Comparison of GLR and Invariance Methods Applied to Adaptive Target Detection in Structured Clutter

Hyung Soo Kim and Alfred O. Hero

COMMUNICATIONS & SIGNAL PROCESSING LABORATORY
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109–2122

November, 2000

Technical Report No. 323

Comparison of GLR and Invariance Methods Applied to Adaptive Target Detection in Structured Clutter

Hyung Soo Kim and Alfred O. Hero

Abstract

This paper addresses a target detection problem for which the covariance matrix of the unknown Gaussian clutter background has block diagonal structure. This block diagonal structure is the consequence of the target lying along a boundary between two statistically independent clutter regions. Here, we design adaptive detection algorithms using both the generalized likelihood ratio (GLR) and invariance principles. By exploiting the known covariance structure a set of maximal invariants is obtained. These maximal invariants are a compression of the image data which retain target information while being invariant to clutter parameters. We consider three different assumptions on knowledge of the clutter covariance structure: both clutter types totally unknown, one of the clutter types known except for its variance, and one of the clutter types completely known. By means of simulation, the GLR and maximal invariant (MI) tests are shown to outperform the previously proposed invariant test by Bose and Steinhardt which is derived from a similarly structured covariance matrix. Numerical comparisons are presented which illustrate that the GLR and MI tests are complementary, i.e. neither test strategy uniformly outperforms the other over all values of SNR, number of chips, and false alarm rate. This suggests that it may be worthwhile to combine these two tests into a hybrid test to obtain overall optimal performance.

I. Introduction

In this paper, adaptive detection algorithms are developed for imaging radar (IR) targets in *structured* clutter by exploiting both the generalized likelihood ratio (GLR) principle and the invariance principle. In automatic target recognition (ATR), it is important to be able to reliably detect or classify a target in a manner which is robust to target and clutter variability yet maintains the highest possible discrimination capability. The GLR and invariance principles are worthwhile approaches since they often yield good constant false alarm rate (CFAR) tests. The GLR principle implements the intuitive estimate-and-plug principle: replacing all unknowns in the likelihood ratio (LR) test by their maximum likelihood estimates (MLEs). The GLR is known to be asymptotically optimal, i.e. GLR becomes uniformly most powerful (UMP) in that it attains the highest probability of correct detection for given probability of false alarm (P_{FA}), as the number of observations become much larger than the number of unknown clutter parameters [1]. In contrast, application of the invariance principle seeks to project away the clutter parameters by compressing the observations down to a lower dimensional statistic while retaining the maximal amount

This work was supported in part by the Air Force Office of Scientific Research under MURI grant: F49620-97-0028. A preliminary version of this work appeared in "Adaptive target detection across a clutter boundary: GLR and maximally invariant detectors," *Proceedings of IEEE International Conference on Image Processing*, Vancouver, Canada, Sep. 2000.

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 (Email: kimhs,hero@eecs.umich.edu).

of information for discrimination of the target [2], [3], [4], [5]. This statistic is called the maximal invariant and, if one is lucky, the form of the most powerful (MP) LR test based on the maximal invariant does not depend on the nuisance parameters, resulting in a uniformly most powerful invariant (UMPI) test [6], [7]. When properly applied, the invariance principle can yield adaptive target detection algorithms which outperform the GLR test, sometimes by a significant margin. As we will show in this paper, for the problem of target detection in unknown but structured clutter background, this margin of improvement can be quite significant at low signal-to-noise ratio (SNR) for small fixed P_{FA}.

A common assumption in homogeneous but uncertain clutter scenarios is that the target is of known form but unknown amplitude in Gaussian noise whose covariance matrix is totally unknown or *unstructured*. This assumption induces parameter uncertainty for which the general multivariate analysis of variance (GMANOVA) model applies and optimal and suboptimal detection algorithms can be easily derived using the GLR principle [8], [9], [10].

When some structure on the covariance matrix is known a priori, improvements over this GLR test are possible. For example, in the context of antenna arrays for detection of an impinging wavefront in clutter whose covariance matrix has Toeplitz structure, Fuhrmann [11] showed through simulation that a significant improvement over the GLR test [9] is achieved by incorporating a Toeplitz constraint into the covariance estimation step in the GLR detector. As previously mentioned, an alternative approach to estimation of the partially known noise covariance is to project the data down to a test statistic whose noise-alone distribution does not depend on this covariance. Bose and Steinhardt [12] proposed such an invariant detector and showed that it outperforms the GLR for unstructured covariance when the noise covariance matrix is assumed to have a priori known block diagonal structure. In [13], the difficult deep hide scenario was considered where the target parks along a known boundary separating two adjacent clutter regions, e.g. an agricultural field and a forest canopy. It was shown there that under the reasonable assumption that the two clutter types are statistically independent, the induced block diagonal covariance structure can be used to derive an invariant test with performance advantage similar to Bose and Steinhardt's test.

In this paper, we derive the form of the GLR for block structured covariance. Then the invariant approach considered in [12] and [13] is developed in the context of imaging radar for deep hide targets, and compared to the GLR. In this context, the spatial component has clutter covariance matrix $\mathbf{R}$, which decomposes into a block diagonal matrix under an independence assumption between the two clutter regions, and the target is assumed to come from a known set of signatures of unknown amplitudes and orientations. Several cases, denoted in decreasing order of uncertainty as Cases 1, 2 and 3, of block

diagonal covariance matrices are examined:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_A & \mathbf{O} \\ \mathbf{O} & \mathbf{R}_B \end{bmatrix} \tag{1}$$

- Case 1: $\mathbf{R}_A > 0$, $\mathbf{R}_B > 0$
- Case 2: $\mathbf{R}_A > 0$, $\mathbf{R}_B = \sigma^2 \mathbf{I}$ where $\sigma^2 > 0$
- Case 3: $\mathbf{R}_A > 0$, $\mathbf{R}_B = \mathbf{I}$

where the subscripts denote the two different regions A and B. Case 1 corresponds to two completely unknown clutter covariance matrices $\mathbf{R}_A$ and $\mathbf{R}_B$, and Case 2 corresponds to one clutter covariance $\mathbf{R}_A$ completely unknown and the other $\mathbf{R}_B$ known up to a scale parameter. As shown in [12] the known clutter covariance matrix in $\mathbf{R}_B$, represented by the matrix $\mathbf{I}$, can be taken as the identity matrix without loss of generality. Case 3 corresponds to $\mathbf{R}_B$ known exactly. Cases 2 and 3 arise, for example, in application where one of the clutter regions is well characterized. For real valued observations, the GLR method is shown to have explicit form for each of Cases 1, 2 and 3, involving the roots of a 4th order algebraic equation. For complex valued observations, 4th order algebraic equations for real and imaginary parts of the target amplitude must be solved numerically. The maximal invariant statistics for Cases 1 and 2 were previously derived by Bose and Steinhardt and invariant tests were proposed based on these statistics in [12]. We treat Cases 1-3 in a unified framework and propose alternative maximal invariant (MI) tests which are better adapted to the deep hide target application. We show via simulation that there are regimes of operation which separate the performance of the GLR and MI tests. When there are a large number of independent snapshots of the clutter, the MLEs of the target amplitude and the block diagonal clutter covariance are reliable and accurate, and the GLR test performs as well as the MI test. Conversely, when a limited number of snapshots are available and SNR is low, the MLEs are unreliable and the MI test outperforms the GLR test. This property is also confirmed by the real data example, i.e. the MI test can detect weaker targets than the other tests when the number of snapshots is few.

In Section II, the image model for the detection problem is introduced and a canonical form is obtained by coordinate transformation. We then review the principles of GLR and invariance in Section III. Kelly's GLR test [9] for an unstructured covariance matrix is derived as an illustration of these two principles. Section IV then reviews the application of these principles to detect a target across a clutter boundary. There we also extend the detection problem from a single target to one of multiple targets. Finally, the relations between the GLR and MI tests are explored by analysis and by simulation.

II. Image Model

A. Radar Imaging

Detection of targets in radar images is a multi-stage process involving pre-processing, image formation from raw data, and formation of a test statistic. In this section, we review some issues regarding imaging radars and their image outputs to which we can apply detection algorithms. Further information on the image formation problem for radar can be found in remote sensing books such as [14], [15].

Radiometric sensors are usually divided into two groups according to their modes of operation: passive sensors or radiometers, and active sensors such as imaging or non-imaging radar. Most imaging radars used for remote sensing are divided into two groups: the real-aperture systems that depend on the beamwidth determined by the baseline of a fixed antenna, and the synthetic-aperture systems that form a virtual baseline with a moving antenna. Synthetic-aperture radar (SAR) systems can acquire fine resolution using a small antenna with spatial resolution independent of the radial distance to target, which has made spaceborne imaging radar with fine resolution feasible [14]. Although we restrict our attention to detection problems for radar images, the same techniques can be applied to passive sensors such as long-wave infrared or thermal radiometers.

In active imaging radars, the returned signals are processed to extract complex images of target reflectivity which consist of magnitude and phase information. Fig. 10 displays the magnitude of such a complex valued SAR image which has been converted into decibels (dB). It is common to model a complex valued radar image as linear in the target with additive Gaussian distributed clutter. Examples where a Gaussian model is justified for terrain clutter can be found in [16]. Even in cases when such a model is not applicable to the raw data, a whitening and local averaging technique can be implemented to obtain a Gaussian approximation, e.g. by subtracting the local mean from the original image while minimizing the third moment of the residual image [17], [18].

Assume that the complex image has been scanned or reshaped to a column vector $\underline{x}$. If multiple snapshots (chips) $\underline{x}_1, \ldots, \underline{x}_n$ of the terrain are available, they can be concatenated into a spatio-temporal matrix $\mathbf{X}$ with columns $\{\underline{x}_i\}_{i=1}^n$. Let $\underline{s}$ be the reshaped target vector to be detected in a clutter background $\mathbf{N}$ having independent, identically distributed (i.i.d.) Gaussian columns with zero mean. Then we have the simple image model

$$\mathbf{X} = a\, \underline{s}\, \underline{b}^H + \mathbf{N}$$

where a is an unknown target amplitude and $\underline{b}^H$ accounts for the articulation of the target vector into the snapshot sequence, e.g. possible chip locations of the target. In spatially scanned radar images, the vector $\underline{b}^H$ would be equal to $[1, 0, \ldots, 0]$ if the target presence is to be detected in the first image chip

(1st column of $\mathbf{X}$). In this case, this column will be called primary data while the rest of $\mathbf{X}$ will be called secondary data. In multi-spectral images, described in [15] and studied by Reed [17], [18], $a\underline{s}$ can be thought of as a vector of unknown spectral intensities and $\underline{b}^H$ represents the known spatial signature. We will consider a more general p target model in the next subsection.

The most common assumption for clutter is that its spatial covariance matrix is completely unknown. This assumption makes derivation of the GLR decision rule easy and leads to a CFAR test for which the false alarm rate is independent of the actual covariance of the clutter [9], [19]. However, when available, inclusion of side information about the noise clutter covariance will give improved performance, even though the derivation of the GLR is often rendered more difficult.

B. Problem Formulation

Let $\{\underline{x}_i\}_{i=1}^n$ be n statistically independent $m \times 1$ complex Gaussian vectors constructed by raster scanning a set of n 2-D images. We call each of these vectors subimages or chips and assume that they each have identical $m \times m$ clutter covariance matrices $\mathbf{R}$ but with possibly different mean vectors (targets). Concrete examples are: multi-spectral images where subimages correspond to a scene at n different optical or radar wavelengths; multiple pulse SAR images where repeated probing of a scene produces a sequence of n subimages; polarized L or C band SAR where subimages correspond to $m = 3$ polarization components (HH, HV, VV) at n different spatial locations; or n non-contiguous spatial cells of a single IR/SAR image.

As explained above, by concatenating these n vectors we obtain the measurement image matrix $(m \times n)$ $\mathbf{X} = [\underline{x}_1, \ldots, \underline{x}_n]$. This matrix can be modeled as follows

$$\mathbf{X} = \mathbf{S}\,\underline{a}\,\underline{b}^H + \mathbf{N} \tag{2}$$

where $\mathbf{S} = [\underline{s}_1, \ldots, \underline{s}_p]$ is an $m \times p$ matrix consisting of signature vectors of p possible targets, $\underline{a} = [a_1, \ldots, a_p]^T$ is a $p \times 1$ unknown target amplitude vector for p targets, and $\underline{b}^H = [b_1, \ldots, b_n]$ is a $1 \times n$ target location vector which accounts for the presence of target(s) in each subimage. Also we assume that $\mathbf{N}$ is a complex multivariate Gaussian matrix with i.i.d. columns: $vec\{\mathbf{N}\} \sim \mathcal{CN}(\underline{0}, \mathbf{R} \bigotimes \mathbf{I}_n)$ where $\underline{0}$ is an $mn \times 1$ zero vector, $\mathbf{I}_n$ is an $n \times n$ identity matrix, and $\bigotimes$ is the Kronecker product. This model is a simplified form of the GMANOVA model

$$\mathbf{X} = \mathbf{S}\,\mathbf{A}\,\mathbf{B} + \mathbf{N}$$

where $\mathbf{S}\,(m \times p)$ of rank p and $\mathbf{B}\,(p \times n)$ of rank p are known matrices, and $\mathbf{A}\,(p \times p)$ is a matrix of unknown amplitudes [20]. In this paper, the simpler form (2) will be used throughout and correspond to the assumption that the target location vectors (rows of $\mathbf{B}$) for p targets are all identical, i.e. the target components in different subimages differ only by a scale factor.

The detection problem is to seek the presence of target(s) for $\mathbf{S}$ and $\underline{b}$ known, $\underline{a}$ unknown, and the independent columns of $\mathbf{N}$ having the unknown covariance matrix $\mathbf{R}$. By applying coordinate rotations to both of the column space and the row space of $\mathbf{X}$ we can put the image model into a convenient canonical form as in [19]. Let $\mathbf{S}$ and $\underline{b}$ have the QR decompositions

$$\mathbf{S} = \mathbf{Q}_S \begin{bmatrix} \mathbf{T}_S \\ \mathbf{O} \end{bmatrix}, \ \underline{b} = \mathbf{Q}_b \begin{bmatrix} t_b \\ \underline{0} \end{bmatrix}$$

where $\mathbf{Q}_S(m \times m)$, $\mathbf{Q}_b(n \times n)$ are unitary matrices, $\mathbf{T}_S$ is a $p \times p$ upper-triangular matrix, and t_b is a scalar. Multiplying $\mathbf{X}$ on the left and right by $\mathbf{Q}_S^H$ and $\mathbf{Q}_b$, respectively, we have the canonical representation

$$\tilde{\mathbf{X}} = \mathbf{Q}_S^H \, \mathbf{X} \, \mathbf{Q}_b = \begin{bmatrix} \mathbf{T}_S \\ \mathbf{O} \end{bmatrix} \underline{a} \, \begin{bmatrix} t_b^H & \underline{0}^H \end{bmatrix} + \tilde{\mathbf{N}}$$

where $\tilde{\mathbf{N}}$ is still n-variate normal with $vec\{\tilde{\mathbf{N}}\} \sim \mathcal{CN}(\underline{0}, (\mathbf{Q}_S^H \, \mathbf{R} \, \mathbf{Q}_S) \bigotimes \mathbf{I}_n)$ and the target detection problem is not altered since $\mathbf{R}$ is unknown. Now the transformed data has the partition

$$\tilde{\mathbf{X}} = \begin{bmatrix} \underline{x}_{11} & \mathbf{X}_{12} \\ \underline{x}_{21} & \mathbf{X}_{22} \end{bmatrix}$$

where $\underline{x}_{11}$ is a $p \times 1$ vector, $\underline{x}_{21}$ is a $(m-p) \times 1$ vector, $\mathbf{X}_{12}$ is $p \times (n-1)$, and $\mathbf{X}_{22}$ is $(m-p) \times (n-1)$. Note that $\mathbf{Q}_S^H$ and $\mathbf{Q}_b$ have put all the target energy into the first p pixels of the first subimage, $\underline{x}_{11}$. This canonical form is identical to the one found in [21]. In the sequel, unless stated otherwise, we will assume that the model has been put into this canonical form.

For the special case of $p = 1$ (single target), this model reduces to the one studied by Kelly [9]

$$\mathbf{X} = a \, \underline{\varepsilon}_1 \, \underline{e}_1^T + \mathbf{N} \tag{3}$$

where a is an unknown complex amplitude, $\underline{e}_1 = [1, 0, \ldots, 0]^T$ is the $n \times 1$ unit vector, and the known target signature is transformed into an $m \times 1$ unit vector $\underline{\varepsilon}_1$. With the model (3) we can denote the unknowns by the unknown parameter vector $\theta = \{a, \mathbf{R}\} \in \Theta$ where Θ is the prior parameter range of uncertainty. Let Θ_0 and Θ_1 partition the parameter space into target absent (H_0) and target present (H_1) scenarios: $\Theta_0 = \{a, \mathbf{R} : a = 0, \mathbf{R} \in \text{Hermitian}(m \times m)\}$, $\Theta_1 = \{a, \mathbf{R} : a \neq 0, \mathbf{R} \in \text{Hermitian}(m \times m)\}$. Then the general form for the detection problem is expressed via the two mutually exclusive hypotheses:

$$\begin{aligned} H_0 \quad &: \quad \mathbf{X} \sim f(\mathbf{X}; \theta_0), \ \theta_0 = \{0, \mathbf{R}\} \in \Theta_0 \\ H_1 \quad &: \quad \mathbf{X} \sim f(\mathbf{X}; \theta_1), \ \theta_1 = \{a, \mathbf{R}\} \in \Theta_1. \end{aligned}$$

Now, following [12], we extend (3) to the structured covariance case. Consider Case 1 in Section I. This is the scenario where a target parks along a known boundary of the two unknown but independent clutter

regions A and B. Then the target signature $\underline{s}$ is partitioned as

$$\underline{s} = \begin{bmatrix} \underline{s}_A(m_A \times 1) \\ \underline{s}_B(m_B \times 1) \end{bmatrix}$$

where $m_A + m_B = m$, and the unitary matrices $\mathbf{Q}_{S_A}$ and $\mathbf{Q}_{S_B}$ can be obtained from the QR decompositions of $\underline{s}_A$ and $\underline{s}_B$, respectively. Using

$$\mathbf{Q}_S = \begin{bmatrix} \mathbf{Q}_{S_A} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q}_{S_B} \end{bmatrix}$$

in the canonical transformation,the model is then composed of two parts from regions A and B

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{bmatrix} = a \begin{bmatrix} \tilde{\underline{s}}_A \\ \tilde{\underline{s}}_B \end{bmatrix} \underline{e}_1^T + \begin{bmatrix} \mathbf{N}_A \\ \mathbf{N}_B \end{bmatrix} \tag{4}$$

where $\tilde{\underline{s}}_A = \left[s_A^H, 0, \ldots, 0\right]^H$ and $\tilde{\underline{s}}_B = \left[s_B^H, 0, \ldots, 0\right]^H$. $\mathbf{N}_A$ $(m_A \times n)$ and $\mathbf{N}_B$ $(m_B \times n)$ are independent Gaussian matrices with unknown covariance matrices $\mathbf{R}_A$ $(m_A \times m_A)$ and $\mathbf{R}_B$ $(m_B \times m_B)$, respectively. The target detection problem is now simply stated as testing $a = 0$ vs. $a \neq 0$ in (4).

III. Detection Theory

A. Hypothesis Testing

Signal detection is a special case of hypothesis testing theory in statistical inference [1]. As explained in the previous section, given an observation $\mathbf{X}$ a decision has to be made between two hypotheses corresponding to presence of a target (H_1) or no target (H_0), respectively. Hypotheses are divided into two classes according to the parameter θ underlying probability density functions (pdfs) of each of the hypotheses. When θ can take on only one value, fixed and known, under each hypothesis, the hypotheses are said to be simple. In this case, the pdf $f(\mathbf{X};\theta)$ is known given H_0 or H_1. Otherwise, hypotheses are said to be composite. In singly composite hypotheses, only one of Θ_0 or Θ_1 contains a set of values of θ, and in doubly composite hypotheses, both Θ_0 and Θ_1 contain a set of values of θ. Thus composite hypotheses only specify a family of pdfs for $\mathbf{X}$. There are two general strategies for deciding between H_0 and H_1: the Bayesian strategy and the frequentist strategy. In Bayes approach to detection, priors are assigned to the probabilities of H_0 and H_1, and it is assumed that θ is a random variable or vector with a known pdf $f(\theta)$. After assigning a set of costs to incorrect decisions, the Bayes objective is to find and implement a decision rule which has minimum average cost or risk. In many situations, however, it is difficult to assign these priors. Moreover, the Bayes approach assures good performance only for the average parameter values over Θ_0 and Θ_1. The Bayes approach does not control the performance of a test for any specific parameter value which can arise. Also it provides no guaranteed protection against false alarm (deciding H_1 when H_0 is true) and miss (deciding H_0 when H_1 is true). For simple hypotheses,

$\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta_1\}$, the frequentist approach attempts to maximize the conditional probability of detection (P_D) for fixed conditional probability of false alarm (P_{FA}), or vice versa, resulting in the Neyman-Pearson test. Specifically, a maximum tolerable level α of false alarm is specified as

$$P_{FA}(\theta_0) \leq \alpha$$

where $P_{FA}(\theta) = P(\text{decide} H_1 | H_0; \theta)$ and a decision rule is selected which attains the largest P_D among all tests of level α. This test is also called the MP test [22].

In testing simple hypotheses, both Bayes and Neyman-Pearson criteria lead to the same decision rule involving the LR test. The only difference lies in the selection of the thresholds applied to the test statistic. However, in a general problem of testing composite hypotheses the two approaches differ substantially. For non-random θ, if such a test existed, a frequentist would select a test of level α satisfying

$$\max_{\theta \in \Theta_0} P_{FA}(\theta) \leq \alpha$$

which has the highest $P_D(\theta)$ among all tests of level α and for all $\theta \in \Theta_1$. Such a test is called a UMP test. Whenever a UMP test exists, it works as well as if we knew θ. Usually, however, such a test does not exist since the optimal decision region of an LR test depends on the unknown parameters $\{\theta_0, \theta_1\}$, and consequently we need to adopt other alternative strategies. For instance, in testing doubly composite hypotheses, most detectors will have their P_{FA} and P_D varying as functions of unknown $\theta \in \Theta_0$ and $\theta \in \Theta_1$, respectively. In such cases, two classes of strategies can be used: one is to optimize alternative decision criterion and the other is to constrain the form of detectors to a class for which a UMP test may exist. Several methods utilizing one of these strategies are listed below. Some examples of alternative criterion are the minmax test and the locally most powerful (LMP) test. A minmax test is a test of level α which maximizes the worst case power $\min_{\theta \in \Theta_1} P_D(\theta)$. In order to implement this test, we need to find the least favorable density which maximizes P_D while constraining the specific level of P_{FA}. However, the performance of a minmax test can be overly conservative especially if least favorable priors concentrate on atypical values of θ. Furthermore, the least favorable priors may be difficult to find in practice. The main idea behind a LMP test is to find the MP test for detecting a small perturbation of parameters from H_0. The LMP test is particularly useful for weak signals. Some examples of constrained classes of tests are unbiased tests, CFAR tests, and invariant tests. Unbiased tests are all tests of level α whose $P_D(\theta)$ is greater than α for all $\theta \in \Theta_1$. CFAR refers to the property that the tests have constant false alarm probability over Θ_0. Sometimes UMP CFAR tests exist when UMP tests do not exist. Finally, invariant tests seek to find a transformation or compression of the data, which results in reducing the effect of nuisance parameters. In many cases, the invariant and unbiased classes of tests contain the minmax optimal test.

B. GLR Principle

Generally, optimality is not guaranteed in a composite hypothesis test which involves highly variable nuisance parameters, and a UMP test rarely exists. When such a test does not exist, a popular sub-optimal strategy is to use the GLR principle. With unknown parameters θ in the likelihood ratio, a logical procedure is to find good estimates of θ under H_0 and H_1, and substitute these estimates into the likelihood ratio statistic as if they were true. This is akin to reformulating H_0 and H_1 as simple hypotheses depending on the estimated θ values for which an MP test always exists. In GLR procedures, the pdfs of the measurement under both hypotheses are maximized separately over all unknown parameters by replacing them with their MLEs. As a function of MLEs, this GLR test is asymptotically UMP since, under broad conditions [23], MLEs are consistent estimators as the number of observations goes to infinity [1]. And in many physical problems of interest, either a UMP test will exist or a GLR test will give satisfactory results. In some instances, however, the optimization or maximization involved in deriving a GLR test may be intractable to obtain in closed form. Moreover, similarly to MLEs, the performance of a GLR test can be poor (not even unbiased) in a finite sample regime.

C. Invariance Principle

The main idea behind the invariance principle is to find a statistic which maximally condenses the data while retaining the discrimination capacity of the original data set. It is instructive to first consider the mechanism of data reduction associated with the minimal sufficient statistic. Recall that a function $\mathbf{T} = T(\mathbf{X})$ of the data is a sufficient statistic for testing between H_0 and H_1 if, for all $\theta_0 \in \Theta_0$ and $\theta_1 \in \Theta_1$, the likelihood ratio Λ depends on $\mathbf{X}$ only through $T(\mathbf{X})$:

$$\Lambda(\mathbf{X}; \theta_0, \theta_1) \triangleq \frac{f(\mathbf{X}; \theta_1)}{f(\mathbf{X}; \theta_0)} = g(T(\mathbf{X}); \theta_0, \theta_1).$$

The sets $\{\mathbf{X} : T(\mathbf{X}) = t\}_t$ can be thought of as sufficiency orbits of $\mathbf{X}$ which specify constant contours of $\Lambda(\mathbf{X})$. Thus a sufficient statistic $T(\mathbf{X})$ indexes the orbits and preserves all information needed to discriminate between H_0 and H_1. Sufficient statistics are not unique and $T(\mathbf{X})$ is a minimal sufficient statistic if it is a function, i.e. a compression, of any other sufficient statistic.

Data reduction via invariance is achieved by finding a statistic $\mathbf{Z} = Z(\mathbf{X})$, called the maximal invariant statistic, which indexes the set values (which we can think of as constant contours) of the set function

$$\tilde{\Lambda}(\mathbf{X}) \triangleq \{\Lambda(\mathbf{X}; \theta_0, \theta_1) : \theta_0 \in \Theta_0, \theta_1 \in \Theta_1\}. \tag{5}$$

To make this practical, a tractable mathematical characterization of this set function must be adopted. This can be accomplished when the probability model has functional invariance which can be characterized by group actions on the measurement space $\mathbf{X}$ and induced group actions on the parameter space Θ. Let

$\mathcal{G}$ be a group of transformations $g : \chi \to \chi$ acting on $\mathbf{X}$. Assume that for each $\theta \in \Theta$ there exists a unique $\bar{\theta} = \bar{g}(\theta)$ such that $f_\theta(g(\mathbf{X})) = f_{\bar{\theta}}(\mathbf{X})$. $\bar{g} \in \bar{\mathcal{G}}$ is called the induced group action on Θ. The above relation implies that the natural invariance which exists in the parameter space of θ implies a natural invariance in the space of measurement $\mathbf{X}$. If we further assume that $\bar{g}(\Theta_0) = \Theta_0$, $\bar{g}(\Theta_1) = \Theta_1$, then the model and the decision problem are said to be invariant to the group $\mathcal{G}$. The orbits of $\mathbf{X}$ under actions of $\mathcal{G}$ are defined by

$$\mathbf{X} \equiv \mathbf{Y} \text{ if } \exists\, g \in \mathcal{G} \text{ such that } \mathbf{Y} = g(\mathbf{X}).$$

The orbits of θ under actions of $\bar{\mathcal{G}}$ are similarly defined. Note that to capture natural invariance of the model, the groups $\mathcal{G}$ and $\bar{\mathcal{G}}$ must have group actions with the largest possible degrees of freedom among all groups leaving the decision problem invariant.

Example: Sufficiency vs. Invariance

To illustrate the statistical reduction or data compression associated with sufficiency and invariance, consider the signal model defined in (3) for real valued measurements. For the special case of $m = 1$, this model reduces to

$$\underline{x} = a\, \underline{e}_1^T + \underline{N}^T$$

where, as before, $\underline{e}_1$ is an $n \times 1$ unit vector and $\underline{N}^T$ is a normal row vector with zero mean and covariance matrix $\sigma^2\mathbf{I}$. The above model corresponds to testing for target presence at a single pixel in a sequence of n snapshots. This is a simple i.i.d. real Gaussian example with unknown parameters $\theta = \{a, \sigma^2\}$, and the pdf of $\underline{x}$ is

$$f(\underline{x}) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{1}{2}\left\{\frac{(x_1 - a)^2}{\sigma^2} + \sum_{i=2}^{n} \frac{x_i^2}{\sigma^2}\right\}\right] \tag{6}$$

where $\underline{x} = [x_1, x_2, \ldots, x_n]$. Since the objective is to decide whether $a = 0$ or $a \neq 0$ when σ^2 is unknown, we define $\theta_0 = \{0, \sigma_0^2\}$ and $\theta_1 = \{a, \sigma_1^2\}$ which are points in $\Theta_0 = \{a, \sigma^2 : a = 0, \sigma^2 > 0\}$ and $\Theta_1 = \{a, \sigma^2 : a \neq 0, \sigma^2 > 0\}$, respectively. The likelihood ratio is

$$\Lambda(\underline{x}; \theta_0, \theta_1) = \frac{f(\underline{x}; a, \sigma_1^2)}{f(\underline{x}; 0, \sigma_0^2)}.$$

With the pdf (6), we can express the log likelihood ratio as a function of $\underline{x}$:

$$\ln \Lambda(\underline{x}; \theta_0, \theta_1) = \frac{a}{\sigma_1^2} \cdot \underline{e}_1^T \underline{x} + \frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2} \cdot \underline{x}^T \underline{x} - \frac{a^2}{2\sigma_1^2} + n \ln \frac{\sigma_0}{\sigma_1}$$

where $a \in \mathbb{R}$ and $\sigma_0^2, \sigma_1^2 > 0$. Thus Λ depends on $\underline{x}$ only through $T(\underline{x}) = \{t_1, t_2\}$ where $t_1 = \underline{e}_1^T \underline{x}$ and $t_2 = \underline{x}^T \underline{x}$. $T(\mathbf{X})$ is a minimal sufficient statistic indexing the sufficiency orbit illustrated in Fig. 1 as a circle in $\mathbb{R}^3$ when $n = 3$. Given t_1 and t_2 we can recover all of the information in the entire n-sample

required to discriminate between different values of the parameter pair $\{a, \sigma^2\}$. Generally for $n > 2$, the sufficiency orbit of $\underline{x}$ is the surface of an $n-1$ dimensional hypersphere defined by the intersection of the surfaces of the hypersphere $\{\underline{x} : \sum_{i=1}^n x_i^2 = t_2\}$ and the hyperplane $\{\underline{x} : x_1 = t_1\}$.

This sufficient statistic is the minimum amount of information required of $\underline{x}$ to estimate the parameters $\{a, \sigma^2\}$. The maximal invariant is the minimum amount of information required to discriminate between the sets of parameter values Θ_0 and Θ_1, i.e. detection of the target. To determine the maximal invariant, the set function (5) of the likelihood ratio needs to be defined first. Since, under H_1, $a \in \mathbb{R}$ and $\sigma_0^2, \sigma_1^2 > 0$ are unknown, the contours of the log likelihood ratio can be equivalently indexed by the parameters

$$\begin{aligned} \tilde{a} &= a/\underline{e}_1^T \underline{x} \\ \tilde{\sigma}_0^2 &= \sigma_0^2/\underline{x}^T \underline{x} \\ \tilde{\sigma}_1^2 &= \sigma_1^2/\underline{x}^T \underline{x}. \end{aligned}$$

That is, we can express the set function (5) as

$$\ln \tilde{\Lambda}(\underline{x}) = \left\{ \frac{\tilde{a}}{\tilde{\sigma}_1^2} \cdot \frac{\left|\underline{e}_1^T \underline{x}\right|^2}{\underline{x}^T \underline{x}} + \frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_0^2}{2\tilde{\sigma}_0^2 \tilde{\sigma}_1^2} - \frac{\tilde{a}^2}{2\tilde{\sigma}_1^2} \cdot \frac{\left|\underline{e}_1^T \underline{x}\right|^2}{\underline{x}^T \underline{x}} + n \ln \frac{\tilde{\sigma}_0}{\tilde{\sigma}_1} : \tilde{a} \in \mathbb{R}; \tilde{\sigma}_0^2, \tilde{\sigma}_1^2 > 0 \right\}.$$

We conclude that the set function (5) is indexed by the scalar:

$$z(\underline{x}) = \frac{\left|\underline{e}_1^T \underline{x}\right|^2}{\underline{x}^T \underline{x}}$$

which is the maximal invariant in this example. In Fig. 2, the invariance orbit is illustrated as a cone in $\mathbb{R}^3$. Each invariance orbit $\{\underline{x} : \left|\underline{e}_1^T \underline{x}\right|^2 / \underline{x}^T \underline{x} = z\}$ is indexed by the equivalent maximal invariant, $x_1^2 / \sum_{i=2}^n x_i^2$, which determines the tangent of this cone. Thus, the compression to a scalar function of $\underline{x}$ provided by the maximal invariant is a more vigorous compression than that provided by the minimal sufficient statistic above. ■

The principle of invariance stipulates that any optimal decision rule should only depend on $\mathbf{X}$ through the maximal invariant $\mathbf{Z} = Z(\mathbf{X})$ which indexes the invariance orbits in the sense that

1. (invariant property) $Z(g(\mathbf{X})) = Z(\mathbf{X}),\ \forall\, g \in \mathcal{G}$
2. (maximal property) $Z(\mathbf{X}) = Z(\mathbf{Y}) \Rightarrow \mathbf{Y} = g(\mathbf{X}),\ g \in \mathcal{G}$.

Clearly, the maximal invariant is not unique. Any other functions of $\mathbf{X}$ related to $Z(\mathbf{X})$ in a one-to-one manner can be maximal invariant. It can also be shown that the probability density $f(\mathbf{Z}; \delta)$ of $\mathbf{Z}$ only depends on θ through a reduced set of parameters $\delta = \delta(\theta)$, which is the induced maximal invariant under $\bar{\mathcal{G}}$. Use of $\mathbf{Z}$ for detection gives the equivalent set of hypotheses

$$\begin{aligned} H_0 &: \mathbf{Z} \sim f(\mathbf{Z}; \delta(\theta_0)),\ \theta_0 \in \Theta_0 \\ H_1 &: \mathbf{Z} \sim f(\mathbf{Z}; \delta(\theta_1)),\ \theta_1 \in \Theta_1. \end{aligned}$$

Since the new parameterization $\delta(\theta)$ is generally a dimension reducing function of θ, use of the reduced data $\mathbf{Z}$ gives us better chances of finding a CFAR test whose false alarm rate is independent of θ. In particular, when $\delta(\theta_0)$ is constant over $\theta_0 \in \Theta_0$, the distribution of $\mathbf{Z}$ under H_0 is fixed and therefore any test based on $\mathbf{Z}$ will automatically be CFAR.

When there exists a group $\mathcal{G}$ that leaves a testing problem invariant, we can restrict our attention to the class of invariant tests where a test function ϕ on $\mathbf{X}$ satisfies

$$\phi(g(\mathbf{X})) = \phi(\mathbf{X}), \forall g \in \mathcal{G}.$$

Any one-to-one function of the maximal invariants produces an equivalent invariant test [20].

D. Example: Unstructured Covariance Case

We will first consider the case where the clutter is totally unknown. Suppose that the measurement matrix is Gaussian with i.i.d. columns each having the unknown covariance matrix and the problem is to decide the presence of a known target in a known subimage with an unknown amplitude. Then we can use the image model in (3), $\mathbf{X} = a\,\underline{\varepsilon}_1\,\underline{e}_1^T + \mathbf{N}$, and its partitioned form

$$\mathbf{X} = [\underline{x}_1\ \mathbf{X}_2] = \begin{bmatrix} x_{11} & \underline{x}_{12} \\ \underline{x}_{21} & \mathbf{X}_{22} \end{bmatrix} \tag{7}$$

where $\underline{x}_1$ is the first subimage which may contain the target and all the target energy has been put into the first pixel x_{11} of this subimage. This is the case studied by Kelly [9], and the results are briefly reviewed here to help illustrate the application of the GLR principle and invariance principle discussed previously. This will help the reader understand more complicated structured models of interest, covered later in this paper.

D.1 GLR Approach

Since the $m \times n$ measurement matrix $\mathbf{X}$ is complex multivariate normal with $m \times n$ mean $E[\mathbf{X}] = a\,\underline{\varepsilon}_1\,\underline{e}_1^T$ and $mn \times mn$ covariance $cov[vec(\mathbf{X})] = \mathbf{R} \bigotimes \mathbf{I}$ as described in Section II-B, the problem is to decide whether a is 0 or not when $\mathbf{R}$ is unknown. If we write the i.i.d. columns of $\mathbf{X}$ as $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n\}$, the pdf of $\mathbf{X}$ is

$$f(\mathbf{X}) = \frac{1}{\pi^{mn}|\mathbf{R}|^n} \exp\left[-(\underline{x}_1 - a\underline{\varepsilon}_1)^H \mathbf{R}^{-1}(\underline{x}_1 - a\underline{\varepsilon}_1) - \sum_{i=2}^{n} \underline{x}_i^H \mathbf{R}^{-1} \underline{x}_i\right]. \tag{8}$$

Obviously, the likelihood ratio involves unknown parameters, a and $\mathbf{R}$, and we derive the GLR by maximizing the likelihood ratio over those parameters, i.e. by replacing them with their MLEs:

$$l_1 = \frac{\max_{\theta \in \Theta_1} f(\mathbf{X}; \theta)}{\max_{\theta \in \Theta_0} f(\mathbf{X}; \theta)} = \frac{\max_a f(\mathbf{X}; a, \hat{\mathbf{R}}_1)}{f(\mathbf{X}; 0, \hat{\mathbf{R}}_0)}$$

where $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_1$ are the MLEs of $\mathbf{R}$ under H_0 and H_1, respectively. It is easily shown:

$$\begin{aligned}\hat{\mathbf{R}}_0 &= \frac{1}{n}\sum_{i=1}^{n}\underline{x}_i\underline{x}_i^H, \\ \hat{\mathbf{R}}_1 &= \frac{1}{n}\left[(\underline{x}_1 - a\underline{\varepsilon}_1)(\underline{x}_1 - a\underline{\varepsilon}_1)^H + \sum_{i=2}^{n}\underline{x}_i\underline{x}_i^H\right].\end{aligned}$$

To ensure these matrices be nonsingular with probability one, we must impose the condition that $n > m$. After some algebra, we obtain the following simple form of the GLR for this example by taking the n-th root of l_1:

$$\sqrt[n]{l_1} = \max_a\left\{\frac{1+\underline{x}_1^H(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{x}_1}{1+(\underline{x}_1 - a\underline{\varepsilon}_1)^H(\mathbf{X}_2\mathbf{X}_2^H)^{-1}(\underline{x}_1 - a\underline{\varepsilon}_1)}\right\}. \tag{9}$$

It remains to maximize this ratio over the unknown complex amplitude a. This can be done by completing the square in the denominator of (9) giving the MLE of the amplitude as

$$\hat{a} = \frac{\underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{x}_1}{\underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{\varepsilon}_1}. \tag{10}$$

Thus the GLR test is equivalent to $1 - 1/\sqrt[n]{l_1}$, denoted T_{Ku}:

$$T_{Ku} = \frac{|\underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{x}_1|^2}{\underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{\varepsilon}_1 \cdot \{1+\underline{x}_1^H(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{x}_1\}}. \tag{11}$$

This test was obtained by Kelly [9] and will be called the unstructured Kelly's test. Kelly also proved in [9] that this test has the CFAR property.

D.2 Invariance Approach

As defined above, an invariant test is a test statistic which is a function of the maximal invariants. Here, we review the derivation of the maximal invariants under the unstructured model described above, and prove that the Kelly's GLR test can be represented with the maximal invariant statistics.

With the previous model, we can define the following group of transformations acting on $\mathbf{X}$ as

$$g(\mathbf{X}) = \begin{bmatrix}\beta_1 & \underline{\beta}_2^H \\ \underline{0} & \mathbf{M}\end{bmatrix}\mathbf{X}\begin{bmatrix}1 & \underline{0}^T \\ \underline{0} & \mathbf{U}\end{bmatrix} \tag{12}$$

where $\beta_1 \neq 0$, $\underline{\beta}_2(1\times(m-1))$ and $\mathbf{M}((m-1)\times(m-1))$ are arbitrary, and $\mathbf{U}((n-1)\times(n-1))$ is a unitary matrix. In order to prove that the decision problem is invariant to this group, it is worthwhile to recall the important property of the Kronecker product that if an $m\times n$ Gaussian matrix $\mathbf{X}$ has mean $E[\mathbf{X}] = \mu$ and covariance $cov[vec(\mathbf{X})] = \mathbf{R}\bigotimes\mathbf{C}$, then $\mathbf{FXH}$ has mean $\mathbf{F}\mu\mathbf{H}$ and covariance $\mathbf{FRF}^H\bigotimes\mathbf{HCH}^H$. With this property and the model $\mathbf{X}$ in (7), we have $g(\mathbf{X}) = \tilde{a}\underline{\varepsilon}_1\underline{e}_1^T + \tilde{\mathbf{N}}$ where $\tilde{a} = \beta_1 a$ and $\tilde{\mathbf{N}}$ is still zero-mean Gaussian with $cov[vec(\tilde{\mathbf{N}})] = \tilde{\mathbf{R}}\bigotimes\mathbf{I}$ where

$$\tilde{\mathbf{R}} = \begin{bmatrix}\beta_1 & \underline{\beta}_2^H \\ \underline{0} & \mathbf{M}\end{bmatrix}\mathbf{R}\begin{bmatrix}\beta_1 & \underline{\beta}_2^H \\ \underline{0} & \mathbf{M}\end{bmatrix}^H.$$

Thus the problem remains unchanged under this group since the unknown amplitude a and covariance matrix $\mathbf{R}$ are just replaced by $\tilde{a}$ and $\tilde{\mathbf{R}}$, respectively. This group is also the group whose actions have the largest possible number of free parameters which guarantee that the decision problem remains unchanged. Indeed if the full linear group of row actions were used, i.e. the first column of the left multiplying matrix in (12) were to be arbitrary, the signal spatial structure $\underline{\varepsilon}_1$ would not be preserved. Likewise, if a larger group of right multiplying matrices than the one in (12) were applied to the columns of $\mathbf{X}$, the independence of the columns of $\mathbf{X}$ or the temporal (chip) structure $\underline{e}_1$ of the signal would not be preserved.

Once the invariant group of transformations is obtained, we can now define a set of statistics, i.e. maximal invariants, which indexes the orbits of $\mathbf{X}$ under this group.

Proposition 1: With the model (7) and the group of transformations (12), the maximal invariant is 2-dimensional:

$$\begin{aligned} z_{1'} &= \underline{x}_1^H (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{x}_1, \\ z_2 &= \underline{x}_{21}^H (\mathbf{X}_{22} \mathbf{X}_{22}^H)^{-1} \underline{x}_{21}. \end{aligned}$$

And $z_{1'}$ can be replaced by

$$z_1 = \frac{\left| x_{11} - \underline{x}_{12} \mathbf{X}_{22}^H (\mathbf{X}_{22} \mathbf{X}_{22}^H)^{-1} \underline{x}_{21} \right|^2}{\underline{x}_{12} \left[\mathbf{I} - \mathbf{X}_{22}^H (\mathbf{X}_{22} \mathbf{X}_{22}^H)^{-1} \mathbf{X}_{22} \right] \underline{x}_{12}^H}$$

since $z_{1'} = z_1 + z_2$.

Proof: Bose and Steinhardt [12]. See the appendix for an independent derivation. ■

To interpret this set of maximal invariants, consider the group of transformations (12) as

$$g(\mathbf{X}) = \begin{bmatrix} g_1(\underline{x}_1, \mathbf{X}_2) \\ g_2(\underline{x}_{21}, \mathbf{X}_{22}) \end{bmatrix} = \begin{bmatrix} \underline{\beta}^H \underline{x}_1 & \underline{\beta}^H \mathbf{X}_2 \mathbf{U} \\ \mathbf{M} \underline{x}_{21} & \mathbf{M} \mathbf{X}_{22} \mathbf{U} \end{bmatrix}$$

where $\underline{\beta}^H = [\beta_1 \ \underline{\beta}_2^H]$. From each group action on the measurement scaled by $\underline{\beta}$ or $\mathbf{M}$, and rotated by $\mathbf{U}$, we can construct a orbit (cone) as in Fig. 2. Then each cone of g_1 and g_2 is indexed respectively by $z_{1'}$ and z_2 which are the ratios of the norm squared along the axis of the cone to that perpendicular to it. z_2 is the sample correlation between primary and secondary data whose distribution is same under H_0 and H_1. Thus it is an ancillary statistic [20]. Also the representation of z_1 gives it an interpretation as the estimated $\underline{s}$-prediction SNR, i.e. the ratio of the magnitude squared of the estimated target error to that of the estimated clutter prediction error, where $\underline{x}_{12} \mathbf{X}_{22}^H (\mathbf{X}_{22} \mathbf{X}_{22}^H)^{-1} \underline{x}_{21}$ is the least-squares estimate of x_{11} given $\underline{x}_{21}$ and $\mathbf{X}_2$.

Any invariant test will be functions of z_1 and z_2, and we can show that the Kelly's test (11) is one of them. As described in Proposition 1, $\underline{x}_1^H (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{x}_1 = z_1 + z_2$ and we have

$$T_{Ku} = \frac{|\underline{\varepsilon}_1^T (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{x}_1|^2}{\underline{\varepsilon}_1^T (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{\varepsilon}_1 \cdot \{1 + z_1 + z_2\}}.$$

Then using the partition in (7) and relations for the inverse of partitioned matrices [24], we can show

$$\begin{aligned}\underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{\varepsilon}_1 &= \{\underline{x}_{12}\left[\mathbf{I}-\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\right]\underline{x}_{12}^H\}^{-1},\\ \underline{\varepsilon}_1^T(\mathbf{X}_2\mathbf{X}_2^H)^{-1}\underline{x}_1 &= \frac{x_{11}-\underline{x}_{12}\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21}}{\underline{x}_{12}\left[\mathbf{I}-\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\right]\underline{x}_{12}^H}.\end{aligned}$$

Thus

$$T_{Ku} = \frac{z_1}{1+z_1+z_2} \tag{13}$$

, which establishes that the GLR test is also an invariant test.

No optimal properties are claimed for this test, and as noted earlier the number of chips, n, must exceed the number m of spatial pixels per chip which can be quite large in many radar applications. Kelly derived the pdf of the test statistic and showed that it depends on the unknown covariance matrix $\mathbf{R}$ only through a SNR involving the unknown signal amplitude a. Thus, under the clutter-alone hypothesis H_0, the pdf of T_{Ku} is not affected by the unknown parameters, and hence the test is CFAR.

IV. Application to Target Straddling Clutter Boundary

In this section, we consider the problem of detecting a known target straddling the boundary of two independent clutter regions. From the model (4), the measurement matrix $\mathbf{X}$ is composed of two different regions A and B and can be partitioned as

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_A\\ \mathbf{X}_B\end{bmatrix} = \begin{bmatrix}\underline{x}_{A1} & \mathbf{X}_{A2}\\ \underline{x}_{B1} & \mathbf{X}_{B2}\end{bmatrix} \tag{14}$$

where $\underline{x}_{A1}$ and $\underline{x}_{B1}$ are the primary vectors which may contain the separated canonical parts of a known target, $\underline{s}_A$ and $\underline{s}_B$, respectively, with the unknown common amplitude a. Under the clutter-alone hypothesis H_0, any of the i.i.d. columns of $\mathbf{X}$ will be multivariate Gaussian with zero mean and a covariance matrix $\mathbf{R}$ having a block diagonal structure as defined in (1).

A. GLR Tests

Let $\{\underline{x}_{A1}, \underline{x}_{A2}, \ldots, \underline{x}_{An}\}$ and $\{\underline{x}_{B1}, \underline{x}_{B2}, \ldots, \underline{x}_{Bn}\}$ represent the i.i.d. columns of the two uncorrelated matrices $\mathbf{X}_A$ and $\mathbf{X}_B$, respectively, then the pdf of $\mathbf{X}$ factors as

$$f(\mathbf{X}) = f(\mathbf{X}_A)f(\mathbf{X}_B)$$

where $f(\mathbf{X}_A)$ and $f(\mathbf{X}_B)$ are defined similarly as (8) for each region. Now the decision problem is to decide whether the primary data contains clutter alone ($a = 0$) or clutter plus target ($a \neq 0$):

$$\begin{aligned}H_0 &: \mathbf{X} \sim f(\mathbf{X}; 0, \mathbf{R}_A, \mathbf{R}_B)\\ H_1 &: \mathbf{X} \sim f(\mathbf{X}; a, \mathbf{R}_A, \mathbf{R}_B)\end{aligned}$$

where $\mathbf{R}_A$ and $\mathbf{R}_B$ are the regional covariances as given in (1). As in the unstructured case, the GLR maximization can be performed for the unknown covariance matrices $\mathbf{R}_A$ and $\mathbf{R}_B$ by replacing them with their MLEs. Here, the required condition for non-singularity of the estimated covariance matrices ($n > m$) is relaxed since we need only $n > \max\{m_A, m_B\}$. This GLR, however, still involves a maximization over the unknown amplitude a in a complex quartic equation and cannot be represented in closed form. However, for real valued data the roots of the quartic equation are explicit. For complex data we implement the GLR tests, derived under the structured cases, using numerical root finding and compare their performance in Section V.

A.1 Case 1: $\mathbf{R}_A > 0, \mathbf{R}_B > 0$

The GLR for this case is just the product of the likelihood ratios from regions A and B:

$$\Lambda_1 = \max_a \frac{f(\mathbf{X}_A; a, \hat{\mathbf{R}}_{A1}) f(\mathbf{X}_B; a, \hat{\mathbf{R}}_{B1})}{f(\mathbf{X}_A; 0, \hat{\mathbf{R}}_{A0}) f(\mathbf{X}_B; 0, \hat{\mathbf{R}}_{B0})}.$$

Next we can apply the results of the unstructured example in Section III-D to both of the two regions A and B separately:

$$\frac{1}{n} \ln \Lambda_1 = \max_a \left\{ \ln \left[\frac{1 + p(0, \underline{s}_A, \mathbf{X}_A)}{1 + p(a, \underline{s}_A, \mathbf{X}_A)} \right] + \ln \left[\frac{1 + p(0, \underline{s}_B, \mathbf{X}_B)}{1 + p(a, \underline{s}_B, \mathbf{X}_B)} \right] \right\} \tag{15}$$

where

$$p(a, \underline{s}_A, \mathbf{X}_A) = (\underline{x}_{A1} - a\underline{s}_A)^H (\mathbf{X}_{A2} \mathbf{X}_{A2}^H)^{-1} (\underline{x}_{A1} - a\underline{s}_A).$$

Now we call (15) GLR 1 which reduces to the GLR in (9) when $\mathbf{R}$ is unstructured and for which the maximization over the quadratic equation in the denominator can be easily achieved.

With this structured model, however, the maximization over a cannot be completed explicitly. But since the maximizing value of the complex amplitude

$$\hat{a} = \arg\min_a \{[1 + p(a, \underline{s}_A, \mathbf{X}_A)] \cdot [1 + p(a, \underline{s}_B, \mathbf{X}_B)]\}$$

involves a product of two positive quadratic equations, we can derive upper and lower bounds to aid in numerical search. Define the local solutions from each region A, B as in (10):

$$\begin{aligned} \hat{a}_A &= \arg\min_a p(a, \underline{s}_A, \mathbf{X}_A) = \frac{\underline{s}_A^H (\mathbf{X}_{A2} \mathbf{X}_{A2}^H)^{-1} \underline{x}_{A1}}{\underline{s}_A^H (\mathbf{X}_{A2} \mathbf{X}_{A2}^H)^{-1} \underline{s}_A}, \\ \hat{a}_B &= \arg\min_a p(a, \underline{s}_B, \mathbf{X}_B) = \frac{\underline{s}_B^H (\mathbf{X}_{B2} \mathbf{X}_{B2}^H)^{-1} \underline{x}_{B1}}{\underline{s}_B^H (\mathbf{X}_{B2} \mathbf{X}_{B2}^H)^{-1} \underline{s}_B}. \end{aligned} \tag{16}$$

Then we know that $\hat{a}$ lies between those local solutions which serve as bounds, and GLR 1 can be implemented or maximized while varying a in such a way so as to guarantee

$$\begin{aligned} \min\{\mathrm{Re}\{\hat{a}_A\}, \mathrm{Re}\{\hat{a}_B\}\} \leq \quad & \mathrm{Re}\{\hat{a}\} \quad \leq \max\{\mathrm{Re}\{\hat{a}_A\}, \mathrm{Re}\{\hat{a}_B\}\}, \\ \min\{\mathrm{Im}\{\hat{a}_A\}, \mathrm{Im}\{\hat{a}_B\}\} \leq \quad & \mathrm{Im}\{\hat{a}\} \quad \leq \max\{\mathrm{Im}\{\hat{a}_A\}, \mathrm{Im}\{\hat{a}_B\}\}. \end{aligned}$$

A.2 Case 2: $\mathbf{R}_A > 0, \mathbf{R}_B = \sigma^2\mathbf{I}$

This case is just as above except that $\mathbf{R}_B$ is assumed to be diagonal with common unknown variance σ^2 along the diagonal. With this assumption, the pdf of $\mathbf{X}_B$ is

$$f(\mathbf{X}_B; a, \sigma^2) = \frac{1}{\pi^{m_B n}\sigma^{2m_B n}} \exp\left[-\frac{1}{\sigma^2}\left\{|\underline{x}_{B1} - a\underline{s}_B|^2 + \sum_{i=2}^{n}|\underline{x}_{Bi}|^2\right\}\right]$$

and the GLR is expressed as

$$\Lambda_2 = \max_a \frac{f(\mathbf{X}_A; a, \hat{\mathbf{R}}_{A1}) f(\mathbf{X}_B; a, \hat{\sigma}_1)}{f(\mathbf{X}_A; 0, \hat{\mathbf{R}}_{A0}) f(\mathbf{X}_B; 0, \hat{\sigma}_0)}.$$

Again MLEs of the variance under both hypotheses can be easily found as

$$\begin{aligned}\hat{\sigma}_1^2 &= \frac{1}{m_B n} q(a, \underline{s}_B, \mathbf{X}_B),\\ \hat{\sigma}_0^2 &= \frac{1}{m_B n} q(0, \underline{s}_B, \mathbf{X}_B)\end{aligned}$$

where

$$q(a, \underline{s}_B, \mathbf{X}_B) = \mathrm{tr}\left\{(\mathbf{X}_B - a\underline{s}_B\underline{e}_1^T)^H(\mathbf{X}_B - a\underline{s}_B\underline{e}_1^T)\right\}. \tag{17}$$

As before, the maximization over a in Λ_2 cannot be completed in closed form. To bound $\hat{a}$, we first consider the GLR over the region B alone which can be simplified to

$$l_2 = \max_a \left\{\frac{q(0, \underline{s}_B, \mathbf{X}_B)}{q(a, \underline{s}_B, \mathbf{X}_B)}\right\}^{m_B n}.$$

We named it l_2 after the previous unstructured GLR test statistic l_1 in (9). Then by rewriting $q(a, \underline{s}_B, \mathbf{X}_B)$ as

$$q(a, \underline{s}_B, \mathbf{X}_B) = |\underline{s}_B|^2 \cdot \left|a - \frac{\underline{s}_B^H \underline{x}_{B1}}{|\underline{s}_B|^2}\right|^2 + \sum_{i=1}^{n}|\underline{x}_{Bi}|^2 - \frac{|\underline{s}_B^H \underline{x}_{B1}|^2}{|\underline{s}_B|^2},$$

we see that the maximizing value of a is

$$\hat{a}_B = \frac{\underline{s}_B^H \underline{x}_{B1}}{|\underline{s}_B|^2} \tag{18}$$

where $\underline{s}_B$ is the canonical target of form $\underline{s}_B = [s_B, 0, \ldots, 0]^T$. Thus we have the equivalent form of this GLR

$$1 - \frac{1}{\sqrt[m_B n]{l_2}} = \frac{|x_{B11}|^2}{\sum_{i=1}^{n}|\underline{x}_{Bi}|^2}. \tag{19}$$

Now back to Λ_2, GLR 2 can be expressed as

$$\frac{1}{n}\ln\Lambda_2 = \max_a \left\{\ln\left[\frac{1 + p(0, \underline{s}_A, \mathbf{X}_A)}{1 + p(a, \underline{s}_A, \mathbf{X}_A)}\right] + m_B \cdot \ln\left[\frac{q(0, \underline{s}_B, \mathbf{X}_B)}{q(a, \underline{s}_B, \mathbf{X}_B)}\right]\right\}, \tag{20}$$

and the maximizing value of a can be found between $\hat{a}_A$ given as (16) for Case 1 and $\hat{a}_B$ given in (18).

A.3 Case 3: $\mathbf{R}_A > 0, \mathbf{R}_B = \mathbf{I}$

Suppose $\mathbf{R}_B$ is exactly known to be an identity matrix. Then from the results of Case 2 we can derive a bound on the maximizing value of a required to implement the GLR. Define the GLR l_3 over $\mathbf{X}_B$ alone:

$$l_3 = \max_a \{\exp[q(0, \underline{s}_B, \mathbf{X}_B) - q(a, \underline{s}_B, \mathbf{X}_B)]\}$$

where q is the same as defined previously in (17). Hence, the MLE of the amplitude a is equal to $\hat{a}_B$ given in (18), and the GLR over $\mathbf{X}_B$ is equivalent to

$$\ln l_3 = |x_{B11}|^2. \tag{21}$$

Thus, finally, we can define GLR 3 using the entire measurement $\mathbf{X}$ as

$$\frac{1}{n} \ln \Lambda_3 = \max_a \left\{ \ln \left[\frac{1 + p(0, \underline{s}_A, \mathbf{X}_A)}{1 + p(a, \underline{s}_A, \mathbf{X}_A)} \right] + \frac{1}{n} \left[q(0, \underline{s}_B, \mathbf{X}_B) - q(a, \underline{s}_B, \mathbf{X}_B) \right] \right\} \tag{22}$$

where the maximization over a can be implemented similarly to Case 2.

B. MI Tests

In this section, we apply the invariance principle to the structured covariance cases studied above and construct a test statistic as a function of the maximal invariants derived. These results parallel those of Bose and Steinhardt [12]. It will be convenient to first define the partition of $\mathbf{X}$ which is refined from (14):

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{bmatrix} = \begin{bmatrix} x_{A11} & \underline{x}_{A12} \\ \underline{x}_{A21} & \mathbf{X}_{A22} \\ x_{B11} & \underline{x}_{B12} \\ \underline{x}_{B21} & \mathbf{X}_{B22} \end{bmatrix}. \tag{23}$$

With this partition, the structured group of transformations induced by each model will be defined as

$$g(\mathbf{X}) = \begin{bmatrix} g_A(\mathbf{X}_A) \\ g_B(\mathbf{X}_B) \end{bmatrix}$$

and the maximal invariants under each group can easily be obtained. For each case, MI test is proposed based on the maximal invariants and compared to the previous results of Kelly [9] and Bose and Steinhardt [12].

B.1 Case 1: $\mathbf{R}_A > 0, \mathbf{R}_B > 0$

In this case, the independent regions A and B both have unknown covariance matrices, and we can construct a structured group of transformations on $\mathbf{X}$ which is extended from (12):

$$g(\mathbf{X}) = \begin{bmatrix} \begin{bmatrix} \beta & \underline{\beta}_A^H \\ \underline{0} & \mathbf{M}_A \end{bmatrix} & \mathbf{X}_A & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \\ \begin{bmatrix} \beta & \underline{\beta}_B^H \\ \underline{0} & \mathbf{M}_B \end{bmatrix} & \mathbf{X}_B & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \end{bmatrix} \tag{24}$$

where $\beta \neq 0$, $\underline{\beta}_A(1 \times (m_A - 1))$, $\underline{\beta}_B(1 \times (m_B - 1))$, $\mathbf{M}_A((m_A - 1) \times (m_A - 1))$ and $\mathbf{M}_B((m_B - 1) \times (m_B - 1))$ are arbitrary, and $\mathbf{U}_A$ and $\mathbf{U}_B$ are $((n-1) \times (n-1))$ unitary matrices. Showing the invariant property of this group is analogous to the unstructured example. With this group, the set of maximal invariants is defined in the following, which is also briefly covered in [12].

Proposition 2: With the model in (4) and the partition in (23), the maximal invariant under the group of transformations in (24) is 5-dimensional:

$$\begin{aligned} z_{A1} &= \frac{|u_A|^2}{D_A}, \\ z_{A2} &= \underline{x}_{A21}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \underline{x}_{A21}, \\ z_{B1} &= \frac{|u_B|^2}{D_B}, \\ z_{B2} &= \underline{x}_{B21}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \underline{x}_{B21}, \\ z_{AB} &= \frac{u_A}{u_B} \end{aligned}$$

where the subscripts denote whether the quantities are computed over the region A, B or both A and B, and

$$\begin{aligned} u_A &= x_{A11} - \underline{x}_{A12}\mathbf{X}_{A22}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \underline{x}_{A21}, \\ u_B &= x_{B11} - \underline{x}_{B12}\mathbf{X}_{B22}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \underline{x}_{B21}, \\ D_A &= \underline{x}_{A12} \left[\mathbf{I} - \mathbf{X}_{A22}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \mathbf{X}_{A22}\right] \underline{x}_{A12}^H, \\ D_B &= \underline{x}_{B12} \left[\mathbf{I} - \mathbf{X}_{B22}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \mathbf{X}_{B22}\right] \underline{x}_{B12}^H. \end{aligned}$$

And z_{AB} can be replaced by

$$z_{AB'} = \frac{|u_A/s_A - u_B/s_B|^2}{D_A/|s_A|^2 + D_B/|s_B|^2}$$

or

$$z_{AB''} = \frac{|u_A/s_A - u_B/s_B|^2}{q_A D_A/|s_A|^2 + q_B D_B/|s_B|^2}$$

where $q_A = 1 + z_{A1} + z_{A2}$ and $q_B = 1 + z_{B1} + z_{B2}$.

Proof: Bose and Steinhardt [12]. See the appendix for an independent derivation. ■

We can see that z_{A1} and z_{A2} correspond to z_1 and z_2 in the unstructured test (Proposition 1) applied to region A, and z_{B1} and z_{B2} correspond to those applied to region B. The coupling term, z_{AB}, $z_{AB'}$, or $z_{AB''}$, not present in the unstructured test, captures the common amplitude a for both regions.

Bose and Steinhardt proposed a natural modification of Kelly's test (11) which reflects the block covariance structure:

$$T_{Ks} = \frac{\left|\underline{s}^H \mathbf{K}^{-1} \underline{x}_1\right|^2}{\underline{s}^H \mathbf{K}^{-1} \underline{s} \cdot \{1 + \underline{x}_1^H \mathbf{K}^{-1} \underline{x}_1\}} \tag{25}$$

where $\underline{x}_1 = [\underline{x}_{A1}^H \ \underline{x}_{B1}^H]^H$, $\underline{s} = [\underline{s}_A^H \ \underline{s}_B^H]^H$ and

$$\mathbf{K} = \begin{bmatrix} \mathbf{X}_{A2}\mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & \mathbf{X}_{B2}\mathbf{X}_{B2}^H \end{bmatrix}.$$

To see that this is a function of maximal invariants derived in Proposition 2, first look at the term in the bracket in the denominator of (25): $1 + \underline{x}_1^H \mathbf{K}^{-1} \underline{x}_1 = 1 + z_{A1} + z_{A2} + z_{B1} + z_{B2}$ using the relation, $z_{1'} = z_1 + z_2$, in Proposition 1. We can simplify the remaining factor in the test using the results of the unstructured example:

$$\frac{\left|\underline{s}^H \mathbf{K}^{-1} \underline{x}_1\right|^2}{\underline{s}^H \mathbf{K}^{-1} \underline{s}} = \frac{\left|(D_A/|s_A|^2)^{-1} u_A/s_A + (D_B/|s_B|^2)^{-1} u_B/s_B\right|^2}{(D_A/|s_A|^2)^{-1} + (D_B/|s_B|^2)^{-1}} \tag{26}$$

where s_A and s_B are the first elements which are only non-zero in $\underline{s}_A$ and $\underline{s}_B$, respectively.

Lemma 1: Suppose that $p \times p$ matrices $\mathbf{D}_A$, $\mathbf{D}_B$ are hermitian and invertible, and $\underline{u}_A$, $\underline{u}_B$ are column vectors of size p, then

$$\begin{aligned} & (\mathbf{D}_A^{-1}\underline{u}_A + \mathbf{D}_B^{-1}\underline{u}_B)^H (\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1} (\mathbf{D}_A^{-1}\underline{u}_A + \mathbf{D}_B^{-1}\underline{u}_B) \\ = \ & \underline{u}_A^H \mathbf{D}_A^{-1} \underline{u}_A + \underline{u}_B^H \mathbf{D}_B^{-1} \underline{u}_B - (\underline{u}_A - \underline{u}_B)^H (\mathbf{D}_A + \mathbf{D}_B)^{-1} (\underline{u}_A - \underline{u}_B). \end{aligned}$$

Proof: See the appendix. ■

Using Lemma 1, the equation in (26) is a special case for $p = 1$. Hence the structured Kelly's test (25) can be expressed as

$$T_{Ks} = \frac{z_{A1} + z_{B1} - z_{AB'}}{1 + z_{A1} + z_{A2} + z_{B1} + z_{B2}}. \tag{27}$$

Alternatively, by looking at the maximal invariant representation of T_{Ks}, we can obtain another invariant test which reduces to the unstructured test (13):

$$T_1 = \frac{\left| \begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} q_A \mathbf{X}_{A2}\mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & q_B \mathbf{X}_{B2}\mathbf{X}_{B2}^H \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}_{A1} \\ \underline{x}_{B1} \end{bmatrix} \right|^2}{\begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} q_A \mathbf{X}_{A2}\mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & q_B \mathbf{X}_{B2}\mathbf{X}_{B2}^H \end{bmatrix}^{-1} \begin{bmatrix} \underline{s}_A \\ \underline{s}_B \end{bmatrix}}. \tag{28}$$

Note that q_A and q_B are placed in the estimated covariance matrix attempting to separate the coupled denominator in (27). Thus T_1 is same as (26) except for $q_A D_A$ and $q_B D_B$ in place of D_A and D_B, respectively, and from Lemma 1 we have

$$T_1 = \frac{z_{A1}}{1 + z_{A1} + z_{A2}} + \frac{z_{B1}}{1 + z_{B1} + z_{B2}} - z_{AB''} \tag{29}$$

where the different coupling term $z_{AB''}$ is used instead of $z_{AB'}$. This MI test will be shown to outperform (27) for some situations.

B.2 Case 2: $\mathbf{R}_A > 0, \mathbf{R}_B = \sigma^2 \mathbf{I}$

Now suppose $\mathbf{R}_B = \sigma^2 \mathbf{I}$ with unknown σ^2, then the invariant group of transformations in this case is

$$g(\mathbf{X}) = \begin{bmatrix} \begin{bmatrix} \beta & \underline{\beta}_A^H \\ \underline{0} & \mathbf{M}_A \end{bmatrix} & \mathbf{X}_A & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \\ \beta & \mathbf{X}_B & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \end{bmatrix} \tag{30}$$

since $\mathbf{X}_B$ still remains Gaussian under this group except that a and σ^2 are replaced by $\tilde{a} = \beta a$ and $\tilde{\sigma}^2 = (\beta\sigma)^2$. Similarly to (24), the same scaling factor β captures the common amplitude in both regions.

Proposition 3: With the partition in (23), the maximal invariant under the group of transformations in (30) is composed of

$$\begin{aligned} z_{A1} &= \frac{|u_A|^2}{D_A}, \\ z_{A2} &= \underline{x}_{A21}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \underline{x}_{A21}, \\ z_B &= \frac{|x_{B11}|^2}{\sum_{i=1}^n |\underline{x}_{Bi}|^2}, \\ z_{AB} &= \frac{u_A}{x_{B11}} \end{aligned}$$

where u_A and D_A are same as defined in Proposition 2. But, since the maximal invariant is not unique, we can also define alternative forms for z_B and z_{AB}: z_B can be replaced by

$$z_{B'} = \frac{|x_{B11}|^2}{|\underline{x}_{B12}|^2 + |\underline{x}_{B21}|^2 + |\mathbf{X}_{B22}|_F^2},$$

and z_{AB} can be replaced by either of

$$z_{AB'} = \frac{|u_A/s_A - x_{B11}/s_B|^2}{\rho D_A/|s_A|^2 + v_1/|s_B|^2}$$

where

$$\begin{aligned} \rho &= \frac{1}{(n-m_A)(1+z_{A2})}, \\ v_1 &= \frac{|\underline{x}_{B12}|^2 + |\underline{x}_{B21}|^2 + |\mathbf{X}_{B22}|_F^2}{m_B n - 1}, \end{aligned}$$

or

$$z_{AB''} = \frac{|u_A/s_A - x_{B11}/s_B|^2}{q_A D_A/|s_A|^2 + v_2/|s_B|^2}$$

where q_A is same as defined in Proposition 2 and

$$v_2 = \frac{1}{m_B}\sum_{i=1}^{n} |\underline{x}_{Bi}|^2.$$

Proof: Bose and Steinhardt [12]. See the appendix for an independent derivation. ■

z_{A1} and z_{A2} are same as those in Proposition 2, and the coupling terms are associated with the common scaling β for a. Finally, z_B or $z_{B'}$ are the maximal invariant for the case that only region B is considered.

Bose and Steinhardt derived identical maximal invariants in the context of array detection problems and the above results can all be found in [12]. In [25], a representation for the joint pdf of the maximal invariants is derived which gives insight into the marginal distributions: z_{A1}, z_{A2} and z_B as F-statistics, and z_{AB} as complex Cauchy. Based on these statistics an invariant test was proposed in [12] which was shown to be approximately CFAR:

$$T_{BS} = \frac{\left| \begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} \rho \mathbf{X}_{A2}\mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & v_1 \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}_{A1} \\ \underline{x}_{B1} \end{bmatrix} \right|^2}{\begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} \rho \mathbf{X}_{A2}\mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & v_1 \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \underline{s}_A \\ \underline{s}_B \end{bmatrix}} \tag{31}$$

where ρ and v_1 are as in Proposition 3. To see the maximal invariant representation, we write this test as

$$T_{BS} = \frac{\left|(\rho D_A/|s_A|^2)^{-1} u_A/s_A + (v_1/|s_B|^2)^{-1} x_{B11}/s_B\right|^2}{(\rho D_A/|s_A|^2)^{-1} + (v_1/|s_B|^2)^{-1}}$$

then from Lemma 1 we have

$$T_{BS} = (n-m_A) z_{A1}(1+z_{A2}) + (m_B n - 1) z_{B'} - z_{AB'}. \tag{32}$$

However, we can construct another invariant test statistic by considering the structures of both the GLR test (20) and the MI test 1 (28):

$$T_2 = \frac{\left| \begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} q_A \mathbf{X}_{A2} \mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & v_2 \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}_{A1} \\ \underline{x}_{B1} \end{bmatrix} \right|^2}{\begin{bmatrix} \underline{s}_A^H & \underline{s}_B^H \end{bmatrix} \begin{bmatrix} q_A \mathbf{X}_{A2} \mathbf{X}_{A2}^H & \mathbf{O} \\ \mathbf{O} & v_2 \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \underline{s}_A \\ \underline{s}_B \end{bmatrix}} \tag{33}$$

where ρ and v_1 in (31) are replaced by q_A and v_2 defined in Proposition 3. Then this MI test 2 has a maximal invariant form of

$$T_2 = \frac{z_{A1}}{1 + z_{A1} + z_{A2}} + m_B \cdot z_B - z_{AB''}. \tag{34}$$

Thus the weighting between the terms from region A and region B is maintained as in (20), and this test reduces exactly to the unstructured tests: (13) for $\mathbf{X}_A$ alone or (19) for $\mathbf{X}_B$ alone. This reduction does not hold for the Bose and Steinhardt's test (32).

B.3 Case 3: $\mathbf{R}_A > 0, \mathbf{R}_B = \mathbf{I}$

For this case, the invariant group of transformations is defined as

$$g(\mathbf{X}) = \begin{bmatrix} \begin{bmatrix} \beta & \underline{\beta}_A^H \\ \underline{0} & \mathbf{M}_A \end{bmatrix} & \mathbf{X}_A & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \\ & \mathbf{X}_B & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \end{bmatrix}$$

where, unlike the previous two cases, there is no scaling term on the left of $\mathbf{X}_B$ since the variance is exactly known in $\mathbf{X}_B$ and must not be altered by group actions. Thus $g(\mathbf{X})$ cannot have the common scaling term for the unknown amplitude in both regions, and the set of maximal invariants doesn't include any coupling term from regions A and B.

However, MI test 3 can be induced from MI test 2 (33) by replacing v_2 with $v_3 = n$, and we propose the following MI test 3

$$T_3 = \frac{z_{A1}}{q_A} + \frac{1}{n}|x_{B11}|^2 - \frac{|u_A/s_A - x_{B11}/s_B|^2}{q_A D_A/|s_A|^2 + n/|s_B|^2}. \tag{35}$$

Note that this test also reduces to either of the unstructured cases: (13) for $\mathbf{X}_A$ alone or (21) for $\mathbf{X}_B$ alone.

C. Extension of Tests to One of p Known Targets

In this section, the previous results are extended to the problem of detecting the presence of one target from a known set of p possible targets. Previously, the target signature in the primary vector was assumed

to be exactly known and the problem was to decide whether the one and only signal vector $\underline{s}$ is present or not. In real radar applications, however, a more realistic model can be considered. Suppose that we know the form of the target of interest, but don't know its position or orientation in the subimage. Then different target signature vectors can be constructed according to different positions and orientations in that subimage.

To accommodate this scenario, let the image model have an $m \times p$ matrix $\mathbf{S} = \left[\underline{s}_1, \ldots, \underline{s}_p\right]$ for p target signatures:

$$a\,\mathbf{S}\,\underline{\epsilon}_k\,\underline{e}_1^T + \mathbf{N} \tag{36}$$

where $\underline{\epsilon}_k$ is a $p \times 1$ unit vector $[0, \ldots, 0, 1, 0, \ldots, 0]^T$ and '1' is in position k. Here $k \in \{1, \ldots, p\}$, and $p \leq m$ for unstructured clutter or $p \leq \min\{m_A, m_B\}$ for structured clutter. The model (36) implies that only one of the signatures, $\underline{s}_k$, may be present at a time in the primary vector, and in the structured case this signature vector is written as $\underline{s}_k = \left[\underline{s}_{Ak}^H\ \underline{s}_{Bk}^H\right]^H$.

For the GLR tests (15), (20), and (22), it is easy to extend the results of the single target case to this multiple target case. We only need to replace $\underline{s}_A$ and $\underline{s}_B$ in the GLR tests with p possible target signatures $\underline{s}_{Ak}$ and $\underline{s}_{Bk}$, and maximize over $k = 1, \ldots, p$, i.e. for $i = 1, 2, 3$ indexing each of the block covariance cases discussed above:

$$\max_{k=1,\cdots,p} \frac{1}{n} \ln \Lambda_i(\underline{s}_{Ak}, \underline{s}_{Bk}).$$

Similarly, for the MI tests one can also propose to maximize over the p target signatures. In the following, the invariance procedure is applied to the model in (36) for both the unstructured and structured cases. For the structured cases, only Case 1 is investigated.

C.1 Unstructured Case

First, we consider the case of totally unknown covariance. Since $\mathbf{S}$ is known, we can define the canonical model from (36) as

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H \\ \mathbf{P}_S \end{bmatrix} \{a\,\mathbf{S}\,\underline{\epsilon}_k\,\underline{e}_1^T + \mathbf{N}\} \\ &= a \begin{bmatrix} \underline{\epsilon}_k \\ \underline{0} \end{bmatrix} \underline{e}_1^T + \tilde{\mathbf{N}} \end{aligned}$$

where an $(m-p) \times m$ matrix $\mathbf{P}_S$ is an orthogonal matrix to $(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H$ and $\tilde{\mathbf{N}}$ is still zero-mean Gaussian with i.i.d. columns. We partition $\mathbf{X}$ as before

$$\mathbf{X} = [\underline{x}_1\ \mathbf{X}_2] = \begin{bmatrix} \underline{x}_{11} & \mathbf{X}_{12} \\ \underline{x}_{21} & \mathbf{X}_{22} \end{bmatrix} \tag{37}$$

where the $p \times 1$ vector $\underline{x}_{11}$ may contain any of the target signatures which have been transformed to unit vectors $\{\underline{\epsilon}_k\}_{k=1}^p$. With this model, the group of transformations which preserves the problem is defined as

$$g(\mathbf{X}) = \begin{bmatrix} \Delta & \mathbf{B} \\ \mathbf{O} & \mathbf{M} \end{bmatrix} \mathbf{X} \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U} \end{bmatrix} \tag{38}$$

where Δ is a $p \times p$ diagonal matrix, $\mathbf{B}$ $(p \times (m-p))$ and $\mathbf{M}$ $((m-p) \times (m-p))$ are arbitrary, and $\mathbf{U}$ is an $(n-1) \times (n-1)$ unitary matrix. Note that by putting the model (36) into the canonical form, we must restrict a diagonal matrix Δ in (38) instead of an arbitrary matrix in order to preserve the known canonical form of the signal $\underline{\epsilon}_k$. This group of transformations with larger degrees of freedom will lead to a larger set of maximal invariants in the following proposition compared to the single target case.

Proposition 4: The maximal invariant of the model (37) under the group of transformations in (38) consists of $p+2$ functions of the measurement:

$$\begin{aligned} z_1 &= \underline{u}^H \mathbf{D}^{-1} \underline{u}, \\ z_2 &= \underline{x}_{21}^H (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} \underline{x}_{21}, \\ z_{3k} &= \underline{u}^H \mathbf{D}^{-1} \underline{\epsilon}_k (\underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{u} \end{aligned}$$

where $k = 1, \ldots, p$ and

$$\begin{aligned} \underline{u} &= \underline{x}_{11} - \mathbf{X}_{12}\mathbf{X}_{22}^H (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} \underline{x}_{21}, \\ \mathbf{D} &= \mathbf{X}_{12} \left[\mathbf{I} - \mathbf{X}_{22}^H (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} \mathbf{X}_{22} \right] \mathbf{X}_{12}^H. \end{aligned}$$

Proof: See the appendix. ■

Now the unstructured Kelly's test (11) can be modified by maximizing over the p target signatures $\{\underline{\epsilon}_k\}_{k=1}^p$:

$$T_{Ku} = \max_{k=1,\ldots,p} \frac{\left| \left[\underline{\epsilon}_k^T \ \underline{0}^T \right] (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \underline{x}_1 \right|^2}{\left[\underline{\epsilon}_k^T \ \underline{0}^T \right] (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \begin{bmatrix} \underline{\epsilon}_k \\ \underline{0} \end{bmatrix} \cdot \{1 + \underline{x}_1^H (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \underline{x}_1\}}.$$

We will next express this test as a function of the new maximal invariants. Since z_1 and z_2 are equivalent to those in Proposition 1 except for the dimension, it easily follows that $1 + \underline{x}_1^H (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \underline{x}_1 = 1 + z_1 + z_2$. Also using the inverse of the partitioned matrix [24] on $(\mathbf{X}_2\mathbf{X}_2^H)^{-1}$, we can write

$$\begin{aligned} \left[\underline{\epsilon}_k^T \ \underline{0}^T \right] (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \begin{bmatrix} \underline{\epsilon}_k \\ \underline{0} \end{bmatrix} &= \underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{\epsilon}_k, \\ \left[\underline{\epsilon}_k^T \ \underline{0}^T \right] (\mathbf{X}_2\mathbf{X}_2^H)^{-1} \underline{x}_1 &= \underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{u} \end{aligned}$$

and hence the Kelly's test is an invariant test of form

$$T_{Ku} = \max_{k=1,\ldots,p} \frac{z_{3k}}{1 + z_1 + z_2}.$$

C.2 Structured Case

Next consider Case 1 for the structured model. In this case, the signal model is same as (36), but with the structured target signature matrix. Then, similarly to the above unstructured case, the canonical image model is defined as

$$\mathbf{X} = a\,\underline{\tilde{s}}_k\,\underline{e}_1^T + \mathbf{N} \tag{39}$$

where $\underline{\tilde{s}}_k = \left[\underline{\epsilon}_k^T\ \underline{0}_{(m_A-p)\times 1}^T\ \underline{\epsilon}_k^T\ \underline{0}_{(m_B-p)\times 1}^T\right]^T$. Thus, this canonical form can be partitioned as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{bmatrix} = \begin{bmatrix} \underline{x}_{A11} & \mathbf{X}_{A12} \\ \underline{x}_{A21} & \mathbf{X}_{A22} \\ \underline{x}_{B11} & \mathbf{X}_{B12} \\ \underline{x}_{B21} & \mathbf{X}_{B22} \end{bmatrix}$$

and the invariant group of transformations on $\mathbf{X}$ is

$$g(\mathbf{X}) = \begin{bmatrix} \begin{bmatrix} \Delta & \mathbf{B}_A \\ \mathbf{O} & \mathbf{M}_A \end{bmatrix} & \mathbf{X}_A & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \\ \begin{bmatrix} \Delta & \mathbf{B}_B \\ \mathbf{O} & \mathbf{M}_B \end{bmatrix} & \mathbf{X}_B & \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \end{bmatrix} \tag{40}$$

where we have the same $p \times p$ diagonal matrix Δ for $\mathbf{X}_A$ and $\mathbf{X}_B$ to preserve the signal vector $\underline{\epsilon}_k$ and the same amplitude in region A and B.

Proposition 5: With the model (39) and the group of transformations in (40), the maximal invariant is obtained as

$$\begin{aligned}
z_{A1} &= \underline{u}_A^H \mathbf{D}_A^{-1} \underline{u}_A, \\
z_{A2} &= \underline{x}_{A21}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \underline{x}_{A21}, \\
z_{A3k} &= \underline{u}_A^H \mathbf{D}_A^{-1} \underline{\epsilon}_k (\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A, \\
z_{B1} &= \underline{u}_B^H \mathbf{D}_B^{-1} \underline{u}_B, \\
z_{B2} &= \underline{x}_{B21}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \underline{x}_{B21}, \\
z_{B3k} &= \underline{u}_B^H \mathbf{D}_B^{-1} \underline{\epsilon}_k (\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B, \\
z_{ABk} &= \frac{(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A}{(\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B}
\end{aligned}$$

where

$$\begin{aligned}
\underline{u}_A &= \underline{x}_{A11} - \mathbf{X}_{A12}\mathbf{X}_{A22}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \underline{x}_{A21}, \\
\underline{u}_B &= \underline{x}_{B11} - \mathbf{X}_{B12}\mathbf{X}_{B22}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \underline{x}_{B21}, \\
\mathbf{D}_A &= \mathbf{X}_{A12} \left[\mathbf{I} - \mathbf{X}_{A22}^H (\mathbf{X}_{A22}\mathbf{X}_{A22}^H)^{-1} \mathbf{X}_{A22}\right] \mathbf{X}_{A12}^H, \\
\mathbf{D}_B &= \mathbf{X}_{B12} \left[\mathbf{I} - \mathbf{X}_{B22}^H (\mathbf{X}_{B22}\mathbf{X}_{B22}^H)^{-1} \mathbf{X}_{B22}\right] \mathbf{X}_{B12}^H
\end{aligned}$$

for $k = 1, \ldots, p$. And the coupling term z_{ABk} can be replaced by

$$z_{ABk'} = \frac{\left|(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A - (\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B\right|^2}{(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} + (\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1}}$$

or $z_{ABk''}$ which is equivalent to $z_{ABk'}$ except that $q_A \mathbf{D}_A$ and $q_B \mathbf{D}_B$ are substituted for $\mathbf{D}_A$ and $\mathbf{D}_B$, respectively, where $q_A = 1 + z_{A1} + z_{A2}$ and $q_B = 1 + z_{B1} + z_{B2}$.

Proof: See the appendix. ■

Note that z_{A1}, z_{A2}, z_{B1} and z_{B2} are again equivalent to those in Proposition 2 except for the dimension (p vs. 1).

For Case 1, we had before the structured Kelly's test, T_{Ks} (25), and the MI test, T_1 (28). First, consider T_{Ks} modified to fit the multiple signature model:

$$T_{Ks} = \max_{k=1,\ldots,p} \frac{\left|\tilde{\underline{s}}_k^H \mathbf{K}^{-1} \underline{x}_1\right|^2}{\tilde{\underline{s}}_k^H \mathbf{K}^{-1} \underline{s}_k \cdot \{1 + \underline{x}_1^H \mathbf{K}^{-1} \underline{x}_1\}}$$

where $\underline{x}_1$ and $\mathbf{K}$ are same as defined in (25), but for the target signature, we have structured $\tilde{\underline{s}}_k$ as in (39). Then, as before, we have $1 + \underline{x}_1^H \mathbf{K}^{-1} \underline{x}_1 = 1 + z_{A1} + z_{A2} + z_{B1} + z_{B2}$, and from the results of the previous section and Lemma 1, the remaining term can be written as

$$\frac{\left|\underline{s}_k^H \mathbf{K}^{-1} \underline{x}_1\right|^2}{\underline{s}_k^H \mathbf{K}^{-1} \underline{s}_k} = \frac{\left|\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A + \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B\right|^2}{\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k + \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k}. \tag{41}$$

Using the Woodbury identity it can be verified that (41) is identical to $z_{A3k} + z_{B3k} - z_{ABk'}$. Thus T_{Ks} is a function of maximal invariant of form

$$T_{Ks} = \max_{k=1,\ldots,p} \frac{z_{A3k} + z_{B3k} - z_{ABk'}}{1 + z_{A1} + z_{A2} + z_{B1} + z_{B2}}.$$

MI test can also be modified by replacing the signal vector with $\underline{s}_k$ and maximizing over k. Therefore, the modified T_1 is equivalent to (41) except for $q_A \mathbf{D}_A$ and $q_B \mathbf{D}_B$ replacing $\mathbf{D}_A$ and $\mathbf{D}_B$:

$$T_1 = \max_{k=1,\ldots,p} \frac{\left|\underline{\epsilon}_k^T (q_A \mathbf{D}_A)^{-1} \underline{u}_A + \underline{\epsilon}_k^T (q_B \mathbf{D}_B)^{-1} \underline{u}_B\right|^2}{\underline{\epsilon}_k^T (q_A \mathbf{D}_A)^{-1} \underline{\epsilon}_k + \underline{\epsilon}_k^T (q_B \mathbf{D}_B)^{-1} \underline{\epsilon}_k}.$$

This can also be written as

$$T_1 = \max_{k=1,\ldots,p} \left\{ \frac{z_{A3k}}{1 + z_{A1} + z_{A2}} + \frac{z_{B3k}}{1 + z_{B1} + z_{B2}} - z_{ABk''} \right\}.$$

V. Simulation Results

To analyze the performance of the GLR and MI tests derived under the three structured covariance assumptions, Case 1, 2, and 3, receiver operating characteristic (ROC) curves are generated and compared in this section. Even though the exact distributions of the test statistics are difficult to determine, it is

well known that under H_0 the log GLR test statistic of the form $2 \ln \Lambda$ has asymptotically a chi-square distribution with number of degrees of freedom determined by the number of fixed parameters under H_0 and H_1 [26]. This asymptotic approximation can be used to determine the threshold on the GLR which ensures a given P_{FA}. In each simulation, we generated n 10×10 subimages containing 2 independent clutter regions of area m_A and m_B pixels, respectively, and a 5×5 synthetic canonical target is inserted into the first subimage in such a manner to straddle the boundary of the two different regions. Each of the subimages is then concatenated into a column vector of size 100 to obtain a $100 \times n$ measurement matrix. Each of the ROC curves (P_D vs. P_{FA}) shown below was obtained after 500 simulations.

In the following, the ROC curves are evaluated based on factors such as the target-to-clutter power ratio; the dimensional parameters, m_A, m_B and n; and the prior uncertainty on the spatial covariance $\mathbf{R}$. Case 1, 2 and 3 are considered separately under different assumptions on clutter covariance. In each case, the three GLR tests (15), (20), (22), and the three MI tests (29), (34), (35) matched to one of the three cases are compared. Also shown are ROC curves for the following tests proposed by other authors: Kelly's structured test (27) matched to Case 1, and Bose and Steinhardt's invariant test (32) matched to Case 2. We also experimented with a real image where both of our GLR and MI tests were applied to a SAR clutter image with an inserted real target at various pose angles.

A. Comparison with Different SNRs

First, we compared the detectors by varying SNR in region B (SNR_B) for Cases 1, 2 and 3. In Figs. 3 - 5, the ROC curves of 8 different tests are compared for several SNRs: Structured Kelly's test (27), Bose and Steinhardt's test (32), MI test 1 (29), MI test 2 (34), MI test 3 (35), GLR 1 (15), GLR 2 (20), and GLR 3 (22). For each case, two tests stand out as significantly better than the other six: the GLR and MI tests which are matched to the underlying scenario, e.g. GLR 1 and MI test 1 for Case 1; and GLR 2 and MI test 2 for Case 2. This confirms the results from the previous section. For Case 1, we were able to achieve performance improvement by separating the same coupled denominator for both regions found in the matched Kelly's test (27). For Case 2, the ROC improvement over the matched Bose and Steinhardt's test is explained by the weighting between two different regions which is carefully managed in GLR 2 and MI test 2. Note that, however, neither the GLR nor the MI test uniformly outperforms the other. Of particular interest are the curve crossings in the low P_{FA} regions between the GLR and the MI tests. In Fig. 3 (b), we can observe the gains in P_D of MI test 1 over GLR 1 for $P_{FA} < 0.1$. Moreover, it should be noted that the ROC of the structured Kelly's test is dominated by that of the MI test 1 in the low P_{FA} region and by that of the GLR 1 in the high P_{FA} region. In Case 2 (Fig. 4 (b)), both the MI test 2 and GLR 2 outperform Bose and Steinhardt's matched invariant test and it appears that MI test 2 slightly outperforms GLR 2 for low P_{FA}. These crossings are also observed for mismatched cases: between MI test 1 and GLR 1 in Case 2 (Fig. 4), and between MI test 2 and GLR 2 in Case 1 (Fig. 3

(b)).

For Case 3 (Fig. 5), the ROC curves for GLR 2 approach those of the matched GLR 3 in both (a) and (b) since a large number of pixels ($m_B n = 60 \times 61$) are available to generate good MLEs of the unknown variance in region B. Thus, in the following, we will concentrate on the relative performance of GLR vs. MI tests for Cases 1 and 2.

B. Comparison with Different Windows

In this section, ROC curves are compared with different ratios of m_A/m_B by up and down shifting the 10×10 windows used to collect the subimages along the boundary. In Fig. 6 for Case 1, GLR 1 performs better as m_B decreases since fewer parameters can be more accurately estimated with the same number of chips ($n = 61$): the GLR test has to estimate a covariance matrix ($\mathbf{R}_B$) of size 60×60 in (a), but only of size 40×40 in (c). For the smaller size covariance of (b), the structured Kelly's test is almost as accurate as the GLR and MI tests. Conversely, in Fig. 7 of Case 2, GLR 2 performs better as m_B increases since in this case it only needs to estimate the scalar variance in region B and the number of pixels available increases as m_B increases ((a) $m_B n = 60 \times 61$ vs. (c) $m_B n = 40 \times 61$). Also Bose and Steinhardt's test is more sensitive to the number m_B than MI test 2 and GLR 2, and its ROC falls below even those of the mismatched tests shown in (b) and (c).

The relative advantages of MI vs. GLR tests are more closely investigated in the next two figures. In Figs. 8 and 9, we consider Case 1 and Case 2, respectively. In (a) of both figures, we increased the number of chips n while fixing SNR. Note that the GLR and MI tests have ROCs which are virtually indistinguishable for large n. In (b), however, we fixed n and increased SNR. The P_{FA} positions of the crossings of the ROCs for the GLR and MI tests decreased with increasing SNR. In particular, if one fixes a level of false alarm, say $P_{FA} = 0.1$, then note from Fig. 8 (b) that the GLR test dominates the MI test for SNR = 19 dB while the reverse is true for SNR = 7 dB. This behavior is best explained by the fact that at high SNR, the MLE is an accurate estimate of target amplitude, while at low SNR the MLE degrades significantly. Therefore, the GLR which depends on the accuracy of the MLE for accurate detection breaks down for low SNR.

C. Application to Real Image

Finally, we consider an application to real SAR imagery in Fig. 10. The image shown is a rural scene near Redstone Arsenal at Huntsville, Alabama, reproduced from the data collected using the Sandia National Laboratories Twin Otter SAR sensor payload operating at X band (center frequency = 9.6 GHz, band width = 590 MHz). This clutter image consists of a forest canopy on top and a field on bottom, separated by a coarse boundary. The boundary was hand extracted, and a sequence of 9×7 SLICY

targets at different poses were also hand extracted from the image data in Fig. 11. The images in Fig. 11 correspond to the same target but viewed at different pose angles of azimuth. The elevation of 39° was fixed for all poses. The data from which these images are reproduced was downloaded from the MSTAR SAR database at the Center for Imaging Science (www.cis.jhu.edu).

In a first experiment the target signature at pose of azimuth 163° from Fig. 11 (e) was tested at different positions along the boundary. In Fig. 10, the target is inserted additively with the center at column 305 so that it straddles the boundary. From the realigned image in Fig. 12, we took subimages (chips) along the boundary by centering a 20×20 window at the boundary and sliding it over the image from left to right. Each of these subimages is then concatenated into a column vector of size $m = 400$ where $m_A = 200$ and $m_B = 200$. Since we need at least 200 secondary chips to implement the structured detectors, clutter-alone pixels above and below those 20×20 subimages taken along the boundary were used to generate enough secondary data for region A and B, respectively. Each of the subimages along the boundary was tested as a primary chip, and the test statistics derived under Case 1 were calculated and maximized over each possible location in the subimage. After normalizing the known target signature, we obtained the minimum magnitude of target amplitude required for each test to detect the target at the correct location. The resulting amplitude is the minimum detectable threshold for each of the detectors and these thresholds are shown in Table I for different number of secondary chips $(n-1)$. As can be seen, with a large number of chips $(n-1=250)$, both the GLR and MI tests perform as well as the structured Kelly's test. On the other hand, with a limited number of chips $(n-1=200)$, MI test 1 successfully detects the target down to a significantly lower threshold than for GLR 1 and structured Kelly detectors.

Test	$\lvert a \rvert$	
	$(n-1=250)$	$(n-1=200)$
Structured Kelly	1.407×10^{-2}	1.049×10^{-1}
MI test 1	1.454×10^{-2}	0.609×10^{-1}
GLR 1	1.462×10^{-2}	1.042×10^{-1}

TABLE I

MINIMUM DETECTABLE AMPLITUDES FOR DETECTION OF THE TARGET AT THE CORRECT LOCATION.

As a final experiment we maximized the test statistics over the different target poses in Fig. 11 as well as over all possible locations along the boundary. Again the normalized signature from Fig. 11 (e) was inserted with $|a| = 0.015$, and 250 secondary chips were obtained from the surrounding clutter. Test values for the 3 detectors under Case 1 are obtained using 9 different target signatures. For each test the peak values for 9 target signatures are plotted in Fig. 13. Note that all the tests successfully picked the signature at the true pose and location for this target amplitude.

VI. Conclusion

This paper considered the problem of detecting a target lying across a clutter boundary. Two detection strategies were investigated: the GLR and the MI test procedures. Both detectors have comparable ROC performance when a large number of independent clutter samples are available, but the MI test can outperform the GLR test for a small number of independent clutter samples.

Several issues for further research should be addressed.

- The linearity assumption should be relaxed to accommodate canopy interactions with the target. This is a very challenging problem.
- For applications where the Gaussian clutter assumption may not be appropriate, non-Gaussian models should be investigated such as elliptically symmetric distribution or spherically invariant random vector (SIRV) distribution. Many results exist for GLR and invariant tests in this case [20].
- Since the known boundary assumption may not be realistic, edge/boundary estimation and its interaction with detection should be investigated including sensitivity of detector performance to boundary estimations and tradeoffs between segmentation and detection.
- For spatial acquisition mode SAR, the case should be considered where a target may lie across two different chips.
- The methods described herein can also be applied to other detection problems involving boundary and target interactions. Examples include: detection of cancer nodules imbedded on lung tissues, and detection of astronomical objects through partially turbulent atmospheres.

Appendix

I. Proof of Proposition 1

The maximal invariant should satisfy both the invariant and the maximal properties under the defined group of transformations. Before showing those properties, note that the group action can be partitioned as

$$g(\mathbf{X}) = \begin{bmatrix} \underline{\beta}^H \underline{x}_1 & \underline{\beta}^H \mathbf{X}_2 \mathbf{U} \\ \mathbf{M}\underline{x}_{21} & \mathbf{M}\mathbf{X}_{22}\mathbf{U} \end{bmatrix}$$

where $\underline{\beta}^H = \left[\beta_1 \ \underline{\beta}_2^H\right]$. Then the invariant property follows directly:

$$\begin{aligned} z_{1'}(g(\mathbf{X})) &= \underline{x}_1^H \underline{\beta}(\underline{\beta}^H \mathbf{X}_2 \mathbf{U}\mathbf{U}^H \mathbf{X}_2^H \underline{\beta})^{-1} \underline{\beta}^H \underline{x}_1 \\ &= \underline{x}_1^H (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{x}_1 \\ &= z_{1'}(\mathbf{X}), \\ z_2(g(\mathbf{X})) &= \underline{x}_{21}^H \mathbf{M}^H (\mathbf{M}\mathbf{X}_{22}\mathbf{U}\mathbf{U}^H \mathbf{X}_{22}^H \mathbf{M}^H)^{-1} \mathbf{M} \underline{x}_{21} \\ &= \underline{x}_{21}^H (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} \underline{x}_{21} \\ &= z_2(\mathbf{X}). \end{aligned}$$

Now to show the maximal property, let $z_{1'}(\mathbf{X}) = z_{1'}(\mathbf{Y})$ or

$$\underline{x}_1^H (\mathbf{X}_2 \mathbf{X}_2^H)^{-1} \underline{x}_1 = \underline{y}_1^H (\mathbf{Y}_2 \mathbf{Y}_2^H)^{-1} \underline{y}_1.$$

Then by the Vinograd's theorem [3],

$$(\mathbf{Y}_2 \mathbf{Y}_2^H)^{-\frac{1}{2}} \underline{y}_1 = \mathbf{H} (\mathbf{X}_2 \mathbf{X}_2^H)^{-\frac{1}{2}} \underline{x}_1$$

for some $m \times m$ orthogonal matrix $\mathbf{H}$, and we have

$$\underline{y}_1 = \mathbf{F} \underline{x}_1 \tag{42}$$

where $\mathbf{F} = (\mathbf{Y}_2 \mathbf{Y}_2^H)^{\frac{1}{2}} \mathbf{H} (\mathbf{X}_2 \mathbf{X}_2^H)^{-\frac{1}{2}}$. Also from this result,

$$\underline{y}_1^H (\mathbf{Y}_2 \mathbf{Y}_2^H)^{-1} \underline{y}_1 = \underline{x}_1^H \mathbf{F}^H (\mathbf{Y}_2 \mathbf{Y}_2^H)^{-1} \mathbf{F} \underline{x}_1,$$

thus $(\mathbf{Y}_2 \mathbf{Y}_2^H)^{-1} = (\mathbf{F}\mathbf{X}_2 \mathbf{X}_2^H \mathbf{F}^H)^{-1}$ or

$$\mathbf{Y}_2 = \mathbf{F}\mathbf{X}_2 \mathbf{U} \tag{43}$$

for some $(n-1) \times (n-1)$ orthogonal matrix $\mathbf{U}$. Therefore, from (42) and (43), we have

$$\mathbf{Y} = \mathbf{F}\mathbf{X} \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U} \end{bmatrix}. \tag{44}$$

Next by using z_2, i.e. $\underline{x}_{21}^H (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} \underline{x}_{21} = \underline{y}_{21}^H (\mathbf{Y}_{22}\mathbf{Y}_{22}^H)^{-1} \underline{y}_{21}$, and the Vinograd's theorem again, we have

$$\left[\underline{y}_{21} \ \mathbf{Y}_{22}\right] = [\underline{0} \ \mathbf{M}] \, [\underline{x}_1 \ \mathbf{X}_2] \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U} \end{bmatrix} \tag{45}$$

where $\mathbf{M} = (\mathbf{Y}_{22}\mathbf{Y}_{22}^H)^{\frac{1}{2}} \mathbf{J} (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-\frac{1}{2}}$ for some $(m-1) \times (m-1)$ orthogonal matrix $\mathbf{J}$. Then from (44) and (45), it is verified that $\mathbf{Y} = g(\mathbf{X})$. Therefore, $\{z_{1'}, z_2\}$ satisfies both the invariant and the maximal properties, and hence uniquely indexes the orbits of $\mathbf{X}$ under the group action.

Now we can easily verify that $z_{1'} = z_1 + z_2$ by using the relations for the inverse of a partitioned matrix [24]. Define

$$(\mathbf{X}_2\mathbf{X}_2^H)^{-1} = \begin{bmatrix} \underline{x}_{12}\underline{x}_{12}^H & \underline{x}_{12}\mathbf{X}_{22}^H \\ \mathbf{X}_{22}\underline{x}_{12}^H & \mathbf{X}_{22}\mathbf{X}_{22}^H \end{bmatrix}^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix},$$

then again from that relations:

$$\begin{aligned} V_{11} &= \{\underline{x}_{12}\left[\mathbf{I} - \mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\right]\underline{x}_{12}^H\}^{-1}, \\ V_{12} &= -V_{11}\underline{x}_{12}\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}, \\ V_{21} &= -V_{11}(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\underline{x}_{12}^H, \\ V_{22} &= (\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1} + V_{11}(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\underline{x}_{12}^H\underline{x}_{12}\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}. \end{aligned}$$

Therefore, plugging these values into the equation

$$z_{1'} = x_{11}^H V_{11} x_{11} + \underline{x}_{21}^H V_{21} x_{11} + x_{11}^H V_{12}\underline{x}_{21} + \underline{x}_{21}^H V_{22}\underline{x}_{21},$$

we have

$$\begin{aligned} z_{1'} &= V_{11}\left|x_{11} - \underline{x}_{12}\mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21}\right|^2 + \underline{x}_{21}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21} \\ &= z_1 + z_2 \end{aligned}$$

and hence $\{z_1, z_2\}$ can also serve as the maximal invariant.

II. Proof of Proposition 2

From Proposition 1, we can see clearly that $\{z_{A1}, z_{A2}\}$ is the maximal invariant corresponding to the group of transformations

$$g_A(\mathbf{X}_A) = \begin{bmatrix} \beta_1 & \underline{\beta}_A^H \\ \underline{0} & \mathbf{M}_A \end{bmatrix} \mathbf{X}_A \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \tag{46}$$

and $\{z_{B1}, z_{B2}\}$ to the group of transformations

$$g_B(\mathbf{X}_B) = \begin{bmatrix} \beta_2 & \underline{\beta}_B^H \\ \underline{0} & \mathbf{M}_B \end{bmatrix} \mathbf{X}_B \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \tag{47}$$

where we can only use arbitrary β_1 and β_2 separately for each group. So it suffices to show that z_{AB} is in the maximal invariant set which gives $\beta_1 = \beta_2 = \beta$.

Since the group action (24) can be partitioned as

$$g(\mathbf{X}) = \begin{bmatrix} \beta x_{A11} + \underline{\beta}_A^H\underline{x}_{A21} & (\beta\underline{x}_{A12} + \underline{\beta}_A^H\mathbf{X}_{A22})\mathbf{U}_A \\ \mathbf{M}_A\underline{x}_{A21} & \mathbf{M}_A\mathbf{X}_{A22}\mathbf{U}_A \\ \beta x_{B11} + \underline{\beta}_B^H\underline{x}_{B21} & (\beta\underline{x}_{B12} + \underline{\beta}_B^H\mathbf{X}_{B22})\mathbf{U}_B \\ \mathbf{M}_B\underline{x}_{B21} & \mathbf{M}_B\mathbf{X}_{B22}\mathbf{U}_B \end{bmatrix}$$

with the partition in (14), the following results are first calculated for convenience:

$$\begin{aligned} u_A(g(\mathbf{X})) &= \beta u_A(\mathbf{X}) \\ D_A(g(\mathbf{X})) &= |\beta|^2 D_A(\mathbf{X}), \end{aligned}$$

and

$$\begin{aligned} u_B(g(\mathbf{X})) &= \beta u_B(\mathbf{X}) \\ D_B(g(\mathbf{X})) &= |\beta|^2 D_B(\mathbf{X}). \end{aligned}$$

Then, it is easily verified that z_{AB} is invariant under $g(\mathbf{X})$ since

$$z_{AB}(g(\mathbf{X})) = \frac{\beta u_A}{\beta u_B} = z_{AB}(\mathbf{X}).$$

Now for the maximal property we need to show that

$$z_{AB}(\mathbf{X}) = z_{AB}(\mathbf{Y}) \;\Rightarrow\; \beta_1 = \beta_2$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_A \\ \mathbf{Y}_B \end{bmatrix} = \begin{bmatrix} g_A(\mathbf{X}_A) \\ g_B(\mathbf{X}_B) \end{bmatrix} \tag{48}$$

with g_A in (46) and g_B in (47). Then it is also straightforward since, from $z_{AB}(\mathbf{X}) = z_{AB}(\mathbf{Y})$, we have

$$\frac{u_A}{u_B} = \frac{\beta_1 u_A}{\beta_2 u_B}.$$

Thus $\beta_1 = \beta_2$ and we have proved that

$$\mathbf{Y} = g(\mathbf{X}).$$

Next, $z_{AB'}$ and $z_{AB''}$ can be shown to be the alternative terms for z_{AB} by expressing them as functions of the maximal invariant previously verified. First, we can write

$$z_{AB'} = \frac{|u_B/s_B|^2 \cdot \left| \dfrac{u_A/s_A}{u_B/s_B} - 1 \right|^2}{D_B/|s_B|^2 \cdot \left(\dfrac{D_A/|s_A|^2}{D_B/s_B|^2} + 1 \right)}.$$

Thus, $z_{AB'}$ is a function of the maximal invariant of form

$$z_{AB'} = z_{B1} \cdot \frac{\left| z_{AB} \cdot \dfrac{s_B}{s_A} - 1 \right|^2}{\dfrac{D_A}{D_B} \cdot \dfrac{|s_B|^2}{|s_A|^2} + 1}$$

where s_A and s_B is known, and D_A/D_B is just a supplementary term to z_{AB}. Also $z_{AB''}$ can be represented similarly with the additional terms q_A and q_B which are already functions of the maximal invariant, and this completes the proof.

III. Proof of Proposition 3

We know from Proposition 2 that $\{z_{A1}, z_{A2}\}$ is the maximal invariant to the group of transformations on $\mathbf{X}_A$

$$g_A(\mathbf{X}_A) = \begin{bmatrix} \beta_1 & \underline{\beta}_A^H \\ \underline{0} & \mathbf{M}_A \end{bmatrix} \mathbf{X}_A \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_A \end{bmatrix} \tag{49}$$

where $\beta_1 \neq 0$ is an arbitrary scalar. Therefore, we need to show that z_B is the maximal invariant to the group action on $\mathbf{X}_B$

$$g_B(\mathbf{X}_B) = \beta_2 \mathbf{X}_B \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U}_B \end{bmatrix} \tag{50}$$

where $\beta_2 \neq 0$ is also an arbitrary scalar, and finally $\beta_1 = \beta_2$ with z_{AB}.

First, write z_B as

$$z_B = \frac{|x_{B11}|^2}{tr\{\mathbf{X}_B^H \mathbf{X}_B\}}.$$

Then the invariant property is easily followed:

$$z_B(g(\mathbf{X})) = \frac{|\beta x_{B11}|^2}{tr\{(\beta \mathbf{X}_B)^H (\beta \mathbf{X}_B)\}} = z_B(\mathbf{X})$$

since

$$tr\{\mathbf{A}\} = tr\{\mathbf{P}^H \mathbf{A} \mathbf{P}\}$$

for any $n \times n$ matrix $\mathbf{A}$ and orthogonal matrix $\mathbf{P}$, [24]. Next, for the maximal property, let $z_B(\mathbf{X}_B) = z_B(\mathbf{Y}_B)$, then

$$\frac{|x_{B11}|^2}{tr\{\mathbf{X}_B^H \mathbf{X}_B\}} = \frac{|y_{B11}|^2}{tr\{\mathbf{Y}_B^H \mathbf{Y}_B\}}$$

or

$$x_{B11}^H \left[tr\{\mathbf{X}_B^H \mathbf{X}_B\}\right]^{-1} x_{B11} = y_{B11}^H \left[tr\{\mathbf{Y}_B^H \mathbf{Y}_B\}\right]^{-1} y_{B11}.$$

Thus, from the Vinograd's theorem, we have

$$y_{B11} = \beta_2 x_{B11} \tag{51}$$

where $\beta_2 = \left[tr\{\mathbf{Y}_B^H\mathbf{Y}_B\}\right]^{1/2}\mathbf{H}\left[tr\{\mathbf{X}_B^H\mathbf{X}_B\}\right]^{-1/2}$ for some orthogonal matrix $\mathbf{H}$. Then, we can also write

$$\frac{|\beta_2 x_{B11}|^2}{tr\{\mathbf{Y}_B^H\mathbf{Y}_B\}} = \frac{|x_{B11}|^2}{tr\{\mathbf{X}_B^H\mathbf{X}_B\}}$$

and from this, we have

$$tr\{\mathbf{Y}_B^H\mathbf{Y}_B\} = tr\{(\beta_2\mathbf{X}_B)^H(\beta_2\mathbf{X}_B)\}$$

or

$$\mathbf{Y}_B = \beta_2\mathbf{X}_B\mathbf{U} \tag{52}$$

for some unitary matrix $\mathbf{U}$. From (51) and (52), we can say $\mathbf{Y}_B = g_B(\mathbf{X}_B)$ as in (50) and z_B is the maximal invariant under g_B on $\mathbf{X}_B$. In addition, since z_B can be written as

$$z_B = \frac{|x_{B11}|^2}{|x_{B11}|^2 + |\underline{x}_{B12}|^2 + |\underline{x}_{B21}|^2 + |\mathbf{X}_{B22}|_F^2} = \frac{1}{1 + 1/z_{B'}},$$

$z_{B'}$ is also a maximal invariant which can be substituted for z_B.

Now it is quite simple to prove z_{AB} as in the proof of Proposition 2. As before, the invariant property is easily verified since

$$z_{AB}(g(\mathbf{X})) = \frac{\beta u_A}{\beta x_{B11}} = z_{AB}(\mathbf{X}),$$

and for the maximal property, we have

$$\frac{u_A(\mathbf{X}_A)}{x_{B11}} = \frac{u_A(\mathbf{Y}_A)}{y_{B11}} \tag{53}$$

from $z_{AB}(\mathbf{X}) = z_{AB}(\mathbf{Y})$. Since we have already proved that $\mathbf{Y}_A = g_A(\mathbf{X}_A)$ with g_A in (49) and $\mathbf{Y}_B = g_B(\mathbf{X}_B)$ with g_B in (50), we can write

$$\begin{aligned} u_A(\mathbf{Y}_A) &= \beta_1 u_A(\mathbf{X}_A), \\ y_{B11} &= \beta_2 x_{B11}. \end{aligned}$$

Thus, from (53), $\beta_1 = \beta_2$ and z_{AB} implies the common scaling term β in (30).

Finally, the proof for the alternative terms, $z_{AB'}$ and $z_{AB''}$, are easily followed from the proof of Proposition 2 since both terms are equivalent to $z_{AB''}$ in Proposition 2 except for x_{B11} instead of u_B, and the invariant terms ρ, v_1 and v_2.

IV. Proof of Proposition 4

Since the group action $g(\mathbf{X})$ in (38) can be partitioned as

$$g(\mathbf{X}) = \begin{bmatrix} \Delta\underline{x}_{11} + \mathbf{B}\underline{x}_{21} & (\Delta\mathbf{X}_{12} + \mathbf{B}\mathbf{X}_{22})\mathbf{U} \\ \mathbf{M}\underline{x}_{21} & \mathbf{M}\mathbf{X}_{22}\mathbf{U} \end{bmatrix},$$

the following results are first calculated for convenience:

$$\begin{aligned}
\underline{u}(g(\mathbf{X})) &= (\Delta\underline{x}_{11} + \mathbf{B}\underline{x}_{21}) - (\Delta\mathbf{X}_{12} + \mathbf{B}\mathbf{X}_{22})\mathbf{X}_{22}(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21} \\
&= \Delta\{\underline{x}_{11} - \mathbf{X}_{12}\mathbf{X}_{22}(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21}\} \\
&= \Delta\underline{u}(\mathbf{X}), \\
\mathbf{D}(g(\mathbf{X})) &= (\Delta\mathbf{X}_{12} + \mathbf{B}\mathbf{X}_{22})\left[\mathbf{I} - \mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\right](\mathbf{X}_{12}^H\Delta^H + \mathbf{X}_{22}^H\mathbf{B}^H) \\
&= \Delta\{\mathbf{X}_{12}\left[\mathbf{I} - \mathbf{X}_{22}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\mathbf{X}_{22}\right]\mathbf{X}_{12}^H\}\Delta^H \\
&= \Delta\mathbf{D}(\mathbf{X})\Delta^H.
\end{aligned}$$

Then with the partitioned structure of $\mathbf{X}$ and the above results, we can easily verify the invariant property as follows:

$$\begin{aligned}
z_1(g(\mathbf{X})) &= \underline{u}^H\Delta^H(\Delta\mathbf{D}\Delta^H)^{-1}\Delta\underline{u} \\
&= \underline{u}^H\mathbf{D}^{-1}\underline{u} \\
&= z_1(\mathbf{X}), \\
z_2(g(\mathbf{X})) &= \underline{x}_{21}^H\mathbf{M}^H(\mathbf{M}\mathbf{X}_{22}\mathbf{U}\mathbf{U}^H\mathbf{X}_{22}^H\mathbf{M}^H)^{-1}\mathbf{M}\underline{x}_{21} \\
&= \underline{x}_{21}^H(\mathbf{X}_{22}\mathbf{X}_{22}^H)^{-1}\underline{x}_{21} \\
&= z_2(\mathbf{X}),
\end{aligned}$$

and

$$\begin{aligned}
z_{3k}(g(\mathbf{X})) &= \underline{u}^H\Delta^H(\Delta\mathbf{D}\Delta^H)^{-1}\underline{\epsilon}_k\left[\underline{\epsilon}_k^T(\Delta\mathbf{D}\Delta^H)^{-1}\underline{\epsilon}_k\right]^{-1}\underline{\epsilon}_k^T(\Delta\mathbf{D}\Delta^H)^{-1}\Delta\underline{u} \\
&= \underline{u}^H\mathbf{D}^{-1}(\Delta^{-1}\underline{\epsilon}_k)\left[(\Delta^{-1}\underline{\epsilon}_k)^H\mathbf{D}^{-1}(\Delta^{-1}\underline{\epsilon}_k)\right]^{-1}(\Delta^{-1}\underline{\epsilon}_k)^H\mathbf{D}^{-1}\underline{u} \\
&= \underline{u}^H\mathbf{D}^{-1}\underline{\epsilon}_k(\underline{\epsilon}_k^T\mathbf{D}^{-1}\underline{\epsilon}_k)^{-1}\underline{\epsilon}_k^T\mathbf{D}^{-1}\underline{u} \\
&= z_{3k}(\mathbf{X}).
\end{aligned}$$

Next, for the maximal property, it is easily followed from Proposition 1 that $z_1(\mathbf{X}) = z_1(\mathbf{Y})$ and $z_2(\mathbf{X}) = z_2(\mathbf{Y})$ gives

$$\mathbf{Y} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{M} \end{bmatrix}\mathbf{X}\begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \mathbf{U} \end{bmatrix}$$

where we have a $p \times p$ non-zero matrix $\mathbf{A}$ instead of β_1 in (12) and others are defined in (38). Note that this is a general case of Proposition 1 ($p = 1$) and the proof directly follows that of Proposition 1 except for the dimension.

Now we only need to show that $\mathbf{A}$ is a $p \times p$ diagonal matrix with z_{3k}. Let $\mathbf{A}^{-1} = \left[\underline{\gamma}_1, \ldots, \underline{\gamma}_p\right]$, then

$$\begin{aligned} z_{3k}(\mathbf{Y}) &= \underline{u}^H \mathbf{A}^H (\mathbf{A}\mathbf{D}\mathbf{A}^H)^{-1} \underline{\epsilon}_k \left[\underline{\epsilon}_k^T (\mathbf{A}\mathbf{D}\mathbf{A}^H)^{-1} \underline{\epsilon}_k\right]^{-1} \underline{\epsilon}_k^T (\mathbf{A}\mathbf{D}\mathbf{A}^H)^{-1} \mathbf{A}\underline{u} \\ &= \underline{u}^H \mathbf{D}^{-1} (\mathbf{A}^{-1}\underline{\epsilon}_k) \left[(\mathbf{A}^{-1}\underline{\epsilon}_k)^H \mathbf{D}^{-1} (\mathbf{A}^{-1}\underline{\epsilon}_k)\right]^{-1} (\mathbf{A}^{-1}\underline{\epsilon}_k)^H \mathbf{D}^{-1}\underline{u} \\ &= \underline{u}^H \mathbf{D}^{-1} \underline{\gamma}_k (\underline{\gamma}_k^H \mathbf{D}^{-1} \underline{\gamma}_k)^{-1} \underline{\gamma}_k^H \mathbf{D}^{-1} \underline{u}. \end{aligned}$$

Then from $z_{3k}(\mathbf{X}) = z_{3k}(\mathbf{Y})$,

$$\underline{u}^H \mathbf{D}^{-1} \frac{\underline{\epsilon}_k \underline{\epsilon}_k^T}{\underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{\epsilon}_k} \mathbf{D}^{-1} \underline{u} = \underline{u}^H \mathbf{D}^{-1} \frac{\underline{\gamma}_k \underline{\gamma}_k^H}{\underline{\gamma}_k^H \mathbf{D}^{-1} \underline{\gamma}_k} \mathbf{D}^{-1} \underline{u}.$$

Thus we have

$$\frac{\underline{\epsilon}_k \underline{\epsilon}_k^T}{\underline{\epsilon}_k^T \mathbf{D}^{-1} \underline{\epsilon}_k} = \frac{\underline{\gamma}_k \underline{\gamma}_k^H}{\underline{\gamma}_k^H \mathbf{D}^{-1} \underline{\gamma}_k}$$

, which gives $\underline{\gamma}_k = \delta_k^{-1} \underline{\epsilon}_k$ for some scalar $\delta_k \neq 0$ and $k = 1, \ldots, p$. This means that

$$\mathbf{A} = diag(\underline{\delta})$$

where $\underline{\delta} = [\delta_1, \ldots, \delta_p]$.

V. Proof of Proposition 5

From the proof of Proposition 4, we know that $\{z_{A1}, z_{A2}, z_{A3k}\}$ and $\{z_{B1}, z_{B2}, z_{B3k}\}$ are associated with the groups

$$\begin{aligned} g_A(\mathbf{X}_A) &= \begin{bmatrix} \Delta_A & \mathbf{B}_A \\ \mathbf{O} & \mathbf{M}_A \end{bmatrix} \mathbf{X}_A \begin{bmatrix} 1 & \underline{0}^H \\ \underline{0} & \mathbf{U} \end{bmatrix}, \\ g_B(\mathbf{X}_B) &= \begin{bmatrix} \Delta_B & \mathbf{B}_B \\ \mathbf{O} & \mathbf{M}_B \end{bmatrix} \mathbf{X}_B \begin{bmatrix} 1 & \underline{0}^H \\ \underline{0} & \mathbf{U} \end{bmatrix}, \end{aligned}$$

respectively. So it suffices to show that $\Delta_A = \Delta_B = \Delta$ with z_{ABk}.

First, the invariant property of z_{ABk} directly follows from the properties of $\underline{u}$ and $\mathbf{D}$ on $g(\mathbf{X})$ in the proof of Proposition 4. Next, for the maximal property of z_{ABk}, let $z_{ABk}(\mathbf{X}) = z_{ABk}(\mathbf{Y})$ with

$$\mathbf{Y} = \begin{bmatrix} g_A(\mathbf{X}_A) \\ g_B(\mathbf{X}_B) \end{bmatrix}$$

where $\Delta_A = diag([\delta_{A1}, \ldots, \delta_{Ap}])$ and $\Delta_B = diag([\delta_{B1}, \ldots, \delta_{Bp}])$, then

$$\frac{(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A}{(\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B} = \frac{\delta_{Ak} (\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{u}_A}{\delta_{Bk} (\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1} \underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{u}_B}.$$

Therefore, $\delta_{Ak} = \delta_{Bk}$ for $k = 1, \ldots, p$ and we have proved that $\Delta_A = \Delta_B$.

Finally, we can substitute $z_{ABk'}$ for z_{ABk} since $z_{ABk'}$ is a function of the previously obtained maximal invariant of form

$$z_{ABk'} = z_{B3k} \cdot \frac{|z_{ABk} - 1|^2}{\frac{(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1}}{(\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1}} + 1}$$

where $(\underline{\epsilon}_k^T \mathbf{D}_A^{-1} \underline{\epsilon}_k)^{-1} / (\underline{\epsilon}_k^T \mathbf{D}_B^{-1} \underline{\epsilon}_k)^{-1}$ is just a supplementary term for z_{ABk}. Similarly, $z_{ABk''}$ can also be shown to be a substitute for the coupling term with the additional functions of the maximal invariant q_A and q_B.

VI. Proof of Lemma 1

We can write the equation as

$$\begin{aligned} & (\underline{u}_A^H \mathbf{D}_A^{-1} + \underline{u}_B^H \mathbf{D}_B^{-1})(\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1}(\mathbf{D}_A^{-1}\underline{u}_A + \mathbf{D}_B^{-1}\underline{u}_B) \\ = \quad & \underline{u}_A^H \mathbf{D}_A^{-1}(\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1}\mathbf{D}_A^{-1}\underline{u}_A + \underline{u}_A^H \mathbf{D}_A^{-1}(\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1}\mathbf{D}_B^{-1}\underline{u}_B \\ + \quad & \underline{u}_B^H \mathbf{D}_B^{-1}(\mathbf{D}_B^{-1} + \mathbf{D}_A^{-1})^{-1}\mathbf{D}_A^{-1}\underline{u}_A + \underline{u}_B^H \mathbf{D}_B^{-1}(\mathbf{D}_B^{-1} + \mathbf{D}_A^{-1})^{-1}\mathbf{D}_B^{-1}\underline{u}_B \end{aligned}$$

and from the Woodbury identity, we have either

$$\begin{aligned} (\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1} &= \mathbf{D}_A - \mathbf{D}_A(\mathbf{D}_B + \mathbf{D}_A)^{-1}\mathbf{D}_A, \\ \text{or} \quad (\mathbf{D}_B^{-1} + \mathbf{D}_A^{-1})^{-1} &= \mathbf{D}_B - \mathbf{D}_B(\mathbf{D}_A + \mathbf{D}_B)^{-1}\mathbf{D}_B. \end{aligned}$$

Thus, applying this identity, the equation becomes

$$\begin{aligned} & (\underline{u}_A^H \mathbf{D}_A^{-1} + \underline{u}_B^H \mathbf{D}_B^{-1})(\mathbf{D}_A^{-1} + \mathbf{D}_B^{-1})^{-1}(\mathbf{D}_A^{-1}\underline{u}_A + \mathbf{D}_B^{-1}\underline{u}_B) \\ = \quad & \underline{u}_A^H \mathbf{D}_A^{-1}\underline{u}_A + \underline{u}_B^H \mathbf{D}_B^{-1}\underline{u}_B - (\underline{u}_A - \underline{u}_B)^H (\mathbf{D}_A + \mathbf{D}_B)^{-1}(\underline{u}_A - \underline{u}_B) + L_1 + L_2 \end{aligned}$$

where

$$\begin{aligned} L_1 &= \underline{u}_A^H \left[\mathbf{D}_B^{-1} - (\mathbf{D}_A + \mathbf{D}_B)^{-1} - (\mathbf{D}_A + \mathbf{D}_B)^{-1}\mathbf{D}_A\mathbf{D}_B^{-1}\right] \underline{u}_B, \\ L_2 &= \underline{u}_B^H \left[\mathbf{D}_A^{-1} - (\mathbf{D}_A + \mathbf{D}_B)^{-1} - (\mathbf{D}_A + \mathbf{D}_B)^{-1}\mathbf{D}_B\mathbf{D}_A^{-1}\right] \underline{u}_A. \end{aligned}$$

Now we can remove the extra terms L_1 and L_2 since

$$\begin{aligned} L_1 &= \underline{u}_A^H (\mathbf{D}_A + \mathbf{D}_B)^{-1} \left[(\mathbf{D}_A + \mathbf{D}_B) - \mathbf{D}_B - \mathbf{D}_A\right] \mathbf{D}_B^{-1}\underline{u}_B = 0, \\ L_2 &= \underline{u}_B^H (\mathbf{D}_A + \mathbf{D}_B)^{-1} \left[(\mathbf{D}_A + \mathbf{D}_B) - \mathbf{D}_A - \mathbf{D}_B\right] \mathbf{D}_A^{-1}\underline{u}_A = 0 \end{aligned}$$

and this completes the proof.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*, New York: Wiley, 1968.

[2] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*, New York: Academic Press, 1967.

[3] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, New York: Wiley, 1982.

[4] M. L. Eaton, *Group Invariance Applications in Statistics*, *Regional Conference Series in Probability and Statistics*, vol. 1, Institute of Mathematical Statistics, 1989.

[5] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Reading, MA: Addison-Wesley, 1991.

[6] L. L. Scharf and D. W. Lytle, "Signal detection in Gaussian noise of unknown level: an invariance application," *IEEE Transactions on Information Theory*, vol. IT-17, no. 4, pp. 404-411, July 1971.

[7] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146-2157, August 1994.

[8] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 853-863, November 1974.

[9] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, pp. 115-127, March 1986.

[10] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-28, no. 1, pp. 208-216, January 1992.

[11] D. Fuhrmann, "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 39, no. 10, pp. 2194-2198, October 1991.

[12] S. Bose and A. O. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2164-2175, September 1995.

[13] A. O. Hero and C. Guillouet, "Robust detection of SAR/IR targets via invariance," *Proceedings of IEEE International Conference on Image Processing*, vol. 3, pp. 472-475, October 1997.

[14] F. T. Ulaby, R. K. Moore and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, Reading, MA: Addison-Wesley, 1981.

[15] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, San Diego: Academic Press, 1997.

[16] R. D. De Roo, F. T. Ulaby, A. E. El-Rouby and A. Y. Nashashibi, "MMW radar scattering statistics of terrain at near grazing incidence," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 1010-1018, July 1999.

[17] J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 46-59, January 1987.

[18] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760-1770, October 1990.

[19] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Technical Report 848, Lincoln Laboratory, Massachusetts Institute of Technology, April 1989.

[20] T. Kariya and B. K. Sinha, *Robustness of Statistical Tests*, San Diego: Academic Press, 1989.

[21] K. A. Burgess and B. D. Van Veen, "Subspace-based adaptive generalized likelihood ratio detection," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 912-927, April 1996.

[22] E. L. Lehmann, *Testing Statistical Hypotheses*, New York: Wiley, 1959.

[23] I. A. Ibragimov and R. Z. Hasminskii, *Statistical Estimation: Asymptotic Theory*, New York: Springer-Verlag, 1981.

[24] F. A. Graybill, *Matrices with Applications in Statistics*, 2nd ed., Belmont, CA: Wadsworth, 1983.

[25] S. Bose, "Invariant hypothesis testing with sensor arrays," Ph. D. thesis, Cornell University, Ithaca, NY, January 1995.

[26] P. J. Bickel and K. A. Doksom, *Mathematical Statistics: Basic Ideas and Selected Topics*, San Francisco: Holden-Day, 1977.

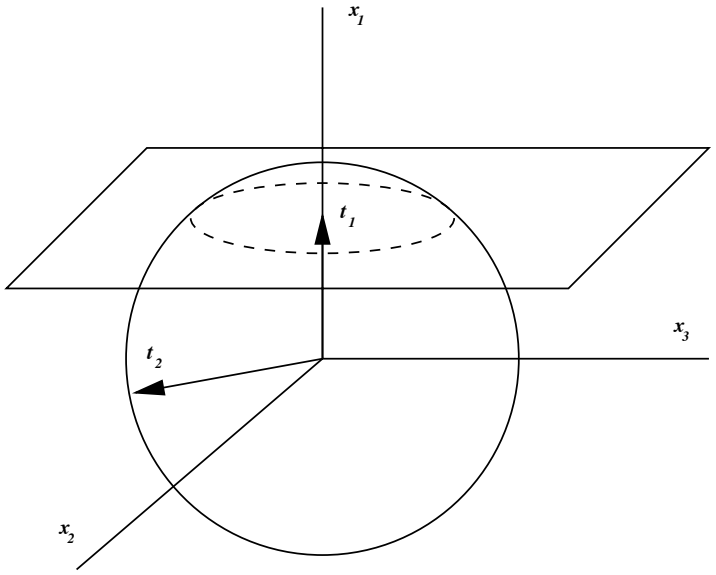


Fig. 1. Sufficiency orbit is a circle in $\mathbb{R}^3$

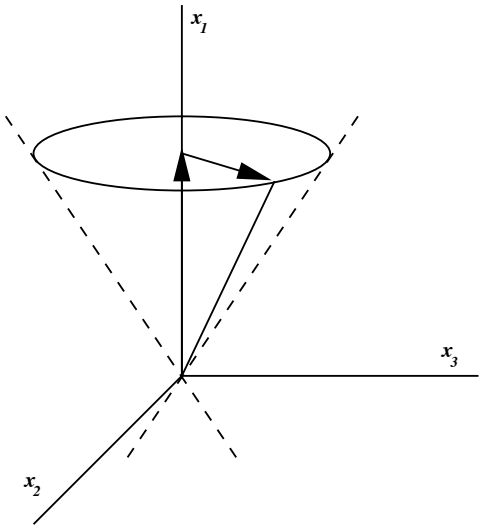
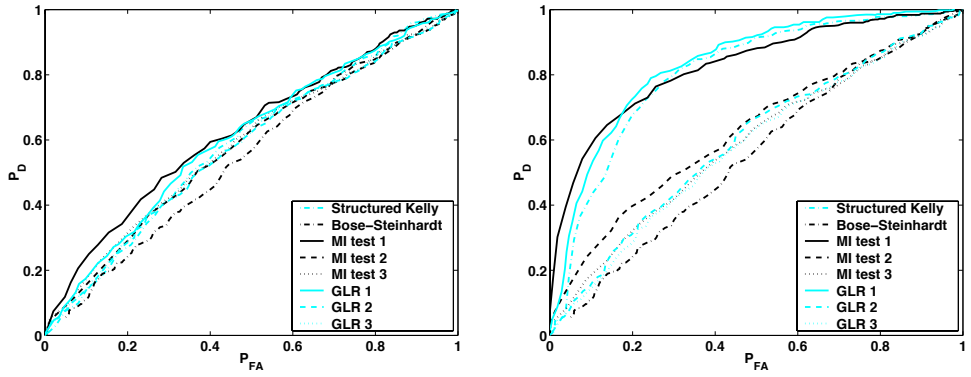
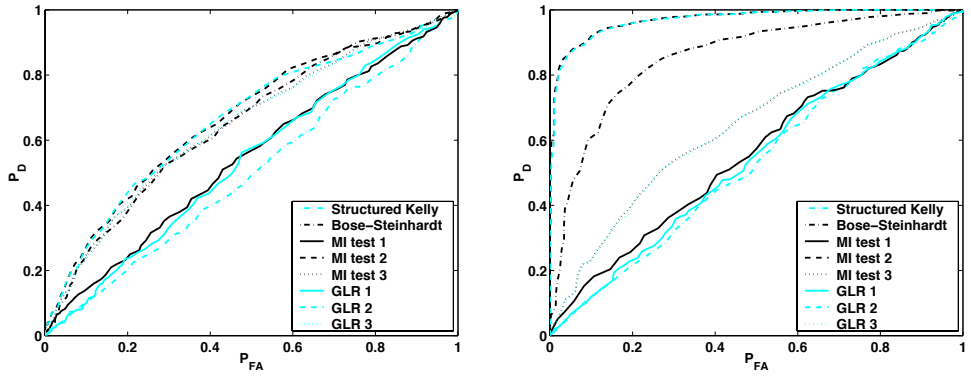


Fig. 2. Invariance orbit is a cone in $\mathbb{R}^3$



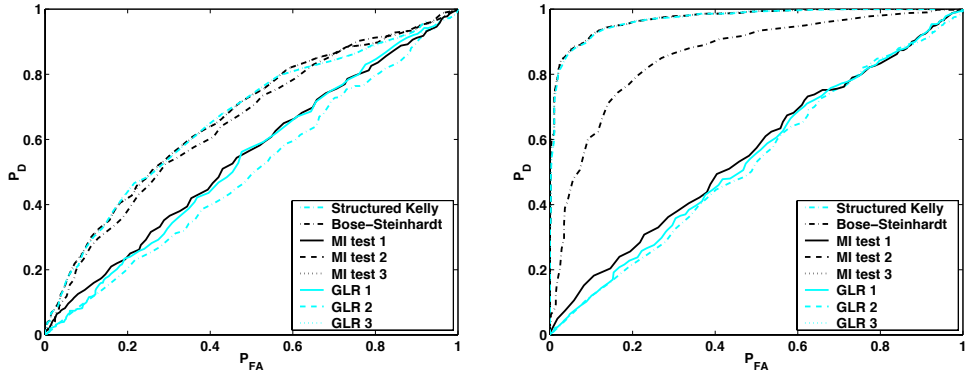
(a) $\mathrm{SNR}_A = 11$ dB, $\mathrm{SNR}_B = 10$ dB (b) $\mathrm{SNR}_A = 11$ dB, $\mathrm{SNR}_B = 22$ dB

Fig. 3. ROC curves for Case 1 with (a) SNR = 14 dB, (b) SNR = 22 dB ($m_A = 50, m_B = 50, n = 51$).



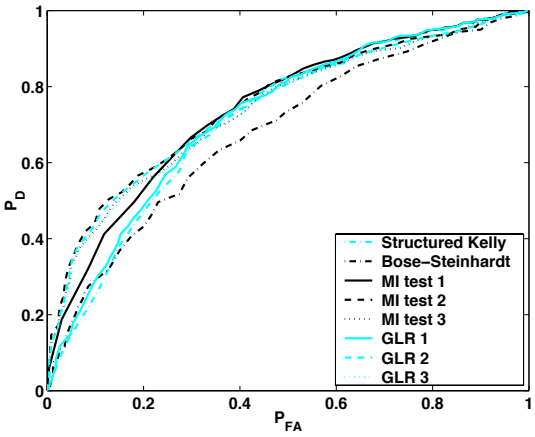
(a) $\mathrm{SNR}_A = 3$ dB, $\mathrm{SNR}_B =$ -6 dB (b) $\mathrm{SNR}_A = 3$ dB, $\mathrm{SNR}_B = 8$ dB

Fig. 4. ROC curves for Case 2 with (a) SNR = 4 dB, (b) SNR = 10 dB ($m_A = 40, m_B = 60, n = 61$).

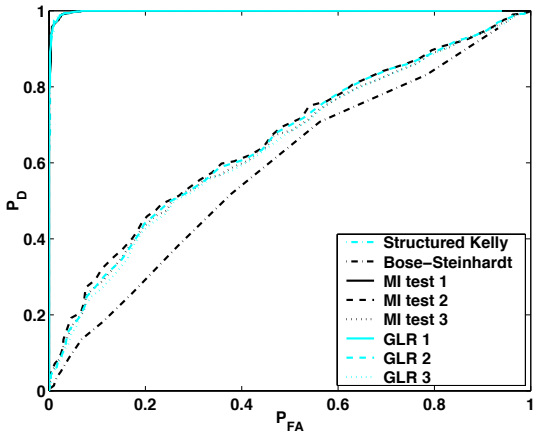


(a) $\mathrm{SNR}_A = 3$ dB, $\mathrm{SNR}_B =$ -6 dB (b) $\mathrm{SNR}_A = 3$ dB, $\mathrm{SNR}_B = 8$ dB

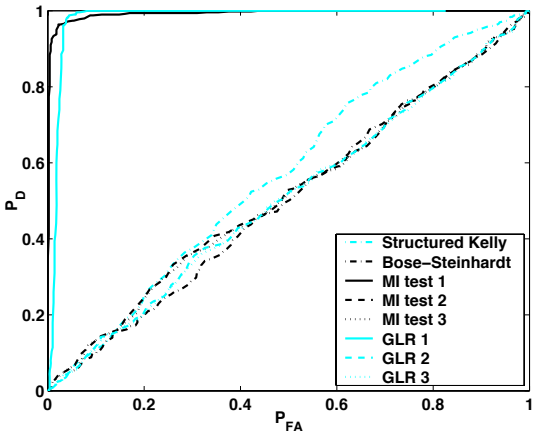
Fig. 5. ROC curves for Case 3 with (a) SNR = 4 dB, (b) SNR = 10 dB ($m_A = 40, m_B = 60, n = 61$).



(a) $m_A = 40, m_B = 60$

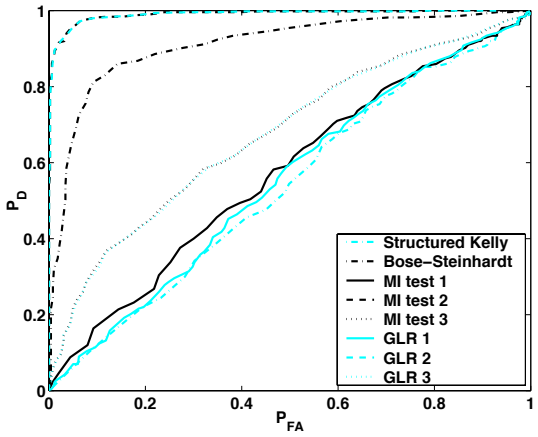


(b) $m_A = 50, m_B = 50$

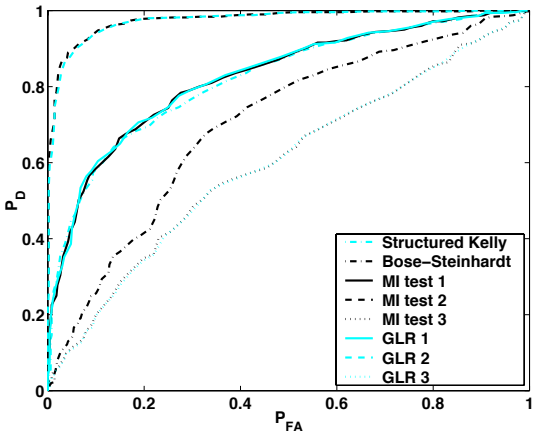


(c) $m_A = 60, m_B = 40$

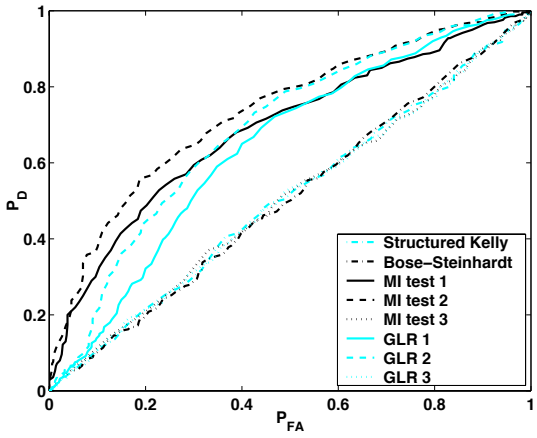
Fig. 6. ROC curves for Case 1 with different ratios of m_A/m_B, and SNR = 19 dB ($n = 61$).



(a) $m_A = 40, m_B = 60$

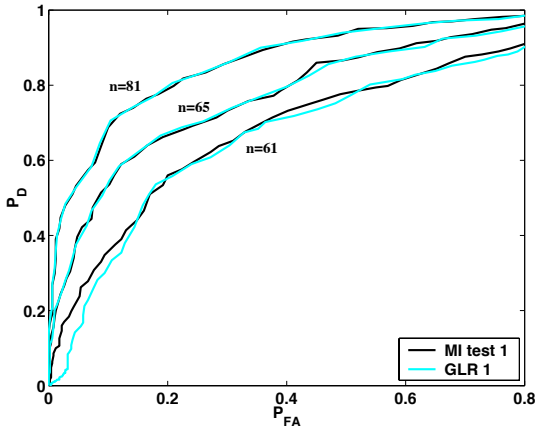


(b) $m_A = 50, m_B = 50$

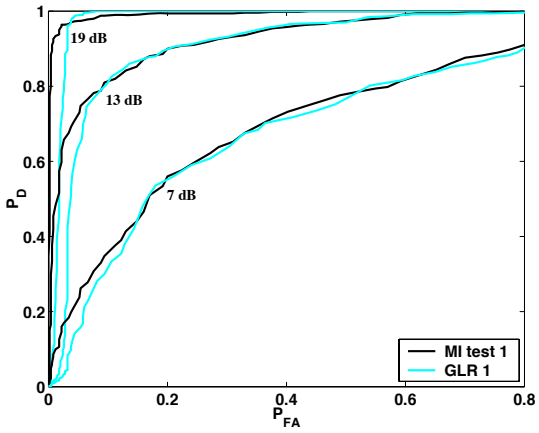


(c) $m_A = 60, m_B = 40$

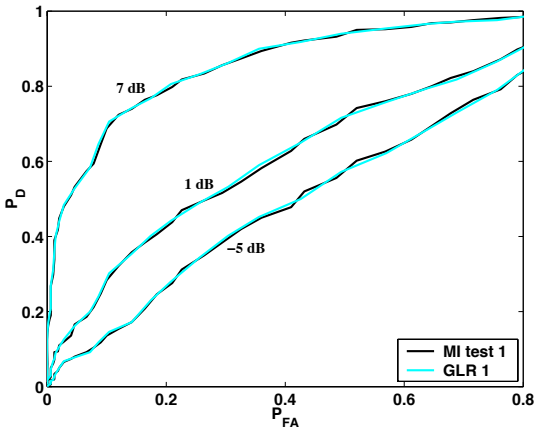
Fig. 7. ROC curves for Case 2 with different ratios of m_A/m_B, and SNR = 10 dB ($n = 61$).



(a) SNR = 7 dB

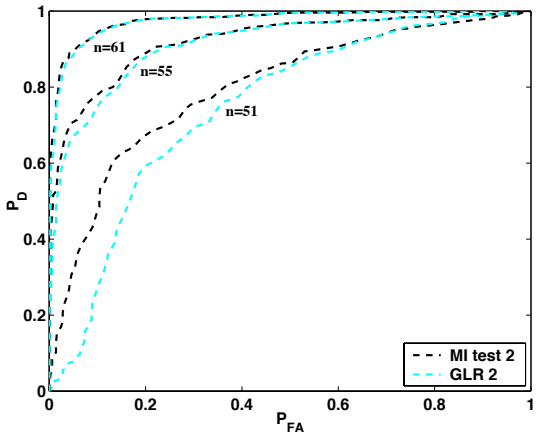


(b) $n = 61$

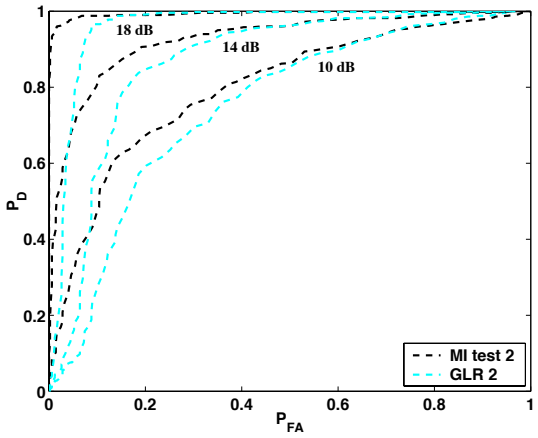


(c) $n = 81$

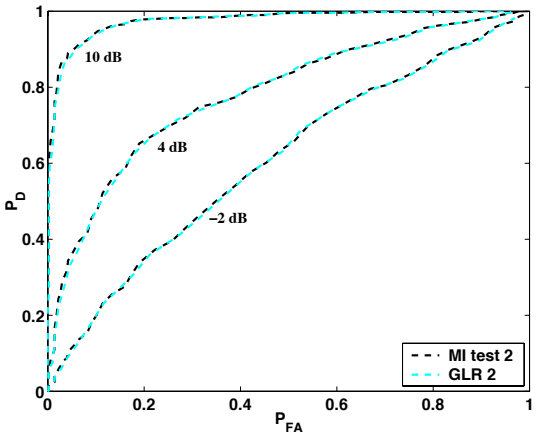
Fig. 8. Comparison of GLR and MI tests for Case 1 by (a) varying n with fixed SNR, (b) increasing SNR with small n, and (c) decreasing SNR with large n ($m_A = 60, m_B = 40$).



(a) SNR = 10 dB



(b) $n = 51$



(c) $n = 61$

Fig. 9. Comparison of GLR and MI tests for Case 2 by (a) varying n with fixed SNR, (b) increasing SNR with small n, and (c) decreasing SNR with large n ($m_A = 50, m_B = 50$).

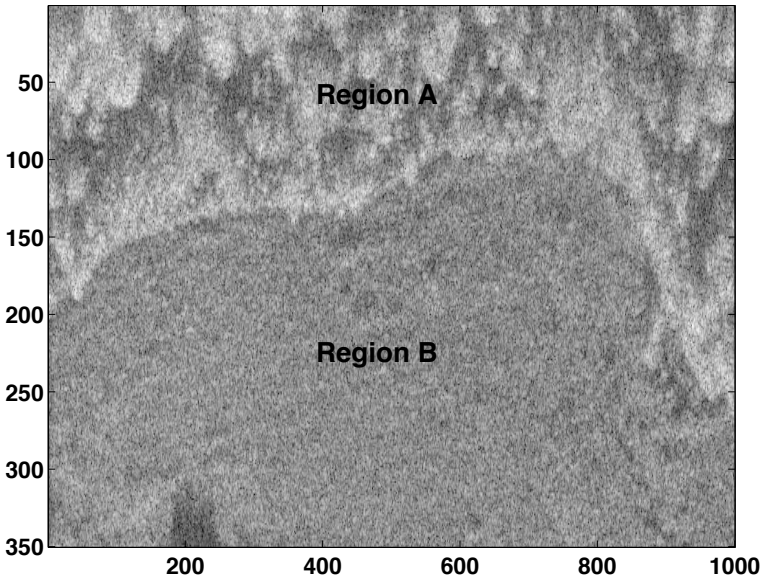


Fig. 10. SAR clutter image with target in Fig. 11 (e) straddling the boundary at column 305.

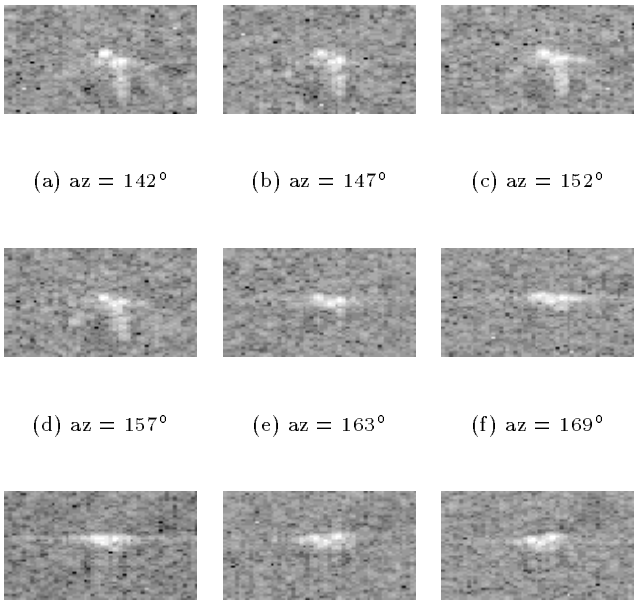


Fig. 11. SLICY canonical target images at elevation 39° and different azimuth angles (az). Image in (e) is inserted in Fig. 10.

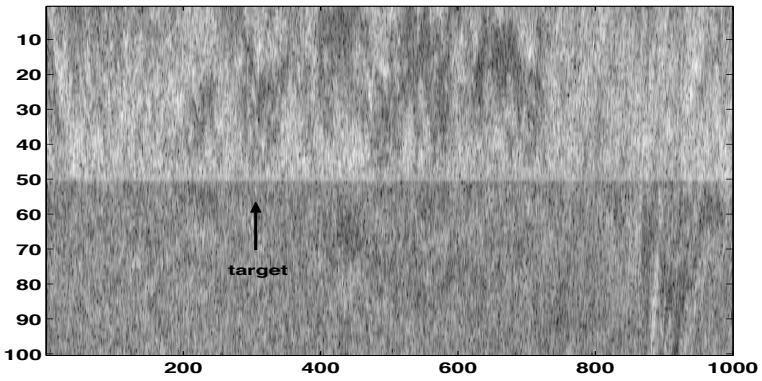
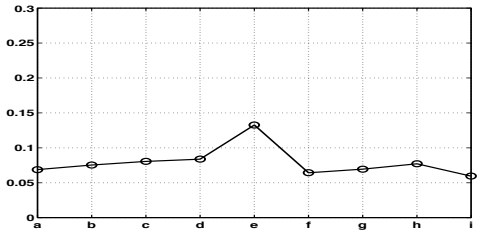
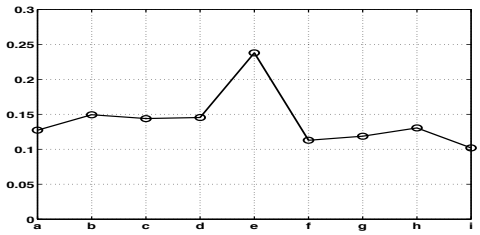


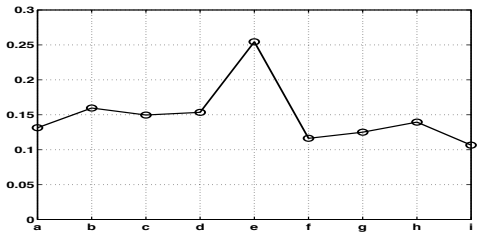
Fig. 12. Image realigned along the extracted boundary. SLICY target is located at column 305 with $|a| = 0.015$. This target is just above the minimal detectable threshold for the three tests investigated in Fig. 13.



(a) Structured Kelly's test values



(b) MI test values



(c) GLR test values

Fig. 13. Peak values obtained by (a) structured Kelly's test, (b) MI test 1 and (c) GLR 1 for 9 different target images in Fig. 11 ($|a| = 0.015, n - 1 = 250$).