noise is nongaussian, e.g., Class A or Class B (compound) Poisson processes [5].

It is accordingly of interest to consider the generic situation where signal and noise are combined nonadditively at or within the reception process. Here we shall consider the basic optimum binary threshold detection problem $H_1:S \otimes N$ vs. $H_0:N$ under these circumstances (where $\otimes$ denotes "combination," not necessarily an additive one) of signal with the accompanying noise. Our goal is to obtain the canonical extension of our earlier, "classical" threshold results when signal and noise are additive, and where the noise samples may also be considered statistically independent [6].

Our Report is organized as follows: Section 2 summarizes the earlier optimum threshold binary detection algorithms for $H_1:S+N$ vs. $H_1:N_1$ i.e. the "on-off" cases, which serves as a guide to our
extension of the theory to the nonadditive cases $H_i:S \otimes N$ vs. $H_0:N$ in Section 3 [7]. Section 4 completes our treatment with a brief discussion of the results and possible next steps.

2. The Classical Binary On-Off Case: Threshold Optimum Detection for Additive Signals and Noise:

For independent noise samples, we obtain the earlier canonical results [6] for optimum threshold detection, respectively for coherent and incoherent reception, from the generalized likelihood ratio [cf. Eq. (19.14), [1]]*, now for additive signal and noise, viz.:

$$\Lambda_j(x|\theta) = \mu \left( \frac{F_j(x|\theta)}{F_j(x|0)} \right)_s$$

$$= \mu \left( \frac{W_j(x-s|\theta)}{W_j(x|\theta)} \right)_N,$$

where

$$x = \left[ \frac{X_j}{\sqrt{\psi}} \right]; \mu \equiv p/q; W_j(x|\theta) = J^m - \text{order pdf of } x:H_0:N \text{ alone, and } j = (m,n) = \text{[space (m), time (n)] index}$$

Here $F_j(x|\theta) = F_j(x|s(\theta))$ represents the conditional pdf of $x$, given $s = [S_j/\sqrt{\psi}]$, where $\langle \cdot \rangle_s \rightarrow \langle \cdot \rangle_\theta$ in turn represent the average over the signal-bearing parameters $\theta$, e.g. $\langle \cdot \rangle_\theta = \int \langle \cdot \rangle_\theta w(\theta) d\theta$. As we shall see from the procedures of Sec. 3 ff. for the nonadditive cases, the additive case is a necessary preliminary.

Next, for threshold operation we proceed first to expand the numerator of (2.1) in powers of $\theta$, (through $\theta^4$) to obtain**

$$\Lambda_j = \mu \sum_{n=0} \frac{(-1)^n}{w_j n!} \left[ \left( \theta \cdot \nabla_x \right)^n \right]_\theta w_j(x|^N)_{x=\theta}$$

$$= \mu \left\{ 1 + \left[ -\sum_{i=1} \langle \theta_i \rangle \frac{\partial w_j}{w_j \partial x_i} + \frac{1}{2!} \sum_{ij} \langle \theta_i \theta_j \rangle \frac{1}{w_j} \frac{\partial^2 w_j}{\partial x_i \partial x_j} - \frac{1}{3!} \sum_{ijk} \langle \theta_i \theta_j \theta_k \rangle \frac{1}{w_j} \frac{\partial^3 w_j}{\partial x_i \partial x_j \partial x_k} \right.\right.$$

$$+ \frac{1}{4!} \sum_{ijkl} \langle \theta_i \theta_j \theta_k \theta_l \rangle \frac{1}{w_j} \frac{\partial^4 w_j}{\partial x_i \partial x_j \partial x_k \partial x_l} + O(\theta^5) \left. \right\}$$

where

* As defined by (2.1) and used in [1] and subsequently by the author. The "generalized likelihood ratio" or GLR in some other usage refers to the case where $\theta$ is replaced by the conditional maximum likelihood estimate (CMLE) $\hat{\theta}$ in (2.1) cf. Eqs. (20.173), 20.174, p.932, 910, and discussion [1].

** The results of Section 2 here stem directly from the author's Chapter 4, [7].
\[ y_j = \frac{\partial}{\partial x_j} \log w_j = \frac{1}{w_j} \frac{\partial w_j}{\partial x_j} = w_j^{(i)} / w_j; \quad \frac{\partial^2 w_j}{\partial x_i \partial x_j} = \frac{w_j^{(ij)}}{w_j} = \frac{\partial^2 \log w_j}{\partial x_i \partial x_j} + \frac{w_j^{(i)} w_j^{(j)}}{w_j^2} = z_{ij} + y_{ij}, \]

with \( z_{ij} = \frac{\partial^2 \log w_j}{\partial x_i \partial x_j} \),

and

\[ \frac{1}{w_j} \frac{\partial^2 w_j}{\partial x_i \cdots \partial x_q} = \frac{w_j^{(i-\cdots-q)}}{w_j}, \quad (Q) \equiv (i,j,k,\ldots,q), \text{ etc.} \]

Here all indices are double \( i = (m_1,n_1), j = (m_2,n_2), \text{ etc. for space-time sampling.} \)

Our next step is to expand \( \log \Lambda_n \), using \( \log (1 + u) = u - u^2/2 + u^3/3 - u^4/4 + \cdots, \) where \( u \) equals the expression in the brackets \( [ \] \) in (2.2). Collecting terms of \( O(\theta, \theta^2, \theta^3, \theta^4) \) separately, we obtain, after some algebra, the desired general threshold result for \( \log \Lambda_j \) with additive signals and noise:

\[ \log \Lambda_j = \log \mu - \frac{1}{2} \sum_i \langle \theta_i \rangle y_i + \frac{1}{2^2} \sum_{ij} \left\{ \langle \theta_i \theta_j \rangle (y_i y_j + z_{ij}) - \langle \theta_i \rangle \langle \theta_j \rangle y_i y_j \right\} \]

(2.3a)

\[ - \frac{1}{2!} \sum_{ij} \left\{ \langle \theta_i \theta_j \rangle \frac{w_j^{(ij)}}{w_j} - 3 \langle \theta_i \rangle \langle \theta_j \rangle y_i (y_j y_k + z_{jk}) + 2 y_i y_j y_k \langle \theta_i \rangle \langle \theta_j \rangle \right\} \]

(2.3b)

\[ - \frac{1}{3!} \sum_{ijk} \left\{ \langle \theta_i \theta_j \theta_k \rangle \frac{w_j^{(ijk)}}{w_j} - 3 \langle \theta_i \rangle \langle \theta_j \rangle \langle \theta_k \rangle (z_{ij} + y_i y_j) (z_{jk} + y_j y_k) - 4 \langle \theta_i \rangle y_i \langle \theta_j \rangle \langle \theta_k \rangle w_j^{(ijk)} \right\} \]

(2.3c)

+ 12 \langle \theta_i \rangle \langle \theta_j \rangle y_i \langle \theta_k \rangle (z_{jk} + y_j y_k) - 6 \langle \theta_i \rangle \langle \theta_j \rangle \langle \theta_k \rangle y_i y_j y_k \right\} + O(\theta^5). \]

The reason we carry the expansion of \( \log \Lambda_j \) through \( O(\theta^4) \) is that for incoherent threshold reception terms \( O\left(O(\theta^3, \theta^4)\right)_{\text{coh}} \) are needed to provide the proper bias and ensure asymptotic optimality (AO), while for coherent threshold reception terms \( O(\theta^2) \) are sufficient for this purpose. [The implications of these terminations of the series form of \( \log \Lambda_j \) are discussed further in Sections 4.4 and 4.6 of [7].]

From the results (2.3a)-(2.3c), extended to include spatial sampling, we can write the threshold algorithms \( g^*(x) \) formally in these general cases:

I. Coherent Detection \( (\langle \theta \rangle \neq 0) \):

\[ \log \Lambda_j \bigg|_{\text{coh}} = g^*_j(x)_{\text{coh}} = \hat{B}_{j-\text{coh}} + \log \mu - \langle \hat{\theta} \rangle y; \quad \langle \hat{\theta} \rangle y = \sum_j \langle \theta_j \rangle y_j; \quad \log \mu = \log(p/q), \]

(2.4)

where

\[ y = \begin{bmatrix} y_j \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_j} \log w_j (x_1, \ldots, x_J) \end{bmatrix}; \quad z = \begin{bmatrix} z_{ij} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2}{\partial x_i \partial x_j} \log w_j \end{bmatrix} \]

(2.4a)

are respectively (column) vector and square \((J \times J)\) matrix, with the bias* 

* \( K_\theta \) is dimensionless, and is not normalized so that \( (K_\theta)_{mm} = 1. \)
\[ \hat{B}_{j,\text{coh}} = \frac{1}{2!} \left\langle \left( \mathbf{y} K \mathbf{y} + \langle \tilde{\mathbf{z}} \theta \rangle \right) \right\rangle_{H_0} \; ; \; K_\theta = \left[ \langle \theta, \theta_j \rangle - \langle \theta \rangle \langle \theta_j \rangle \right] = \langle \theta \tilde{\theta} \rangle - \langle \theta \rangle \langle \tilde{\theta} \rangle ; \; \langle \theta \tilde{\theta} \rangle = \rho_\theta. \] (2.4b)

Again, \( w_j(x) = w_j(x_{11}, x_{12}, \ldots, x_{mn}, \ldots, x_{jj}) \) is the joint, \( J \)-th order pdf of the accompanying noise, and \( \mu = p/q \) is the usual ratio of \textit{a priori} probabilities that the data sample contains, or does not contain, a signal.

\[ \text{II. Incoherent Detection, } \langle \theta \rangle = 0, \; \langle \theta, \theta_j, \theta_k \rangle = 0: \]

\[ \log \Lambda_j|_{\text{inc}} = g_j^*(x)_{\text{inc}} = \hat{B}_{j,\text{inc}} + \log \mu + \frac{1}{2!} \left\langle \mathbf{y} K \mathbf{y} + \langle \tilde{\mathbf{z}} \theta \rangle \right\rangle ; \; K_\theta = \rho_\theta = \langle \theta \tilde{\theta} \rangle, \; \langle \theta \rangle = 0; \; \langle \theta, \theta_j, \theta_k \rangle = 0. \] (2.5)

where now the bias term is given by

\[ \hat{B}_{j,\text{inc}} = \frac{1}{4!} \left\langle \sum_{j'k'} \left\{ \langle \theta_j, \theta_{j'}, \theta_k \rangle w_{jj'kk'} \right\} \frac{w_{jj'kk'}}{w_j} - 3 \left\langle \sum_{j'} \langle \theta_j, \theta_{j'} \rangle (y_j y_{j'} + z_{jj'}) \right\rangle \right\rangle_{H_0}, \] (2.5a)

in which \( w_{jj'kk'} = \partial^4 w_j / \partial x_j \partial x_{j'} \partial x_k \partial x_{k'} \) and \( \langle \rangle \) denotes the statistical average under the null hypothesis \( H_0 \). Note that the condition for incoherent threshold detection here requires the vanishing of the signal's joint third moment, which in turn implies certain symmetries and uniformities in the joint pdf's of the signal parameters: \( \langle a_{0j} a_{0j'} a_{0j'} \rangle = 0 \), for instance, or \( \langle s_j s_{j'} s_{j'} \rangle = 0 \). This latter condition is usually (and more readily) satisfied, certainly for narrow-band signals. Finally, we emphasize that these canonical structures apply equally well for broad-band, narrow-band, and purely random signals (where, of course, reception is necessarily incoherent), in generalized noise which can be nonstationary, and nonadditive with the signal. [In the case of purely random signals questions of singular detection, i.e., perfect detection with finite samples, may in principle arise, although practically this is not physically possible.]

Finally, in view the threshold nature of the above expansions and their (Bayesian) optimum, likelihood ratio source, we call the resulting algorithms \textit{locally optimum Bayes detectors} (LOBD's).

\textbf{2.1 LOBD's: Independent Noise Samples}

Except in the case of gauss noise, the general joint pdf \( w_j(x) \) of Eqs. (2.1)-(2.5) is not known, nor is analytically tractable, for most physical situations. Thus, in order at this point to proceed further in a canonical way we must invoke the critical condition that the noise samples are \textit{statistically independent}, e.g., \( w_j(x) = \prod_{j=1}^{J} w_l(x_j) \). This condition can usually be well approximated in practice, at least for temporal sampling, but often involves discarding useful correlation information about the noise field sampled spatially by the receiving array. [This limitation can be overcome to a major extent by the author's approach described below in Sec. 4.7, [7], and [8].]

With independent noise samples we find (cf. Sec. 7, [9], and Appendixes A.1, A.2 of [10]) the following explicit canonical LOBD algorithms, corresponding to (2.4) and (2.5)*:

* Eqs. (2.6a,b), Eqs. (2.7a,b) make use of the results of footnote (*), proceeding Eq. (2.4b) above.
I. Coherent detection:

\[ g_j(x)_{coh} = B_{j-coh}^* - \sum_{j=1}^{J} \langle \theta_j \rangle_j; \quad \langle \theta_j \rangle = \langle a_{0j} s_j \rangle = \langle a_{0n}^{(m)} s_{n}^{(m)} \rangle; \quad \langle \theta_j \rangle \neq 0 ; \quad (2.6a) \]

II. Incoherent detection:

\[ g_j(x)_{inc} = B_{j-inc}^* + \frac{1}{2} \sum_{j,j'} \delta_{jj'} \langle \theta_j \theta_{j'} \rangle; \quad \langle \theta_j \rangle = 0, \quad \delta_{jj'} = \delta_{nn'} \cdot \quad (2.6b) \]

Here it is convenient to anatomize \( \langle \theta_j \theta_{j'} \rangle \) as follows:

\[ \langle \theta_{mn} \theta_{m'n'} \rangle = \frac{\bar{a}_0^2 m_{nn'} \rho_{m'n'}^{(mm')} \delta_{mm'}^{(mm')}}{a_{0n}^{(m)} a_{0n'}^{(m')} \sqrt{a_0^2}}, \quad (2.6c) \]

where \( \bar{a}_0^2 \) is an amplitude covariance with \( \bar{a}_0^2 \) chosen to maximize \( \bar{m} \), so that \( |\bar{m}| \leq 1 \) while \( \bar{\rho} \) is a waveform covariance \( |\bar{\rho}| \leq 1 \), cf. (2.2) et seq.

The biases \( (B_j^*) \) are specifically found to be, after considerable manipulation (cf. Appendix A.1, A.2 of [10] for details):

\[ B_{j-coh}^* = \log \mu - \frac{1}{2} L^{(2)} \sum_j \langle \theta_j \rangle^2 = \log \mu + B_{j-coh}^* \quad (2.7a) \]

\[ B_{j-inc}^* = \log \mu - \frac{1}{8} \sum_{j,j'} \langle \theta_j \theta_{j'} \rangle^2 \left( \left[ L^{(4)} - 2 L^{(2)^2} \right] \delta_{jj'} + 2 L^{(2)^2} \right) = \log \mu + B_{j-inc}^* , \quad (2.7b) \]

in all of which

\[ \begin{align*}
\gamma & = \frac{d}{dx} \log w_1(x) \bigg|_{x=x_j}; \quad \bar{\gamma} = \frac{d}{dx} \bar{\varphi} \bigg|_{x=x_j}; \quad L^{(2)} = \langle \bar{\varphi} \rangle_{\theta_0}; \quad L^{(4)} = \int \left( \frac{w''}{w_1^2} \right) w_1 dx = \langle (I + \bar{\varphi})^2 \rangle_{\theta_0}, \quad (2.7c) \end{align*} \]

where \( L^{(2)}, L^{(4)} \geq 1 \).

Direct evaluation of the variances of the \( g_j^* \) under \( H_0, H_1 \) shows that the detection parameter, \( (\sigma_{0j}^2) \), defined by

\[ \begin{align*}
\text{var}_0 g_j^* & = (\sigma_{0j}^2)^2 = \frac{-2 B_j^*}{\text{var}_1 g_j^*}. \quad (2.8a) \end{align*} \]
obeys (2.8a), where the optimum bias \( \hat{B}_j + \log \mu \) is determined according to the procedures described above indicated generally by (2.4b), (2.5a). The associated threshold signal condition is established from the requirement that

\[
\var_g \sigma_j^2 = \left( \sigma_{0j}^2 \right)^2, \text{ e.g., } \sigma_j^2 = \left( \sigma_{0j}^2 + F_j(\theta) \right)^2, \text{ or, } \left| F_j(\theta) \right| < \left( \sigma_{0j}^2 \right)^2, \quad (2.9)
\]

cf. [10], Eq. (2.2a). This condition sets an upper limit on \( \left\langle \theta^2 \right\rangle = \left\langle a_0^2 \right\rangle \ll 1 \), e.g., for a least upper bound on the "weak" input signals for which (2.6a,b) is essentially valid. In practice, of course, we always require \( \left\langle a_0^2 \right\rangle > 0 \) under \( H_0 \).

In the case where the accompanying noise is solely gaussian, then

\[
w_j(x) = e^{-x^2/2} / \sqrt{2\pi}, \quad (2.10)
\]

and \( \left\langle x_j \right\rangle = x_j, \theta = -1 \), with \( L^{(2)} = 1, L^{(4)} = 2 \). Usually, however, the noise is nongaussian: for clutter we may expect periods of Class A-like noise, for example, so that \( \chi(x) \) is no longer linear. [Examples of \( \chi(x) \) are shown in Figs. 7.1-7.2b of [10] for the important canonical classes on nongaussian noise described by Class A and B noise models.] Clearly, the statistical character of the generalized noise, e.g., Class A, B, etc. embodied here in the pdf \( w_j(x) \), can drastically modify the \( (ZMNL) \) transfer response \( y = \chi(x) \) vis-à-vis the linear characteristic \( y = -x \) for the normal noise often encountered in conventional applications.

### 3. Generalization: The Classical Binary On-Off Case: Threshold Optimum Detection for Nonadditive Signal and Noise:

In Section 2 above we have outlined the derivation of the binary "on-off" LOBD's for the usual cases where the input signal and noise are additive. As noted in Section 1 above, there may be situations where signal and noise are no longer additive. Accordingly, let us proceed to derive the LOBD detector algorithms for this latter case*. For this purpose we follow the approach used (2.1)-(2.3), to obtain the formal analogues of (2.4) and (2.5). This is accomplished by again expanding the likelihood ratio (2.1) about the null signal \( \theta = 0 \), now represented by the Taylor's series.

\[
\Lambda_j(x|\theta) = \mu \sum_{n=0}^{\infty} \frac{1}{F_j(x|0)} \cdot \frac{1}{n!} \left[ \left( \theta \cdot \nabla \theta \right)^n F_j(x|\theta) \right]_{\theta=0}^{\theta=\theta_j}, \quad J = MN, \quad (3.1)
\]

where, of course, \( F_j(x|0) = w_j(x)_N \) in (2.1), and now (3.1) is to be compared with (2.2). [The indexes \((i, j)\) etc. in what follows are "double indexes": \( i = (m_1, n_1), j = (m_2, n_2), \) etc., representing space and time samples as before, cf. (2.1b).] For the additive cases we observe directly that \( -\theta \cdot \nabla \theta = \theta \cdot \nabla \theta \). For the non-additive cases the various derivatives in (3.1) are represented by

\[
F_j^{(ijk \ldots)}(x|\theta) = \frac{\partial^{(ijk \ldots)}}{\partial \theta_i^j \partial \theta_2 \partial \theta_3 \ldots} F_j(x|\theta) \equiv \left[ g_j^{(ijk \ldots)}(x|\theta) \cdot F_j(x|0) \right], \quad (3.1a)
\]

this last defining the \( g_j^{(ijk \ldots)}(x|\theta) \). Here the \( F_j^{(ijk \ldots)} \) replace the \( w_j^{(ijk \ldots)}(x)_N \) of the additive régimes (2.2a), so that now

* The results of Section 3 here follow from Section 4.3.4 of [7]. See also, Maras, [11].
\[ y_i \rightarrow \hat{y}_i = g_f(x|0) = \left( \frac{\partial}{\partial \theta'_i} \log F_j \right)_{\theta'_i = 0} = \left( \frac{1}{F_j} F_j^{(i)} \right)_{\theta'_i = 0}; \]

\[ y_i y_j + z_{ij} \rightarrow \hat{y}_i \hat{y}_j + \hat{z}_j = g_f^{(ij)}(x|0); \quad \frac{w_j^{(ijk\ldots)}}{w_j} \rightarrow g_j^{(ijk\ldots)}(x|0), \]

with

\[ \hat{z}_{ij} \equiv \left( \frac{\partial^2}{\partial \theta'_i \partial \theta'_j} \log F_j \right)_{\theta'_i = 0} \left( \frac{1}{F_j} \frac{\partial^2}{\partial \theta'_i \partial \theta'_j} F_j \right)_{\theta'_i = 0} = \left( \frac{F_j^{(ij)}}{F_j} \right)_{\theta'_i = 0} = \hat{z}_j + \hat{\gamma}_j \hat{y}_j = g_j^{(ij)}(x|0), \]

in the general development of log \( \Lambda_j \) in (2.3a,b,c) above in powers \( O(\theta^p), p = 1,\ldots,4 \). Again, it is remarked that this degree (\( p \leq 4 \)) of expansion ultimately may be needed for suitable termination of the LOBD series.

To proceed further, let us again postulate **independent noise samples**, so that

\[ F_j(x|\theta) = \prod_{j=1}^J F_1(x_j|\theta_j); \quad F_j(x|0) = \prod_{j=1}^J F_1(x_j|0) = \prod_{j=1}^J w_1(x_j), \]

and thus we have

\[ y_j \rightarrow \hat{y}_j = \left[ \frac{\partial}{\partial \theta'_j} \log F_{1j}(x_j|\theta'_j) \right]_{\theta'_j = 0} = \frac{1}{F_1(x_j|0)} \left[ \frac{\partial F_1(x_j|\theta'_j)}{\partial \theta'_j} \right]_{\theta'_j = 0} = \frac{F'_{1j, \theta'_j = 0}}{F_{1j, 0}} \equiv \hat{\gamma}_j(x_j), \]

where \( (') \) here refers to the (partial) derivative with respect to \( \theta'_j \), at \( \theta'_j = 0 \), in place of \( (d/dx) \) at \( x = x_j \), cf. (2.7c).

Similarly, we have

\[ \hat{\gamma}_j \equiv \left[ \frac{\partial^2}{\partial \theta'_j \partial \theta'_j} \log(F_{1j} F_{1j}) \right]_{\theta'_j = 0} = \hat{\gamma}_j \delta_{ij}, \quad \text{or} \]

\[ \hat{\gamma}(x_j) = \left[ \frac{\partial}{\partial \theta'_j} \left( \frac{F_{1j}}{F_{1j}} \right) \right]_{\theta'_j = 0} = \left( \frac{F_{1j}}{F_{1j}} \right)_{\theta'_j = 0} \hat{\gamma}_j; \]

\[ g_{f}^{(ij)} \big|_{\theta'_j = 0} = \left[ F_{1j} (F_{1j})_{\theta'_j = 0} \right]_{\theta'_j = 0} = y'_j \left[ \hat{\gamma}_j \right] + \hat{z}_j = \hat{\gamma}_j + \hat{\gamma}_j \hat{\gamma}_j; \]

and

\[ g_{f}^{(ijk\ldots)} \big|_{\theta'_j = 0} = \left[ F_{1j} (F_{1j})_{\theta'_j = 0} \cdots \left( F_{1j} F_{1k} \cdots \right)_{\theta'_j = 0} \right]_{\theta'_j = 0} = \frac{\partial F_{1j}}{\partial \theta'_i}, \quad \text{etc.}, \]

which allows us, using (2.3) above, to write explicitly the extension of (2.6), (2.7) to these nonadditive cases.

This, in turn, is accomplished with the help of the following **regularity conditions**:
The threshold optimal results (for independent noise examples) analogous to the additive cases (2.6a), (2.7a) above, now given in Sec. 3.1 below.

3.1 Coherent and Incoherent Threshold Detection:

We begin with the case of coherent detection (where \( \langle \theta \rangle \neq 0 \)), when signal and noise are nonadditive. Comparison with the additive cases of Section 2 above shows that, in form, we have

I. Coherent Detection:

\[
\log \Lambda_j \bigg|_{coh} = \hat{\Theta}^{*}_{j-coh}(x) = B^{*}_{j-coh} + \langle \hat{\Theta} \rangle \hat{1},
\]

in matrix form, where \( \hat{\Theta} = [\hat{\theta}] \) and \( \hat{1} = [\hat{1}] \). The bias term here is specifically, with the help of the regularity conditions (3.5),

\[
B^{*}_{j-coh} = \log \mu + \frac{1}{2n} \sum_{i=1}^{n} \left\{ \langle \theta_i \theta_i \rangle - \langle \theta_i \rangle \langle \theta_i \rangle \langle \hat{\Theta} (x_i) \hat{\Theta} (x_i) \rangle_{H_0} + \langle \theta_i \theta_i \rangle \langle \hat{\Theta} (x_i) \hat{\Theta} (x_i) \rangle_{H_0} \delta_y \right\},
\]

\[
\therefore B^{*}_{j-coh} = \log \mu - \frac{1}{2} \langle \Theta^{(2)} \rangle \Theta = \log \mu - \frac{1}{2} \sum_{i} \hat{L}^{(2)*}_{i} \langle \theta_i \rangle^2 \hat{L}^{(2)*}_{i} = \left[ \hat{L}^{(2)*}_{i} \delta_y \right], \hat{L}^{(2)*} = \langle \hat{\Theta}^2 \rangle_{H_0}.
\]

(Note that \( \hat{L}^{(2)} = \langle \hat{\Theta}^2 \rangle_{H_0} \)) is a (diagonal) Fisher information matrix, cf. Eq. (22.18) of [1].)

In a similar way we obtain the nonadditive counterpart of (2.6b) and (2.7b), again using the regularity conditions (3.5a,b), viz.:

II. Incoherent Detection, \( \langle \theta_i \rangle = 0; \langle \theta_i \theta_j \rangle = 0 \):

\[
\log \Lambda_j \bigg|_{inc} = \hat{\Theta}^{*}_{j-inc}(x) = B^{*}_{j-inc} + \frac{1}{2n} \langle \Theta \hat{L} + \hat{L}^{'*} \Theta \rangle \Theta \delta_y = B^{*}_{j-inc} + \frac{1}{2n} \sum_{y} \langle \Theta \theta_y \rangle \langle \hat{\Theta} (x_i) \hat{\Theta} (x_j) + \hat{\Theta} (x_i) \delta_y \rangle \}
\]

where the square matrices \( \hat{L}, \hat{L}' \) are specifically

\[
\hat{L} = \left[ \hat{\Theta} \hat{\Theta} \right]; \hat{L}' = \left[ \hat{\Theta} \delta_y \right],
\]

(3.8a)
with now

\[
\bar{b}_{j-\text{inc}}^* = \log \mu - \frac{1}{8} \sum_j \left\langle \theta_i \theta_j \right\rangle^2 \left( \left[ \hat{\ell}_j^{(4)*} - 2 \hat{\ell}_j^{(2)*} \right] \delta_{ij} + 2 \hat{\ell}_i^{(2)*} \hat{\ell}_j^{(2)*} \right),
\]

(3.9)

with

\[
\hat{\ell}_j^{(4)*} = \left[ \hat{\ell}_j^{(4)*} \delta_{ij} \right]; \quad \hat{\ell}_i^{(4)*} = \int_{-\infty}^\infty \left( \frac{F_{ij}^{(\mu)}}{F_{ii}^{(\mu)}} \right)^2 F_{ii}(x_i|0)dx_i = \left\langle \left[ \frac{\delta}{\sqrt{2\pi}} \right]^2 \right\rangle_{H_0}.
\]

(3.9a)

Accordingly, such quantities as \( L^{(2)}, L^{(4)}, (2.7c) \), appearing in the bias terms (2.7a,b) of the additive cases (with independent noise samples) now become \( \hat{L}^{(2)}, \hat{L}^{(4)}, \) cf. (3.7), (3.9a) here. These results stem from the formal structure of (2.3), now extended to these nonadditive cases and paralleling the analysis of Appendices A.1, A.2 of [10].

### 3.2 The Decision Process and Performance:

Our results (3.6) - (3.9) depend only on the continuity and differentiability of the \( F_{ii}^{(\mu)} \). Not only do the regularity conditions (3.5a,b) simplify the structure of the threshold algorithms, they also ensure that the latter are locally asymptotically normal (LAN) and asymptotically optimum (AO) as well. Moreover, direct evaluation shows that the detection parameter \( \hat{\sigma}_{ij}^2 \) (\( \text{var}_{H_0} \hat{g}_j^* \)) becomes here, like (2.8) in the additive cases,

\[
\hat{\sigma}_{ij}^2|_{\text{coh,inc}} = -2 \left( \bar{b}^* - \log \mu \right)|_{\text{coh,inc}}, \quad \text{cf. (2.8), (2.8a)}.
\]

(3.10)

The pdf's of \( \hat{g}_j^* \) are also asymptotically normal (AN), again by the Central Limit Theorem (CLT), cf. Sec. 2.7 - 3 of [1], so that we can write at once

\[
w_i(\hat{g}_j^*|H_0/H_1) = G_j\left( \log \mu \mp \hat{\sigma}_{ij}^2/2, \hat{\sigma}_{ij}^2 \right); \quad G_j(a,b) = \frac{e^{-(y-a)^2/2b^2}}{\sqrt{2\pi b^2}},
\]

(3.11)

(cf. last of Sec. 5.1, [7]), and Le Cam's concept of contiguity (cf. Sec. 5.1, [7]) may be used here also. Equations (3.11) allow us to determine the associated (optimum) detection probability forms

\[
p_d^* = 1 - \beta^* = \int \frac{w_i(y|H_1)dy}{\log K}; \quad \alpha_F^* = \int \frac{w_i(y|H_0)dy}{\log K}, \quad y = \hat{g}_j^*, \text{ with (3.11)}.
\]

(3.12)

where \( p_d^* \) is the (conditional) probability \( (p_d^*/p) \) of correctly detecting the signal when it is present with the noise, while \( \alpha_F^* \) is the (conditional) false alarm probability of incorrectly deciding a signal is present when it is not. (For a detailed treatment, see [6] and [9].)
I. The Decision Process $H_I: S \oplus N$ vs. $H_0: N$

For the typical binary "on-off" decision cases the decision process is

\[
\begin{align*}
\text{Decide } H_I: S \oplus N: & \quad S_{\text{target}}(S_m) \oplus N_c(S_m) + N_A, \text{ if: } g^*_{j} \geq \log K \\
\text{vs.} & \\
\text{Decide } H_0: N: & \quad N_c(S_m) + N_A, \text{ if: } g^*_{j} < \log K,
\end{align*}
\]

where $K (> 0)$ is a threshold setting, determined by the receiver's choice of false alarm probability, $\alpha^*_f$, viz.,

\[
\log K = \log \mu + \sigma^*_{o, i} \sqrt{2} \Theta^{-1}(1 - 2\alpha^*_f) - \sigma^*_{o, i} / 2;
\]

\[
\Theta(y) = (2/\sqrt{\pi}) \int_{0}^{y} e^{-t^2} dt = \text{erf } y.
\]

II. Performance:

Because these LOB detectors are locally asymptotically normal (LAN), we find for the detection parameter $\sigma^*_{o, i}$ that specifically

\[
\sigma^*_{o, i} = \frac{\left(g^*_{j} \right)_{H_1} - \left(g^*_{j} \right)_{H_0}}{\sqrt{\text{var}_{H_0} g^*_{j}}} = -2 \text{[Bias} - \log \mu]\}
= 2 \left(\alpha^*_f\right)_{\text{min-det.}} - \Pi^*_{\text{inc}} \left(\alpha^*_f\right)_{\text{min-det.}} - \Pi^*_{\text{coh}}
\]

\[
\sigma^*_{o, i} = 2 \left(\alpha^*_f\right)_{\text{min-det.}} \quad \text{(3.15)}
\]

cf. (3.10), in which the bias (cf. (2.7a, 2.7b)) and the processing gain $\Pi^*$ are functions of the statistics $(L^{(2)*}, L^{(4)*}, \hat{L}^{(4)*}, \hat{L}^{(4)*})$, cf. (2.7c, 3.7), (3.9a) above, (determined here from (2.6a), (2.6b) applied to (3.15) in these explicit cases). Performance itself for the LAN cases is expressed by the probability of correct detection

\[
P^*_D = pp^*_D = \frac{P}{2} \left[1 + \Theta \left(\frac{\sigma^*_{o, i}}{\sqrt{2}} - \Theta^{-1}(1 - 2\alpha^*_f)\right)\right],
\]

\[
\text{where now}
\]

\[
\alpha^*_f = \frac{1}{2} \left(1 - \Theta \left(\frac{\sigma^*_{o, i}}{\sqrt{2}} + \frac{\log(K/\mu)}{\sigma^*_{o, i}}\right)\right).
\]

From (2.7b), (3.9a) we see that the fundamental statistic now for these cases of (incoherent) envelope (i.e., nonadditive S, N) detection is the first-order pdf $w_i(\hat{L}^{(4)*})$, from which in turn $\hat{L}^{(4)*}$, (3.9a), is required for the evaluation of performance, with $\hat{L}^{*}$, cf. (3.3a), needed for the LOB processing algorithm. It is the evaluation of $\hat{L}^{(4)*}$, needed in $\sigma^*_{o, i}$, from (3.10) and the bias terms, upon which we next focus in Sec. 4 ff.
4. A Radar / Sonar Example: Class A Scatter Models [12], [13]:

It has been recently shown that the Class A noise model is the appropriate model (on physical
grounds), cf. Sec. 5 of [12], [13], for incoherent detection in clutter (radar) and reverberation (sonar).
As an example, let us consider the cases where large scale fluctuations in intensity (perceived at the
receiver) are ignorable, corresponding to small large scale wave surface activity, for radar (sonar) off
the sea surface. Thus, we have explicitly

$$F_{1i,\theta} = e^{-\lambda_{m}} \sum_{m=0}^{\infty} \frac{A_{(m)}^{(m)}}{m!} \frac{E_{m}}{\sigma_{m}^{2}} e^{-(E_{m}^{2} + \theta_{i})/2\sigma_{m}^{2}} I_{0}\left(\frac{\theta_{i} E_{m}}{\sigma_{m}^{2}}\right),$$

(4.1)

which is a Ricean mixture process. The "overlap factor" $A_{(m)}^{(m)}$ represents the average number of
"large" scattering events or sources, at any given time, and includes multiple scatter contributions,
cf. Sec. 5A of [12], [13].

Here specifically,

$$2\hat{\sigma}_{m}^{2} = \frac{(m + \Gamma_{A})}{\Gamma_{A}} \left[1 + \Gamma_{A}\right]; \quad \theta_{i} = a_{0}, \gamma_{i};$$

$$\psi = \Omega_{2A}^{(m)}(1 + \Gamma_{A}); \quad \Gamma_{A} = \sigma_{G}^{2}/\Omega_{2A}^{(m)}$$

$$\sigma_{G}^{2} = \text{intensity of gaussian component}; \quad \Omega_{2A}^{(m)} = \text{intensity of nongauss (scatter) component}$$

(4.2a)

Here specifically,

$$w_{i}(E)_{\theta_{i}=0} = \frac{E}{\hat{\sigma}_{m}^{2}} e^{-E^{2}/2\hat{\sigma}_{m}^{2}};$$

$$E = E/\sqrt{2\psi}; \quad E^{2} = 1, \left(E^{2} = 2; 2\hat{\sigma}_{m}^{2} = 1\right).$$

(4.2b)

It can readily be shown that

$$(3.3a): \hat{\gamma} = 0, \text{ and } \hat{\gamma}_{(2)} = 0, \text{ cf. (3.7)}$$

(4.3)

for $F_{1i,\theta}, \ (4.1)$. This is not unexpected, since for envelope detection [and a uniform phase ($\phi$) for
$w_{i}(E, \phi)_{\theta_{i}=0}$], detection is necessarily incoherent, provided, of course, that the nonadditive signal and
noise is such that as $\theta_{i} \to 0, \hat{\gamma} = 0$, as in (4.3), for (4.1) here.

The statistic $\hat{L}_{A}^{(4)}$, (3.9a), however, is positive and nonvanishing, e.g. $\hat{L}_{A}^{(4)} = \langle\hat{\gamma}^{2}\rangle_{H_{0}}$ from (3.9a),
since $\hat{\gamma} = 0$ here, cf. (4.3). Applying (4.1) to (3.9a) gives finally

$$\hat{L}_{A}^{(4)*} = \int_{0}^{\infty} \left[ e^{-\lambda_{m}} \sum_{m=0}^{\infty} \frac{A_{(m)}^{(m)}}{m!} \left(\frac{E_{m}^{2}}{2\hat{\sigma}_{m}^{2}} - \frac{1}{\hat{\sigma}_{m}^{2}}\right) w_{l m}(E)_{\theta_{i}=0} \right]^{2} dE ,$$

(4.4)
where now

\[ w_{1m}(E)_{\theta,i=0} = \frac{E}{\sigma_m^2} e^{-E^2/2\sigma_m^2}, \text{ cf. (4.2)} \]  \hspace{2cm} (4.4a)

[In the special nonadditive case of a purely Rician pdf, e.g. \(2\sigma_m^2 = 1\) in (4.1) and (4.4), we see that (4.4) reduces to]

\[ \tilde{\Lambda}_{\text{Rica}}^{(4)*} = \int_0^\infty (2E^2 - 2)^2 w_1(E) dE = 4E^4 - 8E^2 + 4 = 4 , \]  \hspace{2cm} (4.5)

from (4.2) above, which checks with earlier work, cf. Table 9.1, [1]. Figure 4.1 shows \(\tilde{\Lambda}_{\text{Rica}}^{(4)*}\) as a function of the gaussian factor \(\Gamma_A\). Note that as this factor becomes large (or \(A_B\) becomes large) the gaussian component of the received noise dominates and is here Rayleigh distributed, e.g., \(\tilde{\Lambda}_{\text{Rica}}^{(4)*} = 4 (= 6 \text{ dB})\) from (4.5), as expected.

---

**Fig. 4.1.** \(\tilde{\Lambda}_{\text{Rica}}^{(4)*}\), Eq. (4.4), as a function of the gaussian factor \(\Gamma_A\), (4.2), for various values of \(A_B^{(m)}\).

When there are two scales to the surface scatter process, viz., a small-scale "speckle process" which rapidly decorrelates in (space and time) and an underlying "slower" modulation (large-scale waves), the intensity \(\psi\) is modulated. The statistics of \(\psi\) are usually described by a \(\Gamma\)-distribution. This, in turn, converts (4.1) into a KA-pdf (Sec. 6 of [12], [13]), which in most instances provides an effective statistical description of the clutter, cf. Fig. (6.2), or (6.1), of [12], [13]. (For the counterpart of \(\tilde{\Lambda}_{\text{Rica}}^{(4)*}\), (4.4), we must then use the appropriate KA pdf. – to be evaluated subsequently.)

---

5. Concluding Remarks:
In the preceding Sections we have outlined the development of optimum (Bayesian) threshold detection algorithms for independent noise samples when signal and noise are both additive (Section 2) and nonadditive (Section 3), and when binary "on-off" decisions are required. Thus, the hypothesis tests here are respectively denoted by: $H_1: S + N$ vs. $H_0: N$ for the additive cases, and by $H_1: S \otimes N$ vs. $H_0: N$ for the general, nonadditive situation, cf. (3.13). It turns out that for the latter the formal results are the same as for the former, provided certain regularity conditions (3.5a, 3.5b) are obeyed, as they usually are. The resulting threshold algorithms are both LOBD (locally optimum Bayes detectors) and asymptotically optimum (AO) as well as locally asymptotically normal (LAN). The pdf's (under $H_0, H_1$) of these algorithms are normal, cf. (3.11), so that the associated detection probabilities $\left( \alpha^*, P_d^* \right)$ are readily obtained, cf. (3.12), and explicitly from (3.16), (3.17).

As a specific example, applicable to radar and sonar scatter, environment, on physical grounds, we introduce the Class A envelope distribution with signal, or Rician mixture model (4.1). It is shown for this model that $\hat{\alpha}_4^{(2)*} = \hat{\alpha}_4^{*} = 0$; $\hat{\alpha}_4^{(3)*}$ is also obtained, cf. (4.4) and Fig. 4.1, the latter representing numerical integration of (4.4)* these include the well-known K-distribution and the author's recent extension of it namely the KA cases [12], [13].

A variety of extensions of these results appears possible: (1), binary threshold detection involving two classes of signals, e.g., $H_1: S_1 \otimes N$ vs. $H_2: S_2 \otimes N$; (2), multiple alternative detection (many different signal classes); and (3), appropriate extensions to threshold estimation [11], as well as to specific classes of signals in noise. We emphasize again here the general character of the analysis, which applies to nongaussian noise and arbitrary combinations of signal with noise (subject to (3.5a,b)).

* The author is indebted to Dr. A. H. Nuttall (NUWC), Code 311 for the calculations.
6. References


