01 Jan 1975

Generalized Likelihood Signal Resolution

John A. Stuller
Missouri University of Science and Technology, stuller@mst.edu

Recommended Citation

J. A. Stuller, "Generalized Likelihood Signal Resolution," *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 276 - 282, Institute of Electrical and Electronics Engineers, Jan 1975.
The definitive version is available at https://doi.org/10.1109/TIT.1975.1055371

Generalized Likelihood Signal Resolution

JOHN A. STULLER

Abstract—**This paper defines an M-ary generalized likelihood ratio test (MGLRT) that overcomes Root's early objection to the application of generalized likelihood ratio testing to the resolution of correlated signals. The proposed test extends the form of a conventional binary generalized likelihood ratio test (GLRT) in a manner that permits a generalization of the minimax properties of the binary test to the M-hypotheses case. When the estimated signals are orthogonal, the test reduces to a sequence of conventional binary tests duplicating the performance of a narrow-band matched filter envelope-detector receiver.**

I. Introduction

A WAVEFORM $r(t)$, $0 \leq t \leq T$, is observed that contains an unknown number of signals in independent white complex Gaussian (WCG) noise $w(t)$ having spectral height N_0. The individual signals have the form $a_n s(t,\boldsymbol{\alpha}_n)$, where a_n is an unknown complex amplitude, $\boldsymbol{\alpha}_n$ is a real vector of either known or unknown parameters, and the function $s(t,(\cdot))$ is known and identical for all signals. The receiver must determine the number of signals present and estimate the signal amplitudes and parameters. This situation imposes no particular problem if the estimated signals are orthogonal; multiple signal detection then simplifies to a series of single signal-detection problems for which a conventional generalized likelihood ratio test (GLRT) works. Correlation among the signals upsets this simplicity and gives rise to the issue of signal resolvability, which is the topic of this paper.

Various modifications of this problem have been discussed in the literature. In an early paper [1], Helstrom obtained an expression for the minimum probability of error in deciding which of two equally probable equal-energy radar signals of known parameters is present in additive white Gaussian noise. Nilsson [2] derived the form of Bayes' processor for simultaneous estimation and detection of radar returns from a collection of stationary point scatterers whose number, cross sections, and ranges are random variables. Root [3] obtained conditions that indicate when one could not expect to resolve two signals of unknown amplitudes and parameters and pointed out a difficulty in using a GLRT for that purpose. Ksienski and McGhee [4] discussed the problem of radar angular resolution in the Bayes context using uniform cost functions for estimation and decision errors. This led them to a maximum likelihood test that contained the difficulties discussed by Root. In view of this, threshold tests based on the signal amplitude estimates were proposed, and their utility was demonstrated experimentally. Helstrom [5] gave the marginal detection probabilities of the individual signals for this type of decision with known parameters. Kemerait and Childers have contributed an interesting paper [6] in which cepstrum techniques are employed for both signal detection and extraction. Computationally efficient suboptimum techniques are described by Lichtenstein and Young [7], [8]. The difficult problem of describing parameter estimation errors is considered by Selin [9].

The present paper shows that the signal resolution problem can in fact be formulated in terms of an M-ary generalized likelihood ratio test (MGLRT) that represents a natural extension of the binary GLRT to the M-hypotheses case. Subject to certain constraints, this test is minimax with respect to variations in *a priori* signal amplitude statistics.

II. Preliminaries

For the purpose of deciding the number of signals present in $r(t)$, the receiver formulates M hypotheses

$$H_i\colon r(t) = \Omega_i(t) + w(t), \qquad i = 0, 1, \cdots, M-1 \quad \text{(1a)}$$

in which $r(t)$ contains exactly k_i (typically i) individual signals under H_i:

$$\Omega_i(t) = \begin{cases} 0, & \text{for } i = 0 \\ \sum_{n=1}^{k_i} a_n s(t,\boldsymbol{\alpha}_n), & \text{for } i = 1, 2, \cdots, M-1 \end{cases} \quad \text{(1b)}$$

where $0 < k_1 < k_2 < \cdots < k_{M-1}$. The parameter k_i-tuple $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{k_i}\} \equiv \{\boldsymbol{\alpha}\}_i$ associated with hypothesis H_i is an element of an arbitrary space $\mathscr{B}(H_i) = \mathscr{B}_i$ of possible real k_i-tuples. If each $\mathscr{B}_i$ contains a single element, we have the known parameter problem. The amplitude k_i-vector $(a_1 a_2 \cdots a_{k_i})^T \equiv \boldsymbol{a}_i$ associated with hypothesis H_i is an unknown point in complex k_i space $\mathscr{C}^{k_i}$. It is implicit in (1) that the values of the signal amplitudes a_n and parameter vectors $\boldsymbol{\alpha}_n$ may differ with each hypothesis. The total number of hypotheses M is given *a priori* or is forced by computational constraints.

The quantities $r(t)$, $a_n s(t,\boldsymbol{\alpha}_n)$, $\Omega_i(t)$, and $w(t)$ in (1) are the complex amplitudes [10, p. 59] of their real narrow-band counterparts. Throughout this paper it will be assumed that the function $s(t,(\cdot))$ is known, is the same for all signals, and is normalized

$$\int_0^T |s(t,\boldsymbol{\alpha}_n)|^2 \, dt = 1 \quad (2)$$

for every value assumable by $\boldsymbol{\alpha}_n$. It will also be assumed that the k_i signals present under H_i are algebraically independent. This condition is required if one is to have any hope of resolving individual signals.

Manuscript received July 25, 1973; revised November 12, 1974. This work was done in collaboration with staff members of the Radio Research Directorate, Communications Research Centre, Department of Communications, Ottawa, Ont., Canada, and was supported in part by National Research Council Grant A8251.
The author is with the Department of Electrical Engineering, University of New Brunswick, Fredericton, N.B., Canada.

In the sequel the waveforms $r(t)$, $s(t,\boldsymbol{\alpha}_n)$, $\Omega_i(t)$, and $w(t)$ will be represented by N-vectors $\boldsymbol{r}$, $\boldsymbol{s}(\boldsymbol{\alpha}_n) \equiv \boldsymbol{s}_n$, $\boldsymbol{\Omega}_i$, and $\boldsymbol{w}$ that are obtained by expanding their associated waveforms in N-term series of arbitrary complex orthonormal functions [11, p. 244]. It is assumed that the orthonormal function set used for this representation is complete on $[0,T]$ in the limit $N \to \infty$. In subsequent work we tacitly assume the limit $N \to \infty$ whenever the result of this operation is well defined [12, p. 274]. We denote the matrix of signal coordinate vectors under H_i as

$$S_i = [\boldsymbol{s}_1 \boldsymbol{s}_2 \cdots \boldsymbol{s}_{k_i}] \tag{3}$$

and use the symbol $\mathscr{S}_i$ to denote the space spanned by the column vectors of S_i. The asterisk (*) is used to denote matrix and vector conjugate transposition. By these representations, (1) and (2) become

$$H_i: \boldsymbol{r} = \boldsymbol{\Omega}_i + \boldsymbol{w}, \qquad i = 0, 1, \cdots, M-1 \tag{4a}$$

where

$$\boldsymbol{\Omega}_i = \begin{cases} 0, & i = 0 \\ S_i \boldsymbol{a}_i, & i = 1, 2, \cdots, M-1 \end{cases} \tag{4b}$$

and

$$|\boldsymbol{s}_n|^2 = 1. \tag{5}$$

Since $w(t)$ is a WCG process with covariance function $N_0\,\delta(\tau)$, $\boldsymbol{w}$ is a WCG vector with covariance matrix $N_0 I$.

III. M-ary Generalized Likelihood Ratio Test

In terms of the vector representation (4), the problem of formulating an M-ary hypotheses-testing rule is to optimally partition the space of all $\boldsymbol{r} \sim r(t)$ into M regions Z_n such that a decision $D_i \sim H_i$ occurs if and only if $\boldsymbol{r}$ falls in region Z_i. In the absence of prior probabilities on $\boldsymbol{a}$, $\boldsymbol{\alpha}$, and H, however, no decision-theoretic solution is available even for the *binary* case with $k_1 = 1$ in which signal resolvability is not an issue. For the binary case, however, a GLRT gives useful results. Let $p(\boldsymbol{r} \mid (\cdot))$ denote the conditional probability density function of $\boldsymbol{r}$. Then the GLRT is

$$\max_{a_1, \boldsymbol{\alpha}_1} \ln \frac{p(\boldsymbol{r} \mid a_1, \boldsymbol{\alpha}_1, H_1)}{p(\boldsymbol{r} \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{10} \tag{6a}$$

where the symbol $\gtrless_{H_0}^{H_1}$ denotes $D = H_1$ or $D = H_0$ depending upon the inequality satisfied. For data of the form (4), (6a) simplifies to

$$l_1 = \max_{\boldsymbol{\alpha}_1} |\boldsymbol{s}^*(\boldsymbol{\alpha}_1)\boldsymbol{r}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} t_{10} = N_0 \gamma_{10} \tag{6b}$$

which represents a conventional narrow-band matched filter square-law envelope-detector receiver. It is well known [10, sec. 9.8] that the performance of this receiver is optimum in the Neyman–Pearson sense for known $\boldsymbol{\alpha}_1$, arbitrary signal energy $|a_1|^2$, and uniformly random signal phase $\theta = \arg[a_1]$. Also, l_1 is statistically independent of signal phase. Therefore, for known $\boldsymbol{\alpha}_1$, the GLRT receiver is the minimax receiver with respect to variations in the *a priori* statistics of θ. In view of this successful performance in the binary case, one is motivated to find a generalization of the GLRT appropriate to M-ary hypotheses.

One approach to the generalization of (6) is to begin with the simpler problem in which: i) the parameters $\boldsymbol{\alpha}_i$ are known; ii) the amplitudes $\boldsymbol{a}_i$ are random with known *a priori* densities $p(\boldsymbol{a}_i)$; and iii) the *a priori* probabilities of the H_i are *not* known. For these conditions it is reasonable to use a decision rule that maximizes conditional probabilities of correct decisions subject to constraints on conditional probabilities of error. For this purpose let P_{ji} denote the conditional probability of decision D_i, given that hypothesis H_j is true, and let E denote the error event. Then the problem is to maximize P_{ii} subject to constraints $\{P[E \mid H_j] = \beta_j; j \neq i\}$, where the β_j are given constants, and

$$P[E \mid H_j] = \sum_{n \neq j} P_{jn}. \tag{7}$$

The solution follows directly using Lagrange multipliers. Since $P[E \mid H_j] = 1 - P_{jj}$, fixing $P[E \mid H_j]$ is equivalent to fixing P_{jj}. Therefore, it is equivalent to maximize

$$\sum_n \lambda_n P_{nn} = \sum_n \lambda_n \int_{Z_n} p(\boldsymbol{r} \mid H_n)\, d\boldsymbol{r} \tag{8}$$

where λ_i is an arbitrary positive constant and the λ_n, $n \neq i$, are Lagrange multipliers. This leads to the test: "Choose H_m, if $\lambda_m p(\boldsymbol{r} \mid H_m) > \lambda_n p(\boldsymbol{r} \mid H_n)$, for all $n \neq m$." We write this test as

$$\lambda_i p(\boldsymbol{r} \mid H_i) \underset{\bar{H}_i}{\overset{\bar{H}_j}{\gtrless}} \lambda_j p(\boldsymbol{r} \mid H_j), \qquad i > j = 0, 1, \cdots, M-2 \tag{9}$$

where $\bar{H}_v$ is the complement of H_v. It is evident from (9) that the decision process requires $M-1$ binary decisions, since if any $\lambda_v p(\boldsymbol{r} \mid H_v)$ is less than another, hypothesis H_v is excluded from further consideration. Dividing both sides by $p(\boldsymbol{r} \mid H_0)$ and taking logarithms, (9) becomes

$$\ln \Lambda_i \underset{\bar{H}_i}{\overset{\bar{H}_j}{\gtrless}} \ln \Lambda_j + \gamma_{ij}, \qquad i > j = 0, 1, \cdots, M-2 \tag{10}$$

where Λ_n is the likelihood ratio

$$\Lambda_n = \frac{p(\boldsymbol{r} \mid H_n)}{p(\boldsymbol{r} \mid H_0)}, \qquad n = 0, 1, \cdots, M-1 \tag{11}$$

and

$$\gamma_{ij} = \ln \lambda_j - \ln \lambda_i. \tag{12}$$

We note that the test thresholds in (12) are related by

$$\gamma_{ij} = \sum_{n=j+1}^{i} \gamma_{n,n-1}, \qquad i > j = 0, 1, \cdots, M-2. \tag{13}$$

There are $M-1$ Lagrange multipliers λ_n that relate to the γ_{ij} and, therefore, to the determination of the $P[E \mid H_j]$. When the γ_{ij} can be found to satisfy the original conditional error constraints, $\{P[E \mid H_j] = \beta_j; j \neq i\}$, a solution to the maximization problem exists, and the test is completely specified. From the symmetry of (8), one sees that for any choice of thresholds constrained by (13) *each* P_{ii}, $i = 0, 1, \cdots, M-1$, will be the maximum possible of any receiver having the same $P[E \mid H_j]$. (A more rigorous proof of this property is given in [13].)

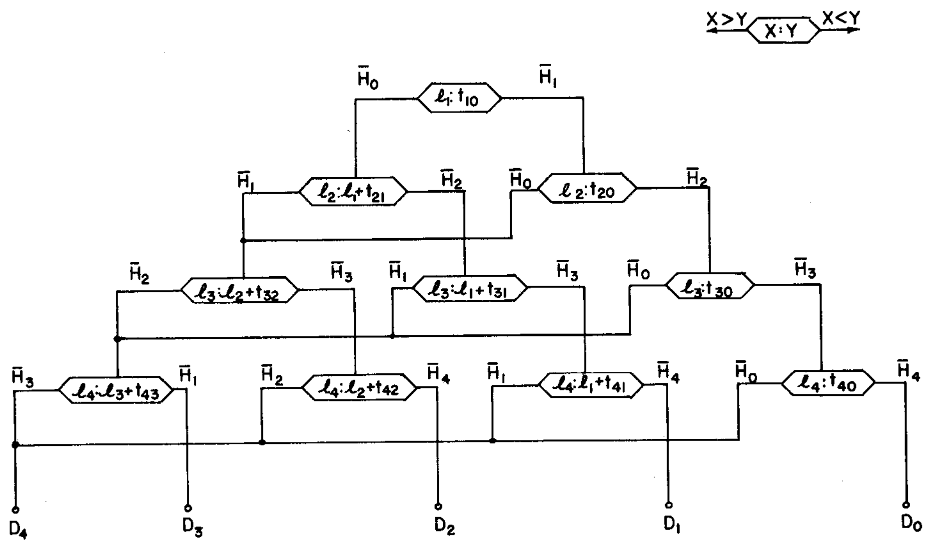

Fig. 1. MGLRT logic for $M = 5$ hypotheses.

Considering now unknown $\boldsymbol{a},\boldsymbol{\alpha}$, we are motivated by the structure of the generalized Neyman–Pearson test (10) to *define* an MGLRT by (10), where Λ_n of (11) is replaced by the generalized likelihood ratio

$$\Lambda_n = \max_{\boldsymbol{a}_n, \{\boldsymbol{\alpha}\}_n} \frac{p(\boldsymbol{r} \mid \boldsymbol{a}_n, \{\boldsymbol{\alpha}\}_n, H_n)}{p(\boldsymbol{r} \mid H_0)}, \quad n = 0, 1, \cdots, M-1 \tag{14}$$

and the γ_{ij} are an arbitrary set of thresholds related by (13). It is implicit in (14) that the maximization with respect to the k_n-tuple $\{\boldsymbol{\alpha}\}_n$ is over the space $\mathscr{B}_n$ of possible parameter k_n-tuples. For data of the form (4) the generalized likelihood ratio becomes

$$\Lambda_n = \max_{\boldsymbol{a}_n, \{\boldsymbol{\alpha}\}_n} \exp\left\{\frac{2}{N_0} \operatorname{Re}[\boldsymbol{\Omega}_n^* \boldsymbol{r}] - \frac{1}{N_0} |\boldsymbol{\Omega}_n|^2\right\}, \qquad n = 0, 1, \cdots, M-1. \tag{15}$$

The value of $\boldsymbol{a}_n$ for which Λ_n is maximum is the maximum-likelihood amplitude estimate

$$\hat{\boldsymbol{a}}_n = [S_n^* S_n]^{-1} S_n^* \boldsymbol{r} \tag{16}$$

where $[S_n^* S_n]^{-1}$ exists by assumption of algebraic independence of the $s(t,\boldsymbol{\alpha}_n)$. It can be seen from (16) that if the true signal is $\boldsymbol{\Omega}_n = S_n \boldsymbol{a}_n$, then $\hat{\boldsymbol{a}}_n = \boldsymbol{a}_n + \boldsymbol{e}$ where the estimation error $\boldsymbol{e}$ is zero mean complex Gaussian with covariance $N_0[S_n^* S_n]^{-1}$. Substitution of (16) into (15) yields

$$\Lambda_n = \max_{\{\boldsymbol{\alpha}\}_n} \exp\left\{\frac{\boldsymbol{r}^* Q_n \boldsymbol{r}}{N_0}\right\}, \qquad n = 0, 1, \cdots, M-1 \tag{17}$$

where Q_n is the idempotent Hermitian matrix

$$Q_n = \begin{cases} 0, & n = 0 \\ S_n[S_n^* S_n]^{-1} S_n^*, & n = 1, 2, \cdots, M-1. \end{cases} \tag{18}$$

From (17) and the idempotence of Q_n, (10) becomes

$$l_i \underset{\bar{H}_i}{\overset{\bar{H}_j}{\gtrless}} l_j + t_{ij}, \qquad i > j = 0, 1, \cdots, M-2 \tag{19}$$

where

$$l_n = \max_{\{\boldsymbol{\alpha}\}_n} |Q_n \boldsymbol{r}|^2 \tag{20}$$

and

$$t_{ij} = N_0 \gamma_{ij}. \tag{21}$$

The decision logic of (19) is illustrated in Fig. 1 for $M = 5$ hypotheses.

It is easily verified from the form of Q_n in (18) that $Q_n \boldsymbol{r}$ is the projection of $\boldsymbol{r}$ onto the column space $\mathscr{S}_n$ of S_n, and from (4b) and (16) that this projection is the estimated composite signal $\boldsymbol{\Omega}_n$. The statistic l_n is the energy in this projection. Therefore, the search for $\{\boldsymbol{\alpha}\}_n$ to maximize l_n may be interpreted as a search for that signal space $\mathscr{S}_n(\{\hat{\boldsymbol{\alpha}}\}_n)$ that contains the maximum energy.

It will be useful in the sequel to rewrite the MGLRT in the form

$$l_{ij} \underset{\bar{H}_i}{\overset{\bar{H}_j}{\gtrless}} t_{ij}, \qquad i > j = 0, 1, \cdots, M-2 \tag{22}$$

where

$$l_{ij} = \max_{\{\boldsymbol{\alpha}\}_i} \min_{\{\boldsymbol{\alpha}\}_j} \boldsymbol{r}^* Q_{ij} \boldsymbol{r} = l_i - l_j \tag{23}$$

with

$$Q_{ij} = Q_i - Q_j. \tag{24}$$

It is interesting to note that (19) can also be "derived" from the Bayes test for the random $\boldsymbol{a},\boldsymbol{\alpha}$ case obtained by Nilsson [2] upon replacing Nilsson's conditional mean signal estimates by maximum likelihood estimates and his thresholds by arbitrary γ_{ij} that satisfy (13). The two "derivations" of the proposed test lend some encouragement for its use. However, this is countered by Root's [3] objection to applying a GLRT to the signal resolution task: for any $i > j$, hypothesis H_i includes hypothesis H_j as a special case, where $k_i - k_j$ of the signals have zero amplitudes. Therefore, for zero t_{ij}, the GLRT will always select H_i (eventually H_{M-1}) rendering its application useless. Since either decision H_i or H_j can occur for $t_{ij} > 0$, however, Root's objection seems more concerned with the threshold values used than with the basic test structure. In fact, the same considerations apply to the conventional binary GLRT described by expression (6).

IV. Properties of M-ary Generalized Likelihood

Preliminary insight into the significance of test (19) is obtained by considering the implications of the previous observation that l_n is the energy in the signal space $\mathscr{S}_n$. The energy in any space does not depend upon the particular coordinates assumed to define that space. (Note the invariance of Q_n to the transformation $S_n \to S_nT$, where T is nonsingular.) Therefore, it is the space itself rather than the coordinates $s(\boldsymbol{\alpha}_n)$ that enters the decision rule, and hypotheses (4) are equivalent to

$$H_i: \boldsymbol{r} = \boldsymbol{\Omega}_i + \boldsymbol{w} \tag{25}$$

where $\boldsymbol{\Omega}_i$ is any (i.e., an unknown) vector in $\mathscr{S}_i$. The implication of *unknown* $\boldsymbol{\alpha}_n$, therefore, is that no vector, and hence no direction, in $\mathscr{S}_i$ is preferred *a priori*. This suggests the following.

Theorem 1: If the $\boldsymbol{\alpha}_n$ are known and the $\boldsymbol{\Omega}_n$ are random with spherically symmetric probability density functions (pdf's) $p(\boldsymbol{\Omega}_n)$, then $L = \{l_v;\, v = 1,\cdots, M-1\}$ is a set of sufficient statistics for decision D. Conversely, if any $p(\boldsymbol{\Omega}_n)$ is not spherically symmetric, then L is not a sufficient set.

Proof: Writing

$$\frac{p(\boldsymbol{r} \mid H_n)}{p(\boldsymbol{r} \mid H_0)} = \frac{p(l_n \mid H_n)p(\boldsymbol{y} \mid l_n, H_n)}{p(l_n \mid H_0)p(\boldsymbol{y} \mid l_n, H_0)} \tag{26}$$

where $(l_n, \boldsymbol{y})$ specifies $\boldsymbol{r}$, it follows that l_n is a sufficient statistic if and only if

$$p(\boldsymbol{y} \mid l_n, H_n) = p(\boldsymbol{y} \mid l_n, H_0). \tag{27}$$

$\boldsymbol{r}$ is the sum of its projections on $\mathscr{S}_n$ and the complementary space $\mathscr{S}_n^c$, i.e.,

$$\boldsymbol{r} = \boldsymbol{r}_n + \boldsymbol{r}_c \tag{28}$$

where

$$\boldsymbol{r}_n = Q_n\boldsymbol{r} \tag{29a}$$

and

$$\boldsymbol{r}_c = [I - Q_n]\boldsymbol{r}. \tag{29b}$$

The vector $\boldsymbol{r}_n$ can be expressed in terms of its length $|Q_n\boldsymbol{r}| = \sqrt{l_n}$ and a unit direction vector $\boldsymbol{\xi}_n$

$$\boldsymbol{r}_n = \sqrt{l_n}\, \boldsymbol{\xi}_n. \tag{30}$$

Setting $\boldsymbol{y} = (\boldsymbol{\xi}_n, \boldsymbol{r}_c)$ then permits specification of $\boldsymbol{r}$ by $(l_n, \boldsymbol{y})$. It remains to establish the conditions for which

$$p(\boldsymbol{\xi}_n, \boldsymbol{r}_c \mid l_n, H_n) = p(\boldsymbol{\xi}_n, \boldsymbol{r}_c \mid l_n, H_0). \tag{31}$$

Because $\boldsymbol{w}$ is WCG, $\boldsymbol{r}_c$ is independent of H and $\boldsymbol{r}_n$, therefore, of H, $\boldsymbol{\xi}_n$, and l_n. Therefore, (31) becomes

$$p(\boldsymbol{\xi}_n \mid l_n, H_n) = p(\boldsymbol{\xi}_n \mid l_n, H_0). \tag{32}$$

Under H_0, $\boldsymbol{r}_n$ equals the projection $\boldsymbol{w}_n = Q_n\boldsymbol{w}$ of $\boldsymbol{w}$ onto $\mathscr{S}_n$. Since $\boldsymbol{w}_n$ is WCG, its pdf is spherically symmetric with $\boldsymbol{\xi}_n$ independent of $l_n = |\boldsymbol{w}_n|^2$. Spherical symmetry of $p(\boldsymbol{w}_n)$ implies that $\boldsymbol{\xi}_n$ points uniformly in all directions of $\mathscr{S}_n$ in the sense that the statistics of the projection of $\boldsymbol{\xi}_n$ onto any fixed unit vector $\boldsymbol{\eta} \in \mathscr{S}_n$ do not depend on $\boldsymbol{\eta}$. Since $\boldsymbol{\xi}_n$ is a unit vector, equality holds in (32) if and only if this property is retained when $\boldsymbol{\xi}_n$ is conditioned on l_n and H_n. Given H_n and l_n one has

$$H_n: \boldsymbol{r}_n = \boldsymbol{\Omega}_n + \boldsymbol{w}_n = \sqrt{l_n}\, \boldsymbol{\xi}_n \tag{33}$$

and

$$\boldsymbol{\eta}^*\boldsymbol{\xi}_n = \frac{\boldsymbol{\eta}^*\boldsymbol{\Omega}_n}{\sqrt{l_n}} + \frac{\boldsymbol{\eta}^*\boldsymbol{w}_n}{\sqrt{l_n}}. \tag{34}$$

Since $\boldsymbol{w}_n$ is independent of $\boldsymbol{\Omega}_n$ and l_n is given, the joint characteristic function $\Phi(\omega_1,\omega_2)$ of the random variable pair $(\mathrm{Re}\,\{\boldsymbol{\eta}^*\boldsymbol{\xi}_n\}, \mathrm{Im}\,\{\boldsymbol{\eta}^*\boldsymbol{\xi}_n\})$ equals the product of the joint characteristic functions $\Phi_1(\omega_1,\omega_2)$ and $\Phi_2(\omega_1,\omega_2)$ of the pairs $(\mathrm{Re}\,\{\boldsymbol{\eta}^*\boldsymbol{\Omega}_n/\sqrt{l_n}\}, \mathrm{Im}\,\{\boldsymbol{\eta}^*\boldsymbol{\Omega}_n/\sqrt{l_n}\})$ and $(\mathrm{Re}\,\{\boldsymbol{\eta}^*\boldsymbol{w}_n/\sqrt{l_n}\}, \mathrm{Im}\,\{\boldsymbol{\eta}^*\boldsymbol{w}_n/\sqrt{l_n}\})$, respectively,

$$\Phi(\omega_1,\omega_2) = \Phi_1(\omega_1,\omega_2)\Phi_2(\omega_1,\omega_2). \tag{35}$$

Since $\boldsymbol{w}_n$ is spherically symmetric, $\Phi_2(\omega_1,\omega_2)$ does not depend on $\boldsymbol{\eta}$. If $p(\boldsymbol{\Omega}_n)$ is spherically symmetric, $\Phi_1(\omega_1,\omega_2)$ and hence $\Phi(\omega_1,\omega_2)$ will not depend on $\boldsymbol{\eta}$. Therefore, spherical symmetry of $p(\boldsymbol{\Omega}_n)$ implies spherical symmetry of $p(\boldsymbol{\xi}_n \mid l_n, H_n)$, and equality holds in (32). Conversely, absence of spherical symmetry in $p(\boldsymbol{\Omega}_n)$ implies that $\Phi_1(\omega_1,\omega_2)$ and hence $\Phi(\omega_1,\omega_2)$ depends on $\boldsymbol{\eta}$. Therefore, if $p(\boldsymbol{\Omega}_n)$ is not spherically symmetric, neither is $p(\boldsymbol{\xi}_n \mid l_n, H_n)$ and it cannot equal $p(\boldsymbol{\xi}_n \mid l_n, H_0)$. Q.E.D.

It is interesting to pursue the implications of spherical symmetry in $p(\boldsymbol{\Omega}_n)$ further. Under spherically symmetric $p(\boldsymbol{\Omega}_n)$ equality holds in (27), and the generalized Neyman–Pearson test (10) assumes the form

$$f_i(l_i) \underset{\overline{H}_i}{\overset{\overline{H}_j}{\gtrless}} f_j(l_j) + \gamma_{ij}, \qquad i > j = 0, 1, \cdots, M-2 \tag{36}$$

where

$$f_n(l_n) = \ln \frac{p(l_n \mid H_n)}{p(l_n \mid H_0)} \tag{37}$$

the parameters being assumed known. Under hypothesis H_n and for given $\boldsymbol{\Omega}_n$, the quantity $l_n = |\boldsymbol{\Omega}_n + \boldsymbol{w}_n|^2$ is the sum of the absolute squares of k_n independent complex Gaussian random variables each having variance N_0 and nonzero complex mean. The conditional density function $p(l_n \mid H_n, \boldsymbol{\Omega}_n)$ has, as is well known, the form $q_{k_n}(l_n; |\boldsymbol{\Omega}_n|^2)$ where the function $q_n((\cdot);(\cdot))$ is given in (48). Since $p(l_n \mid H_n, \boldsymbol{\Omega}_n)$ depends only on the energy in $\boldsymbol{\Omega}_n$, it also follows that $p(l_n \mid H_n, |\boldsymbol{\Omega}_n|^2) = q_{k_n}(l_n; |\boldsymbol{\Omega}_n|^2)$. Under the null hypothesis we have similarly, $p(l_n \mid H_0) = q_{k_n}(l_n; 0)$. Therefore,

$$\frac{p(l_n \mid H_n, |\boldsymbol{\Omega}_n|^2)}{p(l_n \mid H_0)} = \Gamma(k_n) \exp\left\{-\frac{|\boldsymbol{\Omega}_n|^2}{N_0}\right\} \frac{I_{k_n-1}\{(2|\boldsymbol{\Omega}_n|\sqrt{l_n})/N_0\}}{[(|\boldsymbol{\Omega}_n|\sqrt{l_n})/N_0]^{k_n-1}}. \tag{38}$$

The function $p(l_n \mid H_n, |\boldsymbol{\Omega}_n|^2)/p(l_n \mid H_0)$ is monotonically increasing in l_n, for every $|\boldsymbol{\Omega}_n|^2$. Considering binary hypotheses with $M = 2$, therefore, one then finds that the optimum test $f_1(l_1) \gtrless \gamma_{10}$ of (36) is identical to the MGLRT $l_1 \gtrless t_{10}$ of (19), for $t_{10} = f_1^{-1}(\gamma_{10})$. Therefore, the MGLRT is *optimum* in the Neyman–Pearson sense for binary hypotheses under known parameters and arbitrary spherically symmetric $p(\boldsymbol{\Omega}_1)$. By inspection of (36) and (38), however, it is clear that for $M \geq 3$ the sufficiency of L does not imply optimality of (19) for arbitrary spherically symmetric $p(\boldsymbol{\Omega}_n)$.

If (19) is to be optimum for $M \geq 3$, the first degree presence of the l_n and the form of (13) requires that the likelihood ratio $\Lambda_n = p(\boldsymbol{r} \mid H_n)/p(\boldsymbol{r} \mid H_0)$ have the form $b_n \exp(al_n/N_0)$, where a and b_n are positive constants independent of $\boldsymbol{r}$. This and the whiteness of $\boldsymbol{r}$ under H_0 in turn implies

$$p(\boldsymbol{r} \mid H_n) = b_n' \exp\left\{-\frac{1}{N_0}\,\boldsymbol{r}^*[I - aQ_n]\boldsymbol{r}\right\} \tag{39}$$

where the constant b_n' is determined by the condition $\int_{-\infty}^{+\infty} p(\boldsymbol{r} \mid H_n)\,d\boldsymbol{r} = 1$. Since

$$\frac{1}{N_0}\,[I - aQ_n] = [M_0Q_n + N_0I]^{-1} \tag{40}$$

where

$$M_0 = \frac{a}{1-a}\,N_0 \tag{41}$$

(39) can be written as

$$p(\boldsymbol{r} \mid H_n) = b_n' \exp\{-\boldsymbol{r}^*[M_0Q_n + N_0I]^{-1}\boldsymbol{r}\}. \tag{42}$$

It is easy to show that the pdf (42) results from $\boldsymbol{r} = \boldsymbol{\Omega}_n + \boldsymbol{w}$, if and only if *a priori* $\boldsymbol{\Omega}_n$ is complex Gaussian with zero mean and covariance matrix M_0Q_n. For such $\boldsymbol{\Omega}_n$, a standard calculation reduces (10) to

$$l_i \underset{H_i}{\overset{H_j}{\gtrless}} l_j + t_{ij}', \qquad i > j = 0, 1, \cdots, M-2 \tag{43a}$$

where

$$t_{ij}' = N_0\,\frac{M_0 + N_0}{M_0}\left\{(k_i - k_j)\ln\frac{N_0 + M_0}{N_0} + \ln\lambda_j - \ln\lambda_i\right\}. \tag{43b}$$

We summarize these results in the following theorem.

Theorem 2: If $M = 2$ and $\{\boldsymbol{\alpha}\}_1$ is known, the MGLRT is optimum for arbitrary k_1 and arbitrary spherically symmetric $p(\boldsymbol{\Omega}_1)$. If $M \geq 3$ and the parameters are known, the MGLRT is optimum if and only if each $\boldsymbol{\Omega}_n$ is *a priori* WCG in $\mathscr{S}_n$ with $E\{|\boldsymbol{\Omega}_n|^2\} = k_nM_0$, where M_0 is a positive scalar.

Attempts to describe the performance of the MGLRT for arbitrary signal configurations have not been successful. However, tractable results have been obtained in the special case for which the signal spaces are related by $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$. These results have special importance since they lead to statements concerning certain minimax properties of (19).

Theorem 3: If in the known parameter case the signal spaces are related by $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$, then $\Delta L = \{l_1, l_{21}, \cdots, l_{M-1,M-2}\}$ is a set of independent random variables, with moment generating functions given by

$$E\{\exp[vl_{i+1,i}] \mid H_m\} = \frac{\exp\{vE_{i+1,i,m}/(1 - vN_0)\}}{(1 - vN_0)^{(k_{i+1} - k_i)}}$$

$$\text{Re}\,\{v\} < \frac{1}{N_0} \tag{44a}$$

where

$$E_{ijm} = \boldsymbol{\Omega}_m{}^*Q_{ij}\boldsymbol{\Omega}_m = E_{im} - E_{jm} \tag{44b}$$

and

$$E_{nm} = |Q_n\boldsymbol{\Omega}_m|^2. \tag{44c}$$

The quantity E_{nm} is the energy in the projection of $\boldsymbol{\Omega}_m$ onto $\mathscr{S}_n$.

Proof: Because $\mathscr{S}_{i+1} \supset \mathscr{S}_i$, a basis for $\mathscr{S}_{i+1}$ may be obtained by augmenting the orthonormal basis for $\mathscr{S}_i$, say $\{\boldsymbol{\phi}_v;\, v = 1,2,\cdots,k_i\}$, with $k_{i+1} - k_i$ orthonormal functions $\{\boldsymbol{\phi}_v;\, v = k_i + 1, \cdots, k_{i+1}\}$ that are orthogonal to $\mathscr{S}_i$. The quantity $Q_{i+1,i}\boldsymbol{r} = Q_{i+1}\boldsymbol{r} - Q_i\boldsymbol{r}$ is the difference vector between the projections of $\boldsymbol{r}$ onto $\mathscr{S}_{i+1}$ and $\mathscr{S}_i$. Therefore,

$$Q_{i+1,i} = \sum_{v=k_i+1}^{k_{i+1}} \boldsymbol{\phi}_v\boldsymbol{\phi}_v{}^* \tag{45}$$

and

$$l_{i+1,i} = \sum_{v=k_i+1}^{k_{i+1}} |\boldsymbol{\phi}_v{}^*\boldsymbol{r}|^2. \tag{46}$$

Each basis $\boldsymbol{\phi}_n$ enters only one element in ΔL. Because of this and the fact that $\boldsymbol{w}$ is WCG, ΔL is a set of independent random variables. The expression for the moment generating function (44) follows from (46) in a straightforward manner using material in [5, ch. VII]. Q.E.D.

It follows from (44) and [14] that l_{ij}, $i > j$, has conditional pdf

$$p(l_{ij} \mid H_m) = q_{k_i - k_j}(l_{ij}; E_{ijm}) \tag{47}$$

where

$$q_n(x;E) = \frac{1}{N_0}\left(\frac{x}{E}\right)^{(n-1)/2} \exp\left\{-\frac{x+E}{N_0}\right\} I_{n-1}\left\{\frac{2\sqrt{xE}}{N_0}\right\}, \quad E \neq 0 \tag{48a}$$

and

$$q_n(x;0) = \frac{x^{n-1}}{N_0{}^n\Gamma(n)} \exp\left\{-\frac{x}{N_0}\right\} \tag{48b}$$

where $x \geq 0$.

Because the MGLRT satisfies Theorems 2 and 3, it is the minimax test with respect to variations in the conditional probabilities $p(\boldsymbol{a}_n \mid \{\boldsymbol{\alpha}\}_n, \{E_{in};\, i = 1,\cdots, M-1\})$, if $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$. This statement can be demonstrated in a conventional manner [15, p. 264] as follows. If the $\boldsymbol{\alpha}$ are known and if $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$, then by Theorem 3, $\boldsymbol{\Omega}_n$ enters the statistics of the l_v only through $\{E_{in};\; i = 1,\cdots, M-1\}$; therefore, the MGLRT performance does not depend on $p(\boldsymbol{a}_n \mid \{\boldsymbol{\alpha}_n\}, \{E_{in};\, i = 1,\cdots, M-1\})$. Furthermore, the performance of any other receiver is inferior to that of the MGLRT for $p(\boldsymbol{\Omega}_n)$ as described in Theorem 2. Therefore, no other receiver performs as well as the MGLRT for most adverse $p(\boldsymbol{a}_n \mid \{\boldsymbol{\alpha}_n\}, \{E_{in}; i = 1,\cdots, M-1\})$ when $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$. Hence the MGLRT is a minimax as claimed. This result generalizes the minimax properties of the binary test (6) to the M-ary case. The simplest example of this generalization is given by the binary case H_0, H_1 in which k_1 is an arbitrary positive integer. Here the test

$$l_1 \underset{H_0}{\overset{H_1}{\gtrless}} t_{10} \tag{49}$$

is minimax with respect to variations in $p(\boldsymbol{a}_1 \mid \boldsymbol{\alpha}_1, E_{11})$, where $E_{11} = |\boldsymbol{\Omega}_1|^2$. The receiver operating characteristic follows easily from (47)

$$P_D = Q_{k_1}\left[\sqrt{\frac{2E_{11}}{N_0}}, \sqrt{2\gamma_{10}}\right] \tag{50a}$$

$$P_F = \frac{1}{\Gamma(k_1)} \int_{\gamma_{10}}^{\infty} t^{k_1 - 1} e^{-t}\, dt \tag{50b}$$

where $Q_k(x,y)$ is the generalized Marcum Q function [10, p. 245]

$$Q_k(x,y) = \int_y^\infty z \left(\frac{z}{x}\right)^{k-1} \exp\left\{-\frac{z^2 + x^2}{2}\right\} I_{k-1}(xz)\, dz. \tag{51}$$

We emphasize that the minimax property discussed in the preceding paragraph requires $\mathscr{S}_1 \subset \mathscr{S}_2 \subset \cdots \subset \mathscr{S}_{M-1}$. If the signal spaces are not related in this manner, the property as stated does not apply. It is easy to show more generally that, because $\boldsymbol{w}$ is WCG, the statistics of the l_n are invariant to the transformation $\boldsymbol{\Omega}_m \rightarrow \boldsymbol{\Omega}_m e^{j\theta}$. Therefore, the MGLRT is minimax with respect to variations in $p(\theta)$ for known $\boldsymbol{\alpha}$. Attempts to find other conditions for which the MGLRT may be minimax have been frustrated by the difficulty in obtaining a general closed form for the joint characteristic function of the l_n.

Theorem 4 is motivated by the question of how to select the MGLRT thresholds. Interestingly, a solution to this problem is most easily obtained when the parameters are *unknown.*

Theorem 4: Let the parameter spaces $\mathscr{B}_n$ permit the possibility of orthogonal $\{\boldsymbol{s}(\hat{\boldsymbol{\alpha}}_n)\}$, and let $\mathscr{B}_{i+1} \supset \mathscr{B}_i$, for $i = 1, 2, \cdots, M-2$. Let $k_i = i$, for $i = 1, 2, \cdots, M-1$. Then the MGLRT reduces to a sequence of conventional binary tests as in (6) provided: i) the estimated signal coordinates $\{\boldsymbol{s}(\hat{\boldsymbol{\alpha}})\}$ are orthogonal; and ii) the test thresholds are given by $t_{ij} = (i-j)t_{10}$.

Proof: For orthogonal $\{\boldsymbol{s}(\hat{\boldsymbol{\alpha}})\}$, l_i becomes

$$\begin{aligned} l_i &= \max_{\{\alpha\}_i} |Q_i \boldsymbol{r}|^2 \\ &= |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_1)\boldsymbol{r}|^2 + |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_2)\boldsymbol{r}|^2 + \cdots + |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_i)\boldsymbol{r}|^2 \end{aligned} \tag{52a}$$

where we have arbitrarily labeled the $\hat{\boldsymbol{\alpha}}$ for

$$|\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_1)\boldsymbol{r}|^2 \geq |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_2)\boldsymbol{r}|^2 \geq \cdots \geq |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_i)\boldsymbol{r}|^2. \tag{52b}$$

Since l_i is associated with the maximum energy ($k_i = i$)-dimensional space $\mathscr{S}_i$, and we are considering orthogonal $\boldsymbol{s}(\hat{\boldsymbol{\alpha}})$, then l_{i+1} will have the form

$$\begin{aligned} l_{i+1} &= \max_{\{\alpha\}_{i+1}} |Q_{i+1}\boldsymbol{r}|^2 \\ &= |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_1)\boldsymbol{r}|^2 + \cdots + |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_i)\boldsymbol{r}|^2 + |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_{i+1})\boldsymbol{r}|^2 \end{aligned} \tag{53a}$$

where the first i terms equal those in (52a) and

$$|\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_{i+1})\boldsymbol{r}|^2 \leq |\boldsymbol{s}^*(\hat{\boldsymbol{\alpha}}_i)\boldsymbol{r}|^2. \tag{53b}$$

With $t_{ij} = (i-j)t_{10}$, the form of l_i ensures that if the lower inequality is satisfied in $l_{i+1} \gtrless l_i + t_{i+1,i}$, then $D = D_i$ immediately. Satisfaction of the upper inequality implies the next test $l_{i+2} \gtrless l_{i+1} + t_{i+2,i+1}$. It follows that the resulting sequence of tests, beginning with $l_1 \gtrless t_{10}$, is a series of conventional binary GLRT's. Q.E.D.

The conditions on the $\mathscr{B}_i$ stated in Theorem 4 are encountered frequently in practice; they are a generalization of the conditions normally arising in the problem of unknown signal arrival times on $[0,T]$. Theorem 4 provides a simultaneously simple and reasonable argument for setting thresholds as $t_{ij} = (i-j)t_{10}$ for such $\mathscr{B}_i$: in the event that the $\boldsymbol{s}(\hat{\boldsymbol{\alpha}})$ are orthogonal, these threshold values result in conventional receiver operation. The theorem also demonstrates that $t_{ij} > 0$ does not necessarily bias the test against the many-target hypothesis. Each dimension in $\mathscr{S}_i$ introduces an independent noise sample in l_i; $t_{ij} > 0$ may be viewed as arising because of this. In view of this result, it appears reasonable in the known parameter case to set $t_{ij} = (k_i - k_j)t_{10}$ for general $\mathscr{B}_i$. This is also suggested by (43b) with $\lambda_i = \lambda_j$, which represents the Bayes minimum $P[E]$ receiver subject to the conditions of Theorem 2 for equally probable $P[H_n]$. Alternatively, for small M one may compute M-ary receiver operating characteristics for general t_{ij}. The problem of optimally choosing the thresholds for arbitrary M for conditions not covered by Theorem 4 remains essentially unsolved, however, since it is linked to the difficult and unsolved problem of describing the MGLRT performance for general signal configurations.

V. Conclusion

In summary, we have shown in this paper that the form of the binary GLRT can be extended in a logical manner to include multiple hypotheses containing correlated signals. We were unable to analyze the test performance in the general case but have noted conditions for which the test is optimum and other conditions for which it is minimax. A pleasing property is that the MGLRT reduces to a sequence of conventional binary GLRT's for uncorrelated signal estimates, and this property has provided a rationale for selecting the test thresholds. Perhaps the most important contribution of this effort is the disclosure of an explicit GLRT structure that has meaningful application to the problem of signal resolution.

Acknowledgment

I am grateful for several helpful questions and criticisms from the anonymous referees.

References

[1] C. W. Helstrom, "The resolution of signals in white, gaussian noise," *Proc. IRE*, vol. 43, pp. 1111–1118, Sept. 1955.
[2] N. J. Nilsson, "On the optimum range resolution of radar signals in noise," *IRE Trans. Inform. Theory*, vol. IT-7, pp. 245–253, Oct. 1961.
[3] W. L. Root, "Radar resolution of closely spaced targets," *IRE Trans. Mil. Electron.*, vol. MIL-6, pp. 197–204, Apr. 1962.
[4] A. A. Ksienski and R. B. McGhee, "Radar signal processing for angular resolution beyond the Rayleigh limit," *Radio Electron. Eng.*, pp. 161–174, Sept. 1967.
[5] C. W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. Elmsford, N.Y.: Pergamon, 1968.
[6] R. C. Kemerait and D. G. Childers, "Signal detection and extraction by cepstrum techniques," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 745–759, Nov. 1972.
[7] M. G. Lichtenstein and T. Y. Young, "The resolution of closely spaced signals," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 288–293, Mar. 1968.
[8] T. Y. Young, "A recursive method for signal resolution," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-5, pp. 46–51, Jan. 1969.
[9] I. Selin, "Estimation of the relative delay of two similar signals of unknown phases in white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-10, pp. 189–191, July 1964.
[10] A. D. Whalen, *Detection of Signals in Noise.* New York: Academic, 1971.
[11] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 3.* New York: Wiley, 1971.
[12] ——, *Detection, Estimation, and Modulation Theory, Part 1.* New York: Wiley, 1968.
[13] J. B. Thomas and J. K. Wolf, "On the statistical detection problem for multiple signals," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 274–280, July 1962.
[14] G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications.* New York: Van Nostrand Reinhold, 1954.
[15] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering.* New York: Wiley, 1965.

Noncoherent Detection of Periodic Signals

ROBERT M. GAGLIARDI, MEMBER, IEEE

***Abstract*—In this paper the optimal Bayes detector for a general periodic waveform having uniform delay and additive white Gaussian noise is examined. It is shown that the detector is much more complex than that for the well-known cases of pure sine waves (i.e., classical noncoherent detection) and narrow-band signals. An interpretation of the optimal processor is presented, and several implementations are discussed.**

CLASSICAL noncoherent detection is generally understood to be the detection of a sine wave with random phase or time delay in additive Gaussian noise. The problem is well documented in communication texts, and the Bayes optimal detector has been derived as both a matched envelope detector and a quadrature correlator-squaring device. These results have been expanded to include narrow-band bandpass signals as well [1]. However, the extension to a general noncoherent problem involving the detection of an arbitrary periodic signal with random time delay has received little attention. The most relevant documentation appears in the radar literature, where the problem is formulated as noncoherent detection of a periodic train of arbitrary radio frequency (RF) pulses [2], but in all cases the narrow-band assumption is imposed in order to derive an interpretable solution.

In this paper the general noncoherent problem is examined, with the objective of determining the processing required by the optimal detector. It should be pointed out that a particular case of practical interest occurs when the periodic signal is a baseband square wave.

Manuscript received April 30, 1974; revised October 31, 1974.
The author is with the Department of Electrical Engineering, University of Southern California, Los Angeles, Calif. 90007.

Let $p(t)$ be a general periodic deterministic signal having period t_0 and bounded energy. The signal is observed for T seconds with a random delay τ in the presence of additive white Gaussian noise $n(t)$. The observation time T will be taken as an integer multiple of t_0 for convenience, although the results become accurate approximations if $T \gg t_0$. The observable can, therefore, be written as

$$v(t) = p(t - \tau) + n(t), \qquad t \in (0,T). \tag{1}$$

For the noncoherent problem, we assume τ is uniformly distributed over $(0,t_0)$. The optimal (Bayes) detector for the signal is desired. Mathematically, the Bayes detector is that which computes the generalized likelihood ratio Λ obtained by averaging over τ. For the observable of (1), this becomes

$$\Lambda = C \int_0^{t_0} \exp\left[\frac{2}{N_0}\int_0^T v(t)p(t-\tau)\,dt\right] d\tau \tag{2}$$

where N_0 is the one-sided noise level and C depends upon $v(t)$ but not on τ. Since C can be computed without the use of $p(t)$, it is brought along simply as a constant in subsequent equations. This property of C also depends on the assumption concerning the relation of observation time and signal period. Since $p(t)$ is periodic, it admits a Fourier expansion, which allows its delayed version to be written as

$$p(t-\tau) = \sum_{k=0}^{\infty} a_k \sin\left[k\left(\frac{2\pi}{t_0}\right)t + \psi_k - k\theta\right] \tag{3}$$

where (a_k,ψ_k) are the harmonic amplitudes and phases of $p(t)$ and $\theta \triangleq 2\pi\tau/t_0$ is the uniformly distributed phase variable over $(0,2\pi)$. The delay τ, therefore, introduces a random phase to each harmonic of $p(t)$, but note that these