

Optimality in detecting targets with unknown location

Fred Nicolls and Gerhard de Jager

Department of Electrical Engineering, University of Cape Town, South Africa.

E-mail: fnicolls@eng.uct.ac.za, gdj@eng.uct.ac.za.

Abstract

An optimal test does not exist for the problem of detecting a known target with unknown location in additive Gaussian noise. A common solution uses a generalised likelihood ratio testing (GLRT) formalism, where a maximum likelihood estimate of the unknown location parameter is used in a likelihood ratio test. The performance of this test is commonly assessed by comparing it to the ideal matched filter, which assumes the target location known in advance. This comparison is of limited utility, however, since the fact that the location is unknown has a significant effect on the detectability of the target. We demonstrate that a uniformly most powerful invariant (UMPI) optimal test exists for a specific class of unknown target location problems, where observations are discrete and shifts are defined circularly. Since this approach explicitly models the location as unknown, an assessment of the suboptimality of competing tests becomes meaningful. It is shown that for certain examples in this class the GLRT performance is negligibly different from that of the optimal test.

Key words: Invariant detection, uniformly most powerful invariant (UMPI) test, generalised likelihood ratio test (GLRT), optimality, cyclic permutation invariance, unknown circular shift, unknown target location, circulant covariance matrices.

1 Introduction

Many problems in signal and image processing involve detecting a known target with unknown location in a sequence of data. This constitutes a composite hypothesis testing problem, with the actual target location playing the role of an unknown parameter under the target present hypothesis.

The most common structured approach to this problem is to use the generalised likelihood ratio test (GLRT) [1], which has been studied at length in the engineering and applied statistics literature. It involves obtaining a maximum

likelihood estimate of the location parameter, and performing a simple hypothesis test on the data under the assumption that the estimate is accurate. The procedure constitutes a special case of a plug-in classifier formulation [2], but is generally favoured on account of asymptotic optimality properties [3].

The popularity of the GLRT for the unknown target location problem is perhaps more accidental than planned: the formulation for transient targets very often leads naturally to an implementation in a sliding window framework. Here a detection statistic is calculated for sequential overlapping intervals of received data, and a decision of target presence made if any of the test values exceeds some threshold. This is a natural procedure to adopt in detection systems which operate continuously through time.

The use of the GLRT in signal processing is usually justified in terms of its asymptotic optimality properties: as estimates of the unknown parameters become more accurate, the GLRT performance approaches that of the ideal matched filter which assumes the values known. It is for this reason that one commonly sees the GLRT performance for a test assessed by comparing it to that of the ideal matched filter. Since the matched filter has perfect knowledge of all quantities involved in the detection it cannot be improved upon, and its performance represents an upper bound on the achievable performance of any test.

The bound is not tight, however, and a test may perform considerably worse than the ideal matched filter while still not being significantly suboptimal. That is, the presence of the unknown parameter can significantly change the detectability properties of the target. A more meaningful comparison requires that the unknown parameters be explicitly modelled as such.

An optimal test does not exist for the general problem of detecting a target with unknown location in additive noise. Such a test *does* however exist for one specific case, in particular where samples form a finite discrete sequence and shifts are defined circularly over the observation interval. In this paper an optimal test is developed under these conditions: an invariant hypothesis testing formulation is followed, where the tests considered are restricted to those which are invariant to cyclic permutations of the observations, and the best test is found within this class. This test is uniformly most powerful invariant (UMPI) under the invariance condition described.

Although the UMPI test has certain advantages over the GLRT for the unknown cyclic permutation target problem, it is argued that its main benefit lies in providing a baseline against which the GLRT can be compared. It is demonstrated for some simple targets that the differences between the GLRT and the optimal UMPI test performance are very small. This provides definitive justification for the traditional GLRT solution to the problem.

The need to define shifts circularly is quite a restriction, although in some signal analysis scenarios the condition is at least approximately appropriate [4]. Many real problems, on the other hand, involve the sampling of continuous-time signals over a finite time interval, which will instead result in targets being truncated at the edges of the observation. Unfortunately, this situation does not exhibit sufficient symmetry to afford a sensible invariance condition, and therefore does not aid in the selection of optimal tests. Nonetheless, if the observation interval is long and the targets short and transient then these edge effects may be expected to play a minimal role, and results obtained for the one case should extend to the other. Therefore results obtained for the GLRT under cyclic permutation invariance may be considered to be at least approximately appropriate for related cases with shifts defined differently.

Even without the significance to unknown location target detection, the test presented is interesting purely from an academic point of view, as an exercise in invariant hypothesis testing. As such, the formulation complements other ideal invariant detector formulations that have appeared in the literature, for example invariance to signal scaling, invariance to subspace interference [5,6], and invariance to unknown covariance elements [7,8].

The structure of this paper is as follows. In Section 2 the cyclic permutation invariant detection problem formulation is presented, and the standard GLRT solution presented. In Section 3 the optimal (UMPI) test for this problem is derived. Section 4 compares the performance of these two tests, using Monte-Carlo simulations for a simple set of signals. Section 5 briefly discusses extensions of the testing principle to other problems.

Throughout this paper, data sequences are represented as column vectors denoted by bold roman characters. The notation $\mathbf{x} : N[\mathbf{m}, \mathbf{C}]$ is used to signify that the random vector \mathbf{x} has a multivariate normal (MVN) distribution, with mean \mathbf{m} and covariance matrix \mathbf{C} .

2 Cyclic permutation invariant detection problem formulation

Suppose N samples x_1, \dots, x_N of data are observed. Under hypothesis H_0 , these samples are independent and identically distributed as $N[0, \sigma^2]$ — a more general case is considered in Section 3.5. Under hypothesis H_1 , some shifted version of the prototype target signal s_1, \dots, s_N is added to the noise samples. For discrete-time observations it is natural to regard shifts as cyclic permutations of the elements, so under H_1 the mean of the observations is some cyclic permutation of s_1, \dots, s_N . Figure 1 demonstrates four possible instances of target presence under the assumptions described.

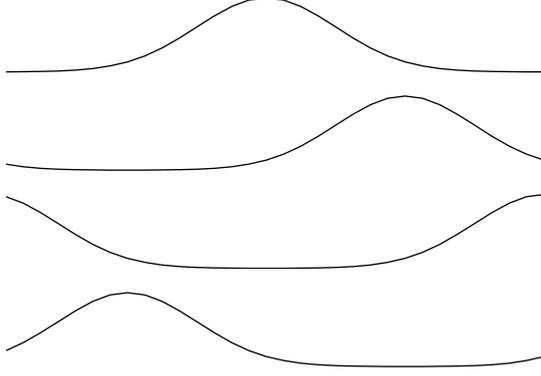


Fig. 1. Instances of four possible targets (Gaussian bumps) which may occur in the observed data interval.

The problem is more simply described in vector notation: letting $\mathbf{x} = (x_1, \dots, x_N)^T$, $\mathbf{n} = (n_1, \dots, n_N)^T$, and $\mathbf{s} = (s_1, \dots, s_N)^T$, the hypotheses are

$$H_0 : \mathbf{x} = \mathbf{n} \quad (1)$$

versus

$$H_1 : \mathbf{x} = \mathbf{P}^\theta \mathbf{s} + \mathbf{n}, \quad (2)$$

where $\mathbf{n} : N[\mathbf{0}, \sigma^2 \mathbf{I}]$ and \mathbf{P} is the cyclic permutation matrix

$$\mathbf{P} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}. \quad (3)$$

It can be seen that premultiplying the column vector \mathbf{x} by \mathbf{P} cyclically permutes the elements one position downwards:

$$\begin{pmatrix} 0 & \cdots & 0 & 1 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} x_n \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}. \quad (4)$$

Thus, under H_1 the mean of \mathbf{x} is some cyclic permutation of the signal vector \mathbf{s} , where the order of the permutation is unknown. The quantity θ is deterministic and unknown, and without loss of generality can be restricted to take on integer values from 0 to $N - 1$.

The conventional method of developing a test for the problem just described would be to use a generalised likelihood ratio test (GLRT) [1], also referred

to as the likelihood ratio test for composite hypotheses in the statistics literature [3]. This testing principle is commonly used in testing under composite hypotheses, where there are unknown (nuisance) parameters in the formulation. Effectively, maximum likelihood estimates are made of the unknown parameters under each hypothesis, and these estimates are substituted back into the associated probability density functions. The resulting density functions are then used in a conventional likelihood ratio test. Using this principle for the unknown target location problem, a maximum likelihood estimate of the unknown parameter θ is obtained from the observed data, under the assumption that H_1 is in force. This estimate is then substituted into a conventional likelihood ratio test for simple hypotheses.

It is easy to show that the GLRT statistic is

$$\Lambda_{\text{GLRT}}(\mathbf{x}) = \ln \frac{\max_{\theta \in [0, N-1]} p_1(\mathbf{x}|\theta)}{p_0(\mathbf{x})}, \quad (5)$$

where $p_0(\mathbf{x})$ is the probability density function (pdf) of the observation \mathbf{x} under H_0 , and $p_1(\mathbf{x})$ the pdf under H_1 . Assuming the distributions given in the previous section,

$$\begin{aligned} \Lambda_{\text{GLRT}}(\mathbf{x}) &= \max_{\theta \in [0, N-1]} -\frac{1}{2\sigma^2} \{(\mathbf{x} - \mathbf{P}^\theta \mathbf{s})^T (\mathbf{x} - \mathbf{P}^\theta \mathbf{s}) - \mathbf{x}^T \mathbf{x}\} \\ &= \frac{1}{\sigma^2} \left\{ \max_{\theta \in [0, N-1]} (\mathbf{P}^\theta \mathbf{s})^T \mathbf{x} + \frac{1}{2} \mathbf{s}^T \mathbf{s} \right\}. \end{aligned} \quad (6)$$

In the GLRT a decision is made by comparing this statistic to a threshold η :

$$\max_{\theta \in [0, N-1]} (\mathbf{P}^\theta \mathbf{s})^T \mathbf{x} \underset{H_0}{\overset{H_1}{\gtrless}} \sigma^2 \eta - \frac{1}{2} \mathbf{s}^T \mathbf{s}. \quad (7)$$

The final quantity on the left of this test is simply the maximum of the inner products between the observation \mathbf{x} and all possible cyclic permutations of the signal \mathbf{s} . Since the null hypothesis is independent of the unknown parameter θ , the threshold can be chosen so that the test has a constant false alarm rate.

3 Most powerful cyclic permutation invariant detection

The GLRT just described has no absolute optimality properties. As discussed in the introduction, an optimal formulation is highly desirable even if only to assess the suboptimality of this test. In this section an invariance argument is used to restrict the form of the tests considered for the problem, and an optimal test is found within this class.

Invariant detection is nicely treated in the book by Scharf [9]. In short, to obtain an optimal invariant test for a problem one needs to (1) find a suitable group of transformations of the observation within which to enforce invariance, (2) find the maximal invariant statistic for the problem under the different hypotheses, along with the corresponding probability densities, and (3) perform a test of hypotheses based on the value of this maximal invariant. The remainder of this section addresses each of these issues in turn, finally yielding the optimal invariant test for the problem under investigation.

3.1 Invariance of the hypothesis testing problem

There is no uniformly most powerful (UMP) test for the hypothesis testing problem under consideration here. In the search for an optimal characterisation, one option is therefore to restrict the class of tests which are to be considered. For the case of detecting a signal with unknown location, it is natural to require that the hypothesis test be constrained such that the same decision be made for arbitrarily shifted versions of any given observation. The transformation group is therefore taken to be

$$\mathcal{G} = \{g(\mathbf{x}) | g(\mathbf{x}) = \mathbf{P}^k \mathbf{x}, k = 0, \dots, N - 1\}. \quad (8)$$

This places an equivalence on the set of observations $\{\mathbf{P}^0 \mathbf{x}, \dots, \mathbf{P}^{N-1} \mathbf{x}\}$, which is natural on account of the symmetry of the elements of the observations under each hypothesis. Thus the observations

$$(x_1, \dots, x_{N-1}, x_N)^T \equiv (x_N, x_1, \dots, x_{N-1})^T \equiv \dots \equiv (x_2, \dots, x_n, x_1)^T \quad (9)$$

are all considered to be equivalent by the detector. From the way the hypothesis testing problem has been formulated, it cannot be said that enforcing this equivalence is restricting the form of the detector in any unreasonable way.

The testing problem is invariant to the group \mathcal{G} . This can be established by considering the distributions of the observation \mathbf{x} under each hypothesis: in both cases \mathbf{x} is MVN with covariance matrix $\sigma^2 \mathbf{I}$, under H_0 the mean is $\mathbf{0}$, and under H_1 the mean is one of the elements in the set $\{\mathbf{P}^0 \mathbf{s}, \dots, \mathbf{P}^{N-1} \mathbf{s}\}$. Consider now an element $g_k(\mathbf{x}) = \mathbf{P}^k \mathbf{x}$ of the group \mathcal{G} . Since this is a linear transformation of \mathbf{x} , the distribution of $\mathbf{y} = g_k(\mathbf{x})$ is $N[\mathbf{P}^k E\{\mathbf{x}\}, \sigma^2 \mathbf{P}^k (\mathbf{P}^k)^T]$, where $E\mathbf{x}$ is the expected value of \mathbf{x} . Noting now that $(\mathbf{P}^k)^T = \mathbf{P}^{N-k} = \mathbf{P}^{-k}$,

$$\mathbf{y} : N[\mathbf{P}^k E\{\mathbf{x}\}, \sigma^2 \mathbf{I}]. \quad (10)$$

Thus under H_0 the mean of the vector \mathbf{y} is $\mathbf{0}$, and under H_1 it is an element of the set $\{\mathbf{P}^k \mathbf{P}^0 \mathbf{s}, \dots, \mathbf{P}^k \mathbf{P}^{N-1} \mathbf{s}\} = \{\mathbf{P}^0 \mathbf{s}, \dots, \mathbf{P}^{N-1} \mathbf{s}\}$. The transformation g_k therefore preserves the form of the distribution and retains the partition of

the parameter space under each hypothesis. Since this conclusion is valid for all elements $g \in \mathcal{G}$, the hypothesis testing problem is invariant- \mathcal{G} .

3.2 Maximal invariant statistic for the problem

Sufficient statistics play an important role in hypothesis testing, in that they summarise the data without ignoring anything that is important to the decision problem. In invariant testing the maximum invariant statistic serves an analogous role: this is a statistic which has the required invariances, but which also captures all the useful information in the observation regarding the decision process.

For the invariant detection problem under discussion, one such statistic can be obtained by defining $\mathbf{P}_{\max}(\mathbf{x})$ to be that function which cyclically permutes the elements of the vector \mathbf{x} until the element with the maximum value is in the first position. Since the permutation is cyclic, the direction of the shift is irrelevant. Note that for the distributions considered here $\Pr\{x_i = x_j\} = 0$ for $i \neq j$, so the maximum element of \mathbf{x} will be unique with probability 1.

The statistic $\mathbf{P}_{\max}(\mathbf{x})$ is clearly invariant to the group \mathcal{G} : since one of the elements of \mathbf{x} is always maximum and elements of \mathcal{G} simply permute the observation \mathbf{x} cyclically, $\mathbf{P}_{\max}[g(\mathbf{x})] = \mathbf{P}_{\max}[\mathbf{x}]$ for all $g \in \mathcal{G}$. Additionally, for the same reasons, the condition $\mathbf{P}_{\max}[g(\mathbf{x}_1)] = \mathbf{P}_{\max}[g(\mathbf{x}_2)]$ means that \mathbf{x}_1 and \mathbf{x}_2 must be related to one another through a cyclic shift, so $\mathbf{x}_2 = g(\mathbf{x}_1)$ for some $g \in \mathcal{G}$. Thus the statistic $\mathbf{P}_{\max}(\mathbf{x})$ is maximal.

As explained by Lehmann [10] and Scharf [9], the significance of this result is that only functions of the maximal invariant statistic have to be considered when looking for a test which is invariant to \mathcal{G} .

The condition $\Pr\{x_i = x_j\} = 0$ for $i \neq j$ is met whenever \mathbf{x} is governed by a true probability density function. Strictly, the condition $\{x_i = x_j\}$ then defines a set of probability measure zero. Note that this does not mean that x_i cannot equal x_j , just that for purposes of probability assignment the event is insignificant. For example, if x and y are jointly normal, then the event $x = y$ is possible but has probability zero. The condition $\Pr\{x_i = x_j\} = 0$ for $i \neq j$ is not met if there is a point probability mass in the observation space (eg. a dirac delta in at least one location). However, this does not occur for any of the densities used in this paper and has not been considered further.

3.3 Distribution of the maximal invariant statistic

The method described by Hogg and Craig [11, p. 142] in relation to order statistics provides a means of determining the distribution of the maximal invariant. Firstly it is reasserted that two elements of \mathbf{x} are equal with probability zero, so the joint probability density of \mathbf{x} can be defined to be zero at all points which have at least two of their coordinates equal — this will not affect any probability arguments based on the distribution. The set \mathcal{A} where the probability density of \mathbf{x} is nonzero can therefore be partitioned into N mutually disjoint sets:

$$\begin{aligned}\mathcal{A}_1 &= \{\mathbf{x} | x_1 = \max(x_1, \dots, x_N)\} \\ &\vdots \\ \mathcal{A}_N &= \{\mathbf{x} | x_N = \max(x_1, \dots, x_N)\}.\end{aligned}\tag{11}$$

The set \mathcal{A}_i is thus the set of all points in \mathbb{R}^N that have no elements equal, and have x_i as their largest element.

Consider the function $\mathbf{y} = \mathbf{P}_{\max}(\mathbf{x})$, where \mathbf{y} is a vector in \mathbb{R}^N with elements defined according to $\mathbf{y} = (y_1, \dots, y_N)^T$. This defines a $1 - 1$ transformation of each of $\mathcal{A}_1, \dots, \mathcal{A}_N$ onto the same set \mathcal{A}_1 . For points in \mathcal{A}_i , the transformation $\mathbf{y} = \mathbf{P}_{\max}(\mathbf{x})$ cyclically permutes the elements of \mathbf{x} upwards by $i - 1$ positions. Thus the inverse function is

$$\mathbf{x} = \mathbf{P}^{i-1}\mathbf{y}\tag{12}$$

which simply rotates them back downwards by the same amount.

Letting J_i be the determinant of the Jacobian of the inverse transformation corresponding to \mathcal{A}_i ,

$$J_i = |\mathbf{P}^{i-1}|.\tag{13}$$

Now by the structure of \mathbf{P}^{i-1} , it is always possible to obtain an identity matrix by means of a number of row exchanges. Thus it must be the case that $J_i = +1$ or $J_i = -1$. Denoting the probability density of \mathbf{x} by $f_x(\mathbf{x})$, the results of this section can be combined to yield the pdf $f_y(\mathbf{y})$ of $\mathbf{y} = \mathbf{P}_{\max}(\mathbf{x})$ as [11, p. 143]

$$f_y(\mathbf{y}) = \begin{cases} \sum_{k=0}^{N-1} f_x(\mathbf{P}^k\mathbf{y}) & y_1 = \max(y_1, \dots, y_N) \\ 0 & \text{otherwise.} \end{cases}\tag{14}$$

This expression can be used to find the distribution of the maximal invariant statistic under each hypothesis. Under H_0 ,

$$f_x(\mathbf{x}) = (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{x}}.\tag{15}$$

Therefore the distribution of $\mathbf{y} = \mathbf{P}_{\max}(\mathbf{x})$ is

$$f_{\mathbf{y}}(\mathbf{y}) = \begin{cases} \sum_{k=0}^{N-1} (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{P}^k \mathbf{y})^T (\mathbf{P}^k \mathbf{y})} & y_1 = \max(y_1, \dots, y_N) \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

Once again using the relation $(\mathbf{P}^k)^T = \mathbf{P}^{-k}$, this can be simplified to

$$f_{\mathbf{y}}(\mathbf{y}) = \begin{cases} N(2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2} \mathbf{y}^T \mathbf{y}} & y_1 = \max(y_1, \dots, y_N) \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

When H_1 is in force, the situation is slightly more complex: now the mean of the observation takes some value in the set $\{\mathbf{P}^\theta \mathbf{s} | \theta = 0, \dots, N-1\}$. The probability density of \mathbf{x} is therefore

$$f_{\mathbf{x}}(\mathbf{x}) = (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{P}^\theta \mathbf{s})^T (\mathbf{x} - \mathbf{P}^\theta \mathbf{s})}, \quad (18)$$

where θ is some integer in the range 0 to $N-1$. Substituting into the expression for $f_{\mathbf{y}}(\mathbf{y})$ gives

$$f_{\mathbf{y}}(\mathbf{y}) = \begin{cases} \sum_{k=0}^{N-1} (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{P}^k \mathbf{y} - \mathbf{P}^\theta \mathbf{s})^T (\mathbf{P}^k \mathbf{y} - \mathbf{P}^\theta \mathbf{s})} & y_1 = \max(y_1, \dots, y_N) \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

The first case in this expression needs to be looked at in more detail: under the condition $y_1 = \max(y_1, \dots, y_N)$,

$$\begin{aligned} f_{\mathbf{y}}(\mathbf{y}) &= \sum_{k=0}^{N-1} (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} - 2\mathbf{s}^T \mathbf{P}^{-\theta} \mathbf{P}^k \mathbf{y} + \mathbf{s}^T \mathbf{s})} \\ &= (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} + \mathbf{s}^T \mathbf{s})} \sum_{k=0}^{N-1} e^{\frac{1}{\sigma^2} \mathbf{s}^T \mathbf{P}^{-\theta} \mathbf{P}^k \mathbf{y}}. \end{aligned} \quad (20)$$

Using the fact that $\mathbf{P}^{-l} = \mathbf{P}^{N-l}$ and noting that the sum in this expression involves exponentials of the inner product between \mathbf{s} and all cyclic permutations of \mathbf{y} , the range of the summation can be modified:

$$\sum_{k=0}^{N-1} e^{\frac{1}{\sigma^2} \mathbf{s}^T \mathbf{P}^{-\theta} \mathbf{P}^k \mathbf{y}} = \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2} \mathbf{s}^T \mathbf{P}^l \mathbf{y}}. \quad (21)$$

This yields the final pdf for \mathbf{y} under H_1 :

$$f_{\mathbf{y}}(\mathbf{y}) = \begin{cases} (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} + \mathbf{s}^T \mathbf{s})} \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2} \mathbf{s}^T \mathbf{P}^l \mathbf{y}} & y_1 = \max(y_1, \dots, y_N) \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

In both cases the density of the maximal invariant is seen to be independent of the unknown parameter θ , as required.

The formulation presented in this section applies to real \mathbf{x} , but analogous results can be obtained when \mathbf{x} is complex. In this case the form for the normal densities in Equations 15 and 18 should be changed to their complex counterparts, and the operator $\mathbf{P}_{\max}(\mathbf{x})$ must be redefined to cyclically permute \mathbf{x} until the element with the largest modulus is in the first position.

3.4 Optimal invariant likelihood ratio test

Once the observation \mathbf{x} has been mapped onto the corresponding maximal invariant statistic, a likelihood ratio test can be performed on this quantity. The likelihood ratio for the problem is

$$\begin{aligned} l(\mathbf{y}) &= \frac{(2\pi\sigma^2)^{-\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\mathbf{y}^T\mathbf{y}} e^{-\frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s}} \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{y}}}{N(2\pi\sigma^2)^{-\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\mathbf{y}^T\mathbf{y}}} \\ &= \frac{1}{N} e^{-\frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s}} \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{y}}. \end{aligned} \quad (23)$$

The log-likelihood ratio is therefore

$$L(\mathbf{y}) = -\ln N - \frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s} + \ln \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{y}}. \quad (24)$$

The best invariant test is to compare this ratio to a threshold η and decide H_1 when exceeded, yielding

$$\ln \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{y}} \underset{H_0}{\overset{H_1}{\geq}} \eta + \ln N + \frac{1}{2\sigma^2}\mathbf{s}^T\mathbf{s}. \quad (25)$$

The resulting test is most powerful out of all tests which share the same invariances. This means that no other test which is invariant to cyclic permutations of the observations can perform better, regardless of the value of the unknown parameter θ . Since the invariance is a reasonable one for the problem, it is fair to assert that this is the optimal test.

Again noting that the summation in the expression for this test goes over terms involving inner products between all cyclic permutations of \mathbf{y} with \mathbf{s} , it is evident that

$$\sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{y}} = \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2}\mathbf{s}^T\mathbf{P}^l\mathbf{x}}. \quad (26)$$

Thus the final expression for the test can be written in terms of the original observed data as

$$\ln \sum_{l=0}^{N-1} e^{\frac{1}{2\sigma^2} \mathbf{s}^T \mathbf{P}^l \mathbf{x}} \underset{H_0}{\overset{H_1}{>}} \eta + \ln N + \frac{1}{2\sigma^2} \mathbf{s}^T \mathbf{s}. \quad (27)$$

A test is uniformly most powerful invariant (UMPI) if it is invariant to a required group transformation, and it is more powerful than any other invariant test *for all possible values* of any unknown nuisance parameters. Since the maximal invariant test statistic contains no unknown parameters, the test just given is trivially UMPI.

It is worth noting that the estimation of the parameter θ is explicit in the GLRT of equation 7. Thus the most likely location of the detected signal is also provided by the test. For the UMPI test, however, the dependence on the parameter is completely eliminated from the problem by the invariance condition. At no point does this test make use of any estimate of θ , either implicitly or explicitly.

3.5 Extension to correlated noise

The previous tests were based on detection in a white noise environment. The results can be extended to the case where the noise has a known circulant covariance matrix. This is a special case of the general stationary condition, where the matrix has a Toeplitz structure. The constraint that the covariance matrix be circulant is required for the hypothesis testing problem remain invariant to cyclic permutations.

In the circulant covariance case, the hypotheses are as in equations 1 and 2, but now the distribution of the noise is $\mathbf{n} : N[\mathbf{0}, \mathbf{C}]$. Applying the (assumed invertible) whitening transformation $\mathbf{z} = \mathbf{C}^{-1/2} \mathbf{x}$ to the observed data, the hypotheses become $\mathbf{z} : N[\mathbf{0}, \mathbf{I}]$ under H_0 and $\mathbf{z} : N[\mathbf{C}^{-1/2} \mathbf{P}^\theta \mathbf{s}, \mathbf{I}]$ under H_1 . Now, if \mathbf{C} is circulant then $\mathbf{C}^{-1/2}$ is also circulant, so $\mathbf{C}^{-1/2} = \mathbf{P}^\theta \mathbf{C}^{-1/2} \mathbf{P}^{-\theta}$ [12]. Writing this as $\mathbf{C}^{-1/2} \mathbf{P}^\theta = \mathbf{P}^\theta \mathbf{C}^{-1/2}$, the distribution under H_1 is $\mathbf{z} : N[\mathbf{P}^\theta \mathbf{C}^{-1/2} \mathbf{s}, \mathbf{I}]$. This can be recognised as the problem of invariant detection of the modified signal $\mathbf{C}^{-1/2} \mathbf{s}$ in white noise. The test given in the previous section can therefore be used in this modified problem, and is once again UMPI.

Finally, it is noted that the components of a random vector with a circulant covariance matrix can be diagonalised by means of the discrete Fourier transform. This can provide a fast method of calculating the required test statistic. Note that the test complexity is the same as that of the GLRT — the only difference is in how the components of the test are combined into the final statistic.

4 Comparison of GLRT and UMPI tests

To compare the GLRT and UMPI tests, the distribution of the test statistics need to be found in each case under the two possible hypotheses. For both tests presented, expressions for the resulting distributions are complicated, and have no convenient and easily calculated closed form. In this section, possible approaches are presented for how the distributions could be estimated if necessary. For the remainder of this paper, however, direct Monte-Carlo methods are used to obtain the relevant test statistics, to avoid having to deal with issues related to accuracy of approximations.

4.1 Distributions of the test statistics

Summarising the results of the previous sections, the GLRT is

$$t_{\text{GLRT}}(\mathbf{x}) = \max_{\theta \in [0, N-1]} (\mathbf{P}^\theta \mathbf{s})^T \mathbf{x} \underset{H_0}{\overset{H_1}{>}} \eta_{\text{GLRT}}, \quad (28)$$

and the UMPI test is

$$t_{\text{UMPI}}(\mathbf{x}) = \ln \sum_{l=0}^{N-1} e^{\frac{1}{\sigma^2} \mathbf{s}^T \mathbf{P}^l \mathbf{x}} \underset{H_0}{\overset{H_1}{>}} \eta_{\text{UMPI}}. \quad (29)$$

The thresholds η_{GLRT} and η_{UMPI} are constants which are chosen to yield the desired test properties. Under a Neyman-Pearson formulation, these constants are commonly chosen to produce a test with a required false alarm rate.

The components of the sum in the UMPI statistic are statistically dependent upon one another, and each have a lognormal distribution. The UMPI statistic is therefore given by the logarithm of the sum of dependent lognormal variates. Sums of dependent and independent lognormal random variables have been discussed at some length in the literature, and no closed form is known for their distribution [13]. It is common therefore to approximate the sum by yet another lognormal distribution (which is appropriate particularly in the tails [14]), and perform identification by means of either moment or cumulant matching [15]. The effectiveness of this approach seems reasonable for some commonly occurring situations discussed in the references.

If the sum in the test statistic $t_{\text{UMPI}}(\mathbf{x})$ does indeed have an approximately lognormal distribution, then the test statistic itself should be approximately normal. To a high degree of accuracy, then, the mean and variance will completely specify the distribution of the statistic under each hypothesis. For the white noise case the methods presented by Schwartz and Yeh [15] can be used

to find the approximate distribution of the test statistics, while for the correlated case the generalisations discussed by Abu-Dayya and Beaulieu [13] and Safak [16] are appropriate. Therefore, although closed-form expressions for the distributions are not available, test thresholds can be selected to yield desired false alarm probabilities, and test power can additionally be calculated for the detector.

The distribution of the GLRT test statistic is even less tractable, and requires general results regarding the density of the maximum of a set of correlated random samples from a normal population. Afonja [17] presents a method for finding the first and second moments of this maximum in terms of the moments of truncated multivariate normal density functions. These can in turn be found if the general correlated multivariate normal probability integral can be calculated over rectangular regions of the probability space [18]. This calculation is computationally expensive and can only be performed over low dimensional spaces in practice. An alternative and more conventional method of finding an approximate distribution of the GLRT test statistic is to use known asymptotic results regarding maximum likelihood estimates in a testing framework [3]. This is however only accurate for low noise observations, which cannot in general be expected in signal processing.

4.2 Comparison for some specific transient targets

To avoid getting engrossed in the details of analytical approximations, the results in this paper are based instead on Monte-Carlo simulations. The distributions of the test statistics can be very easily obtained by applying the tests to a large sample of observations from any specific configuration of target and noise. The first set of three targets used for testing purposes is shown in Figure 2. In total, three different scalings of each of these targets are consid-

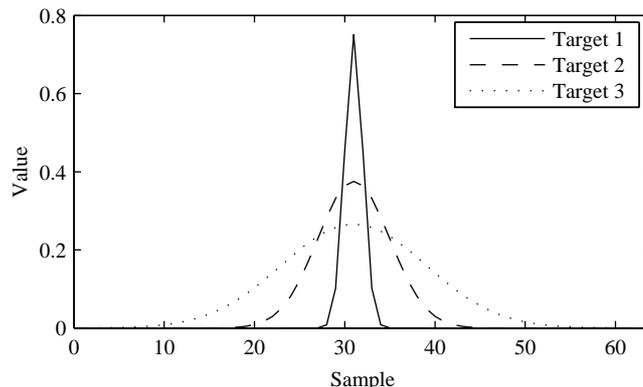


Fig. 2. First set of target signals used in Monte-Carlo simulations. The signals are Gaussians centered on 32 with standard deviations of 2, 4, and 8.

ered, corresponding to energies of 2, 4, and 8. The additive noise used in the simulations is comprised of 64 uncorrelated samples of zero-mean Gaussian noise with unit variance.

Figure 3 shows estimates of the distributions of the test statistics for the specific case where the targets are scaled to have an energy of 8. The line styles in each case correspond to those used for plotting the targets in Figure 2, and in all instances the curves to the right depict the distributions under H_1 . The distributions were estimated using 50000 noise vector samples and 50000

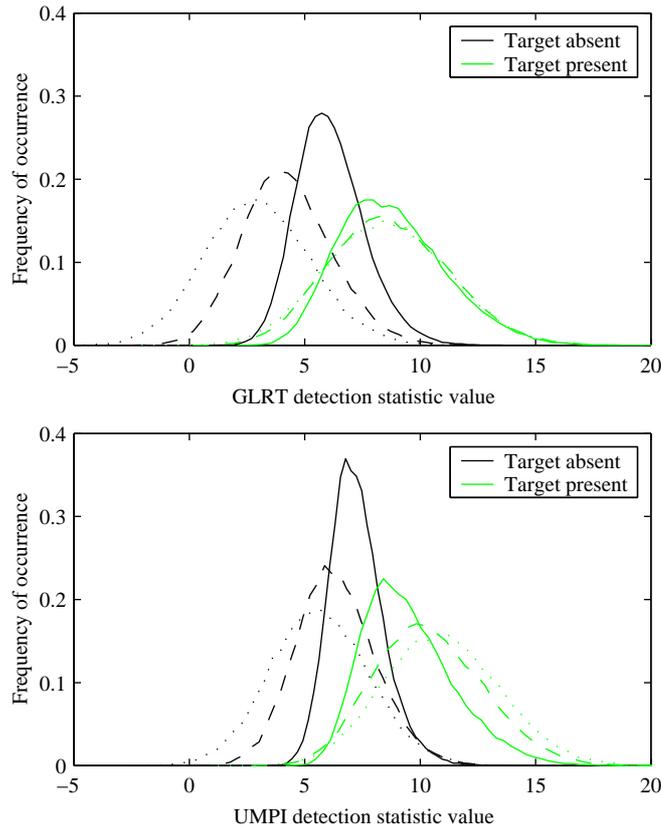


Fig. 3. Distributions of the GLRT statistic (top) and UMPI statistic (bottom) under both H_0 and H_1 for each of the signals tested. The signals were scaled to have an energy of 8.

noise plus target samples for each target at each scaling. The test statistic values given in Equations 28 and 29 were calculated for each of these samples, and frequency histograms generated using 50 bins over the range of values obtained.

As discussed in Section 4, it is common to approximate the sum of lognormal variates by a lognormal distribution. If this assumption were accurate for the scenario presented here, the distribution of the UMPI test statistic values would be approximately normal under both hypothesis. However, in Figure 3 it is clearly seen that the distributions under H_1 are quite considerably skewed

to the right. Blindly applying approximations in this case could be misleading.

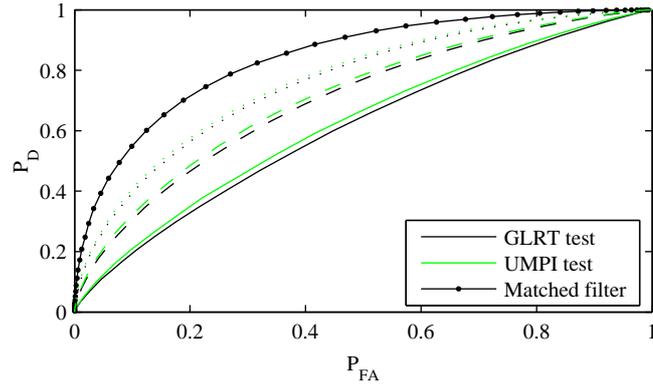
Receiver operating characteristics (ROCs) for the GLRT and UMPI test were obtained for nine cases of interest, namely those of each of the three targets in Figure 2 scaled to have energies of 2, 4, and 8. These ROCs are shown in Figure 4, each plot containing results for one specific energy. The line styles for the curves again relate to those used in Figure 2.

Also shown in the plots in Figure 4 are the ROCs for the ideal matched filter, where the location of the target is known. Since the performance of the matched filter in white noise depends only on the target energy, a single ROC in each plot applies to all of the targets. This matched filter ROC curve represents an absolute upper bound on the performance that any detector can achieve for the problem.

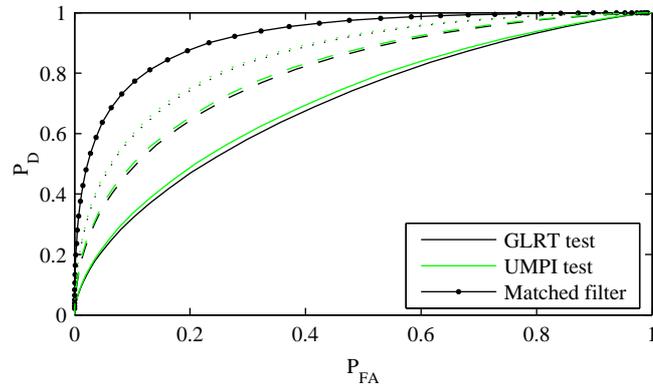
As mentioned in the introduction, in assessing the suboptimality of the GLRT (or any other suboptimal test) for any given problem it is common to compare the ROC for the test to that of the matched filter. If the performance is comparable, then near optimality of the test under analysis is assured. Temporarily ignoring the UMPI results in Figure 4, it may be concluded that the GLRT performance seems to approach optimality as the target to be detected becomes wider (the dotted lines in the plots). In contrast, the performance of the GLRT for narrow targets (the solid lines in the plots) is considerably worse than that of the matched filter. This observation raises the suspicion that the GLRT may be significantly suboptimal in the narrow target case.

The comparison of the GLRT with the matched filter is inconclusive, however, since the two cases assume different unknown quantities. That is, even if the GLRT performance is considerably worse than that of the matched filter, there is no guarantee that this implies suboptimality of the GLRT when the target location is in fact unknown. On the other hand the UMPI test provides an optimal baseline that *is* comparable, since it makes precisely the same assumptions as the GLRT. A comparison of the GLRT to the UMPI test is therefore both justified and meaningful.

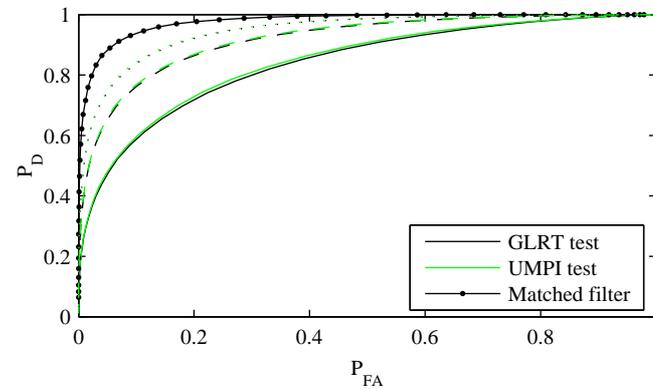
When the GLRT ROC is compared with that of the UMPI test, a very different conclusion emerges: for the unknown target location problem the optimal test performance is highly dependent on the target being detected. The short duration targets are inherently less detectable than the wider targets when their location is unknown, and this causes the difference in performance of the GLRT. Also, since the difference in performance between the GLRT and the UMPI test is marginal, the near optimality of the GLRT for the specific test cases under investigation is assured.



(a) ROC curves for energy 2.



(b) ROC curves for energy 4.



(c) ROC curves for energy 8.

Fig. 4. ROC curves for detection of the three Gaussian test signals scaled to three different energies.

4.3 Comparison for other targets

The difference in performance between the GLRT and the UMPI test is small for Gaussian-shaped targets, which were chosen for their smooth and transient

properties. In this section two other classes of signals are considered: sinusoids of different frequencies, and rectangular pulses with different widths. There are again three signals in each test class, and three scalings of each of these signals are considered.

Figure 5 shows the ROC curves for sinusoidal targets with periods of 16, 8, and 4. The targets are also shown in the figure. The performance of the two

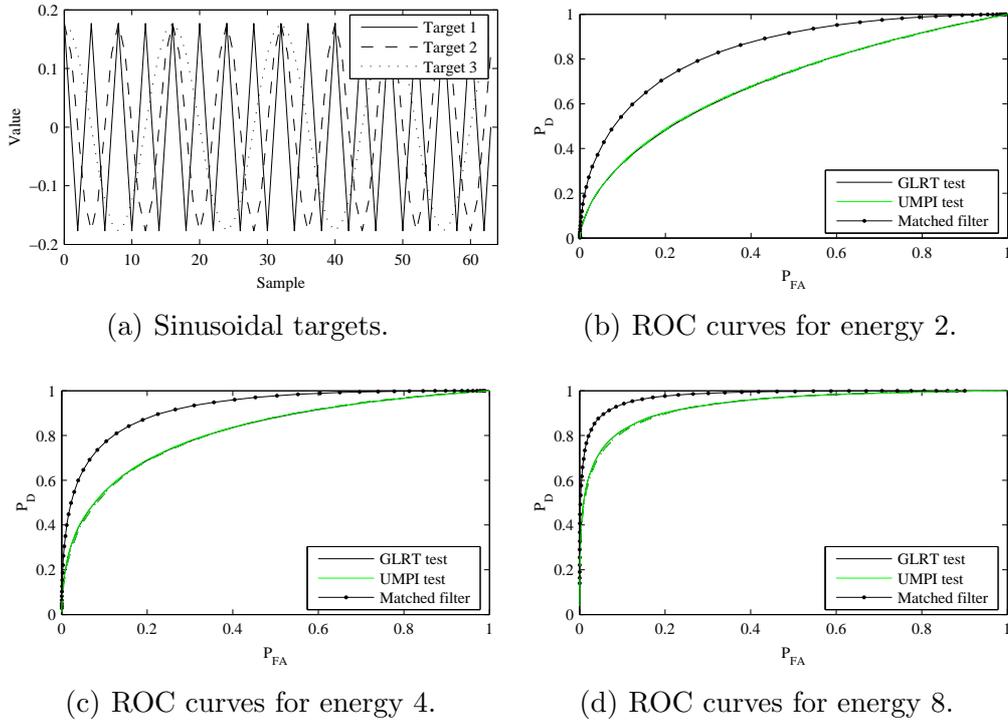


Fig. 5. ROC curves for detection of three sinusoidal test signals scaled to three different energies.

tests are almost identical for this case, again indicating near optimality of the GLRT. Interestingly, the detectability of the signals does not appear to vary over the range of frequencies considered.

Figure 6 shows another set of results, this time for rectangular pulse targets of widths 1, 4, and 8. Comparing these curves with those in Figure 4 it is evident that these targets are less detectable than Gaussian-shaped targets of equivalent energies. Nonetheless, the difference between the performance of the GLRT and UMPI test is again small.

The fact that the differences between the two tests is small in all the cases explored lends support to the near-optimality of the GLRT. The only instance where the UMPI test seems to exhibit slightly better performance is at low signal energies, although the shape of the target also plays a role.

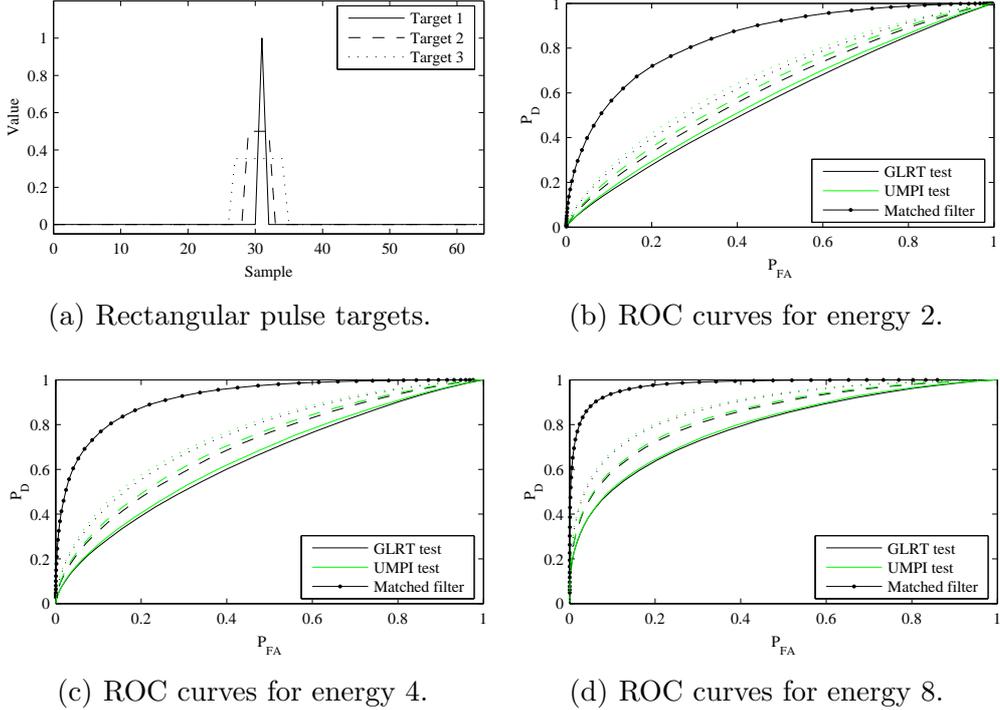


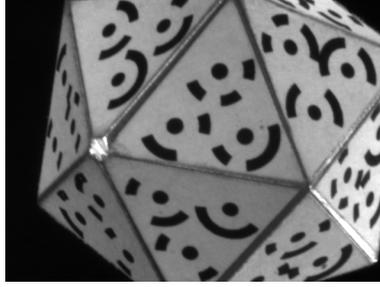
Fig. 6. ROC curves for detection of three rectangular pulse test signals scaled to three different energies.

5 Extensions and applications

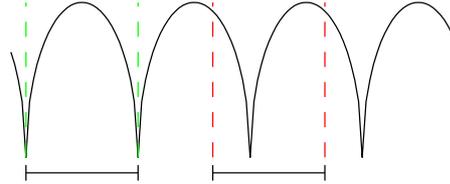
This paper deals with optimality issues related to the detection of targets with unknown location. For this application, forcing invariance to cyclic permutations of the elements of the observation is somewhat contrived, but does provide an indirect means of quantifying optimality of related tests.

In some applications, however, invariance to cyclic permutations of the elements of an observation is *exactly* what is required. For example, Figure 7(a) shows an image of a calibration object used in computer vision — each target marker has a unique circular binary code that needs to be identified [19]. In this case, the task of locating the markers can be performed using generic robust ellipse detectors. To subsequently classify the markers, however, we need to perform hypothesis tests of the observed data against the known binary code configurations, subject to an unknown overall rotation of the marker. To do this, we can sample the pixel data around the located ellipse, stack the sample values into an observation vector, and use this vector in a subsequent hypothesis test. In this case invariance to cyclic permutations corresponds to invariance to rotation of the binary code around the marker center.

In a signal processing example, the UMPI test can be used to optimally differentiate between observations from two periodic signals that repeat over a



(a) Calibration object with binary coded markers.



(b) Two one-period samples from a periodic signal.

Fig. 7. Applications where invariance to cyclic permutation is appropriate.

common and known period. If the observation sample is chosen to consist of an integral number k of fundamental periods, then the resulting signal will contain k cycles of the signal. A shift in the sampling origin then just results in a cyclic permutation of the underlying signal samples (see Figure 7(b)). Invariance to cyclic permutations is therefore a natural condition to impose on tests for this problem, and the UMPI test will consequently be optimal. Note that the optimality of this approach does depend on the signal being sampled at a sufficiently high frequency to avoid aliasing.

In general, invariance to cyclic permutation may be relevant when signals are cyclic in their nature, or exhibit some kind of circular symmetry. Finite-duration samples of cyclostationary random processes fall into this category. These types of signals occur quite often in practical applications.

6 Conclusions

A uniformly most powerful invariant test has been derived for the problem of detecting a known target with unknown cyclic permutation in Gaussian noise with a circulant covariance matrix. The set of permissible test statistics are constrained to be those which are invariant to cyclic permutations of the data. The most powerful test within this class is found, which performs better than any other cyclic permutation invariant test statistic for *all* target locations. References are provided that argue that the distribution of the test statistic

is approximately normal, and outline analytical methods for finding the first two moments.

The UMPI test obtained for the problem is compared with the more conventional solution, namely the GLRT, and it is shown for a number of targets that the differences between the two tests is negligible. In fact, quite extensive investigation has not yielded any cases where the UMPI test significantly outperforms the GLRT. This asserts the near optimality of the GLRT for certain unknown target location problems. The comparison is more meaningful than a direct comparison of the GLRT to the ideal matched filter, which assumes the location parameter known.

The comparison between the UMPI test and the GLRT is strictly only valid when shifts are defined circularly and targets wrap around the edges of the observation interval. However, the comparison should also be approximately valid for linear shifts as long as the edge effects are not too severe. One such example is when the target is short and the observation interval long. In this case, comparing the GLRT and UMPI test performance for circular shifts should provide a good estimate of the suboptimality of the GLRT for linear shifts.

It has been observed by the authors that optimal tests are often derived for the case of known target location, and that the tests are then applied in a sliding window framework in a system that operates continuously through time. The resulting detector has no claim to optimality, however, since the conditions under which it is operating are not the same as those for which it was designed. The methods presented in this paper provide some justification for this ad hoc approach, at least for one simple class of problems.

References

- [1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory (Part I)*, 1968.
- [2] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996.
- [3] M. Kendall, A. Stuart, *The Advanced Theory of Statistics*, 4th Edition, Vol. 2, Charles Griffin & Company, 1979.
- [4] G. B. Giannakis, Cyclostationary signal analysis, in: V. K. Madisetti, D. B. Williams (Eds.), *The Digital Signal Processing Handbook*, CRC Press, 1998, Ch. 17, pp. 17-1-17-31.
- [5] L. L. Scharf, B. Friedlander, Matched subspace detectors 42 (8) (1994) 2146-2157.

- [6] R. T. Behrens, L. L. Scharf, Signal processing applications of oblique projection operators 42 (6) (1994) 1412–1424.
- [7] L. L. Scharf, D. W. Lytle, Signal detection in gaussian noise of unknown level: an invariance application IT-17 (4) (1971) 404–411.
- [8] S. Bose, A. O. Steinhardt, A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices 43 (9) (1995) 2164–2175.
- [9] L. L. Scharf, Statistical Signal Processing: Detection, Estimation, and Time Series Analysis, Addison-Wesley Series in Electrical and Computer Engineering: Digital Signal Processing, 1991.
- [10] E. L. Lehmann, Testing Statistical Hypotheses, 2nd Edition, Wiley Series in Probability and Mathematical Statistics, 1986.
- [11] R. V. Hogg, A. T. Craig, Introduction to Mathematical Statistics, 3rd Edition, 1970.
- [12] P. J. Davis, Circulant Matrices, Pure & Applied Mathematics: A Wiley-Interscience Series of Texts, Monographs & Tracts, 1979.
- [13] A. A. Abu-Dayya, N. C. Beaulieu, Outage probabilities in the presence of correlated lognormal interferers 43 (1) (1994) 164–173.
- [14] W. A. Janos, Tail of the distribution of sums of log-normal variates IT-16 (3) (1970) 299–302.
- [15] S. C. Schwartz, Y. S. Yeh, On the distribution function and moments of power sums with log-normal components 61 (7) (1982) 1441–1462.
- [16] A. Safak, Statistical analysis of the power sum of multiple correlated log-normal components 42 (1) (1993) 58–61.
- [17] B. Afonja, The moments of the maximum of correlated normal and t-variates 34 (1972) 251–262.
- [18] G. M. Tallis, The moment generating function of the truncated multi-normal distribution 23 (1961) 223–229.
- [19] K. Forbes, A. Voigt, N. Bodika, An inexpensive, automatic and accurate camera calibration method, in: Proceedings of the Thirteenth Annual Symposium of the Pattern Recognition Association of South Africa (PRASA 2002, 2002).