Defence Research and Development Canada
Recherche et développement pour la défense Canada

Decision surfaces of binary tests in hyperspectral detection

François Bouffard
DRDC Valcartier

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2013-098
April 2013

Canada

Decision surfaces of binary tests in hyperspectral detection

François Bouffard
DRDC Valcartier

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2013-098
April 2013

IMPORTANT INFORMATIVE STATEMENTS

Abstract

This report is first and foremost an analysis of the decision surfaces associated with common detection statistics used for hyperspectral target detection. The intention is to clarify the process leading to detection using those statistics and show where their differences and similarities lie. The large number of available detection statistics, along with their numerous parameters, makes formal comparisons complicated and sometimes produces ambiguous or uncertain conclusions. This document tries to establish a solid framework on which to base this analysis and then demonstrates how detection statistics can be grouped in distinct classes based on their decision surface's geometry. In doing so, possibly new detection statistics (or methods that can be used to build them) are exposed.

Résumé

Ce rapport se veut principalement une analyse des surfaces de décision associées aux statistiques de détection utilisées couramment en détection hyperspectrale de cibles. L'intention première est la clarification du procédé menant à la détection en utilisant ces statistiques, ainsi que l'exposition de leur différence et similarités. Le grand nombre de statistiques de détection disponibles, couplées avec leurs nombreux paramètres, rend les comparaisons formelles complexes qui parfois produisent des conclusions ambigues ou incertaines. Ce document tente d'établir un formalisme solid sur lequel cette analyse peut se baser et démontre ensuite comment les statistiques de détection peuvent être groupées en classes distinctes en se basant sur la géométrie de leurs surfaces de décision. Ce faisant, des statistiques de détection possiblement nouvelles, ainsi que des méthodes pouvant être utilisées pour les bâtir, seront exposées.

This page intentionally left blank.

Executive summary

Decision surfaces of binary tests in hyperspectral detection

François Bouffard; DRDC Valcartier TR 2013-1213-000; Defence R&D Canada –Valcartier; April 2013.

Background: Defence R&D Canada – Valcartier is involved in the stand-off detection of gases or aerosols in the atmosphere, and liquids or solids on the ground, vehicules and structures, using passive hyperspectral measurements. Algorithms were previously studied and developed in order to automate the detection and identification of a wide range of substances. The detection methods described in the literature use many various formalisms, sometimes coming from different fields (acoustics, radar, etc.) and most often using different notations expressing the same general ideas differently. Comparing the performances of various detection algorithms is difficult because of theses differences and depend entirely on the ability to properly implement the proposed methods.

Principal results: This work presents a unifying formalism which allows the comparison of detection statistics by examination of their most basic geometric features, namely the shape of the decision surfaces themselves. It was found that most popular detection statistics can be grouped in two classes that share the same basic decision surface geometries (either planes or cones). Variations of these statistics give rise to cylindrical, wedge-shaped, closed, or hyperbolic decision surfaces, the exact position of which can be adjusted with few parameters. Extensions to this framework are also proposed, leading to new detection statistics having novel properties.

Significance of results: The main result of this work is the implication that multiple detection statistics can be implemented through the parameterization of a single expression. This greatly simplifies the work required to implement different detection algorithms on various platforms, including specialized hardware such as digital signal processors or graphical processing units. A corollary of this main result is that it demonstrates the futility of testing and comparing the performances of detection statistics from the same class since they can be made to exhibit the exact same decision surface using the appropriate parameters. The extensions to the proposed framework also provide a process to create new decision statistics to solve specific detection problems. Generally, this framework allows an increased comprehension of decision surfaces that will be used to evaluate the potential usefulness of any further detection algorithms.

Future work: This work is principally of a theoretical nature. A simulated or experimental confirmation of the predicted equivalence of many detection statics found in the literature would be interesting as this would allow implementing a large number of statistics with a compact code base, and generally facilitate optimization. Work on the proposed extensions to the framework presented herein, in which new detection statistics could be created, evaluated and compared, could also lead to operational advantages in specific hyperspectral detection scenarios.

Sommaire

Decision surfaces of binary tests in hyperspectral detection

François Bouffard; DRDC Valcartier TR 2013-121 ; R & D pour la défense Canada –Valcartier ; avril 2013.

Contexte : R & D pour la défense Canada – Valcartier est intéressé par la détection de gaz et d'aérosols dans l'atmosphère, ou de liquides et solides sur le sol, les véhicules ou les structures, en utilisant des mesures hyperspectrales passives. Des algorithmes ont été étudiés et développés précédemment pour automatiser la détection et l'identification d'un grand nombre de substances. Les méthodes de détection décrites dans la littérature utilisent des formalismes variés, parfois provenant de champs d'études différents (tels l'acoustique ou le radar) et utilisent le plus souvent des notations incompatibles qui expriment des idées similaires de manière entièrement différente.

Résultats : Ce travail présente un formalisme général qui permet la comparaison des statistiques de détection par l'examen de leur propriétés géométriques les plus fondamentales, c'est-à-dire la forme même de leur surfaces de décision. Il a été déterminé que la plupart des statistiques de détection les plus populaires pouvaient être groupées en deux classes dont les membres partagent la même surface de décision (plans et cones, respectivement). Des variantes de ces statistiques produisent des surfaces de décision cylindriques, en forme de coin, fermées, ou hyperboliques, dont la position exacte peut être ajustée avec un petit nombre de paramètres. Des extensions à ce formalisme sont également proposées, et pouvant mener à de nouvelles surfaces décision possédant des propriétés uniques.

Importance : Le résultat précipal de se travail est la réalisation que de multiples statistiques de décision peuvent être implantées par la paramétrisation d'une seule expression. Ceci simplifie énormément le travail requis pour l'implantation de différents algorithmes de détection sur différentes plateformes, incluant le matériel spécialisé tel les processeurs de signaux digitaux et les processeurs graphiques. Un corollaire de ce résultat est la démonstration qu'il est futile de tester et comparer les performances de détection de statistiques appartenant à la même classe, puisqu'elles peuvent adopter exactement la même surface de décision en ajustant précisément les paramètres adéquats. Les extensions au formalisme proposé suggèrent également une manière de créer de nouvelles statistiques de détection pouvant être utilisées pour des problèmes de détection spécifiques. Généralement, ce formalisme permet une compréhension accrue des surfaces de décision qui sera utilisée pour évaluer l'utilité potentielle de tout algorithme de détection subséquent.

Perspectives : Ce travail est principalement de nature théorique. Une confirmation simulée ou expérimentale de l'équivalence prédite entre plusieurs statistiques de détection retrouvées dans la littérature serait évidemment intéressante, puisque ceci permettrait de calculer un grand nombre de statistiques avec un nombre restreint de lignes de code. L'optimisation de ces statistiques de détection serait également grandement facilitée . Des travaux futurs sur les extensions proposées au formalisme présenté ici, dans lesquels de nouvelles statistique de détection pourraient être créées, évaluées et comparées, pourraient également mener a des avantages opérationels dans des scénarios de détection hyperspectrale spécifiques.

Table of contents

List of figures

List of tables

1 Introduction

This report is first and foremost an analysis of the decision surfaces associated with detection statistics commonly used for hyperspectral target detection. The intention is to clarify the process leading to detection using those statistics and show where their differences and similarities lie. The large number of available detection statistics, along with their numerous parameters, makes formal comparisons complicated and sometimes produce ambiguous or uncertain conclusions. This document tries to establish a solid framework on which to base this analysis and then demonstrates how detection statistics can be grouped in distinct classes based on their decision surface's geometry. In doing so, possibly new detection statistics (or methods that could be used to build them) are exposed.

1.1 Context

Defence R&D Canada – Valcartier has been involved in the passive detection of gases and aerosols in the atmosphere, and liquids and solids on the ground, vehicules and struc-tures, since the mid-90's [1]. Long-wave (thermal) infrared hyperspectral measure-ments, typically acquired using a Fourier transform spectrometer, is the technique of choice to achieve this because of the unique spectral features found in most chemical species of interest. The technique was refined using the CATSI spectroradiometer [2], which was operated by scientists who, aided by analysis software such as GASEM [3], determined if any species of interest was present in a given measurement. In an effort to push the technique into the hands of the Canadian Forces (spearheaded during the CATSI EDM project [4]), algorithms were developed in order to automate the detection and identification of a wide range of substances.

An extensive survey of the literature [5] have shown multiple hyperspectral target detection formalisms, sometimes coming from different fields (acoustics, radar, etc.) and most often using different notations expressing the same general ideas differently. Comparing the performances of various approaches to hyperspectral detection is rendered tedious because of theses variations and depend entirely on the ability to properly implement the proposed algorithms. During the development of the iCATSI and MoDDIFS sensors [6], a new, generic code base for hyperspectral detection was developed[1], which higlighted similarities between many of the detection methods found in the literature.

[1] This will be the topic of a future report.

1.2 Scope

This report is a review of the most common detection statistics in usage for hyperspectral target detection, where the geometry of the associated decision surfaces is explored. It will be shown that most detection statistics can be grouped in two or three broad categories that share basic properties. The framework developed to perform this analysis is then be applied to further explore extensions and possibly altogether new detection statistics. However this report will not compare the actual performance of any detection statistic using either real or simulated data. Comparison will be restricted to the geometry of decision surfaces associated with each statistic; however this can be used to steer performance comparison efforts in new directions.

1.3 Structure of document

In order to enable progress in subsequent sections, Section 2 will first establish notation and review basic concepts such as hypothesis testing and associated performance metrics, generalized likelihood ratio tests and background clutter models. Section 3 will present, in increasing complexity, the detection statistic categorized by the shape of their decision surfaces, begining with planes and up to hyperboloids. The following section (Section 4) explores avenues for creating novel detection statistics by extending the concepts developed previously: combinations of existing decision surfaces, creation of arbitrary axially-symmetric decision surfaces and usage of different background clutter model transformations. Concluding remarks are regrouped in Section 5.

2 Basics

2.1 Notation

We begin by establishing the mathematical notation that will be used throughout this report. Generally speaking, vectors and matrix will be represented by bold, upright characters such as $\mathbf{x}$ (matrices will use capital letters such as $\mathbf{A}$) while scalars will use normal case, italic characters such as a. Idempotent matrices (projectors) will be noted with a $\mathbf{P}$, but sporting an index noting the subspace into which it projects. If they also feature the $\perp$ symbol, then they project into the subspace perpendicular to the one denoted by the index.

The measurement under test will be denoted $\mathbf{x}$. Typically $\mathbf{x}$ will be composed of a background clutter component $\mathbf{b}$, to which is added sensor noise $\mathbf{n}$ and possibly a contribution containing the target signature $\mathbf{s}$. All those vectors are column vectors of size $k \times 1$, where k is the number of spectral elements. In many detection statistic formulations, the measurement under test and the target signatures are linearly transformed for various purposes. For now we note this generic transformation $\mathbf{A}$. We will note the transformed measurement and target signature $\mathbf{Ax} = \mathbf{y}$ and $\mathbf{As} = \mathbf{t}$, respectively.

2.2 Hypothesis testing

Detection of a particular material in a hyperspectral measurement is, most of the time, construed as a binary hypothesis testing problem. The detection procedure aims to choose, between two hypothesis, which one explains the measurement the best. The first hypothesis, often designated the "null" hypothesis and denoted H_0, asserts that the target is not present in the measurement; while the second, designated H_1, asserts that the target is present. In mathematical terms, each hypothesis is associated to a different measurement model. For example, a very common set of measurement models (adequate for the case of translucent or sub-pixel targets in the thermal infrared) is give by:

$$H_0 \,:\, \mathbf{x} = \mathbf{b} + \mathbf{n}\,, \tag{1a}$$

$$H_1 \,:\, \mathbf{x} = \mathbf{b} + a\mathbf{s} + \mathbf{n}\,. \tag{1b}$$

Under hypothesis H_0, the measurement $\mathbf{x}$ is composed of instrument noise $\mathbf{n}$ added to background clutter $\mathbf{b}$. Under hypothesis H_1, the term $a\mathbf{s}$ is added, which is the target signature $\mathbf{s}$ times a strength (or abundance) factor a. In most realistic scenarios, those terms are not known precisely. The instrument noise term $\mathbf{n}$ for example is generally characterized by its covariance matrix $\mathbf{\Sigma_n}$, which is often assumed to be diagonal, if not outright proportional to the identity matrix (i.e. $\mathbf{\Sigma_n} = \sigma_{\mathbf{n}}^2\mathbf{I}$). Likewise, the background clutter term $\mathbf{b}$ is not known exactly. Some models assume that

its statistics (such as covariance matrix $\mathbf{\Sigma_b}$) are known a priori; others need them estimated from background clutter samples. Most models assume $\mathbf{s}$ to be known a priori (which may or may not be a reasonable assumption depending on the target and physics in play), and a to be unknown. In the latter case, deciding between H_0 and H_1 is thus the same as deciding whether $a = 0$ or $a \neq 0$. Such an additive target model can be useful even if it does not faithfully represent the physics of how the target's signature appear in the measurement. For example, radiance coming from gases will typically add a density-contrast product term of the form:

$$\mathrm{DCP} = \Delta \cdot (1 - \mathrm{e}^{-C\mathbf{s}}), \tag{2}$$

which approaches 0 whenever C (a concentration parameter) is small, and saturates to Δ (a contrast parameter) when C becomes very large [7]. This term is not linear in $\mathbf{s}$, but when $C\mathbf{s} \ll 1$, it can be approximated by $\mathrm{DCP} \approx a\mathbf{s}$ with $a = \Delta C$ (hence the name "density-contrast product").

Hypothesis testing thus involves a binary function of the measurement $\mathbf{x}$, $t(\mathbf{x})$, whose value can be either 0 (in which case hypothesis H_0 is favored) or 1 (H_1 favored). Most often, the binary nature of this function is realized by comparing a continuous *detection statistic* $d(\mathbf{x})$ to a fixed threshold τ:

$$\begin{matrix} \\ d(\mathbf{x}) \\ \\ \end{matrix} \begin{matrix} > \\ = \tau \\ < \end{matrix} \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases} . \tag{3}$$

While it is possible to devise ways to perform target detection without resorting to a comparison between a continuous detection statistic and a threshold[2], it is safe to say that this is the most common and studied way of posing and solving the binary testing problem. It is useful to build a picture of the geometry of $d(\mathbf{x})$ (the continuous detection statistic) and $t(\mathbf{x})$ (the actual binary test). Different values of $\mathbf{x}$ are points in a k-dimensional space; $d(\mathbf{x})$ assigns a scalar value for each possible $\mathbf{x}$, creating a scalar field in the measurement space. $t(\mathbf{x})$ assigns either 0 or 1 to the each possible $\mathbf{x}$, by comparing the corresponding scalar value to a threshold. This creates a $(k-1)$–dimensional surface in measurement space separating measurements assigned to hypothesis H_0 from those assigned to H_1. This surface is called the *decision surface* or the *decision boundary*, and specifying its exact shape and position is entirely equivalent to specifying the test $t(\mathbf{x})$. One obvious consequence of this is that hypothesis tests sharing the same decision surface are in fact the same test.

2.3 Performance metrics

Choosing the best hypothesis test $t(\mathbf{x})$ (meaning here the best combination of a detection statistic $d(\mathbf{x})$ and a threshold τ) is a hard problem in general [8]. It depends

[2]For example, classification via clustering is not based on a threshold comparison.

upon the actual measurement model, which may or may not be accurate in the first place. It also depends on the acceptable false alarm rate. Every hypothesis test has a probability of detection p_D (that a measurement containing the target is correctly assigned to H_1) and a probability of false alarm p_F (that a measurement not containing the target is incorrectly assigned to H_1), which both depend upon the chosen threshold τ. Assuming a detection statistics such that $d(\mathbf{x}) \geq 0$, setting the threshold to zero means assigning H_1 to every measurement. This implies $p_D = p_F = 1$ in all conditions. Conversely, choosing an arbitrarily large threshold is akin to assigning H_0 to every measurement, implying $p_D = p_F = 0$. Values of τ between these extremes result in tuples (p_F, p_D) that describe a curve between $(0, 0)$ and $(1, 1)$. This curve is called the *receiver operating characteristics* curve, or *ROC* curve. It is interesting to note that choosing between H_0 and H_1 at random — which is a valid, hypothesis test — produces a ROC curve that is a straight line between $(0, 0)$ and $(1, 1)$, the exact value of p_D and p_F being set by the probability of choosing either hypothesis[3]. This constitutes the worst possible ROC curve, as this hypothesis test does not even take into account the information present in the measurement. Any ROC curve presenting lower probabilities of detection could be interpreted differently by reversing the decision and thus obtaining better-than-random results. It can be shown that ROC curves for threshold-based hypothesis tests must be monotonically increasing, and that ROC curves that are not convex are not optimal [9, 10].

Comparing probabilities of detection (p_D) alone is meaningless when trying to design the best hypothesis test, as it says nothing about the associated false alarm rate, which we wish simultaneously as low as possible. Any comparison of p_D must be done at the same p_F to be meaningful; and some values of p_F (typically very low ones) generally present a greater interest than others. Comparing ROC curves is thus the preferred way to assess the performance of detection tests. One way to reduce this comparison to a single figure is to quote the area below the ROC curve, or variants of this metric in which lower values of p_F carry more weight.

While ROC curves are useful to compare different detection statistics, they do not answer two important questions:

1. Is there, for a given measurement model and associated statistical distributions, a detection statistics that offers an even better performances; and
2. How to find or build it if it is indeed the case.

First, "better performances" must be defined. A detection test that offer the best p_D for all p_F and for all values of the parameters on which $\mathbf{x}$ depends is called the *universally most powerful* (or *UMP*) test. An UMP test for a given detection problem is not guaranteed to exist. However, for a given p_F and given values of all the

[3]This can be implemented as a threshold-based test by randomly generating values between 0 and 1 and using the desired probability of choosing H_0 as the threshold.

parameters on which $\mathbf{x}$ depends, the *locally most powerful* (or *LMP*) test is guaranteed to exist. Finding the UMP or LMP can be straightforward or mathematically intractable depending on the chosen measurement model and statistical distributions associated with its parameters [11].

Other performance metrics can also be sought. For example, some detection statistics provide the same false alarm rate regardless of the values of the parameters affecting the measurement model $\mathbf{x}$ (in particular, noise power $\sigma^2_{\mathbf{n}}$). This property is called *constant false alarm rate*, or *CFAR*. CFAR tests greatly simplify choosing an appropriate threshold, as it does not have to be adjusted whenever measurement conditions change. In the absence of any certitude regarding the optimality of a detection statistics, a statistic which is proven to provide CFAR is quite appealing. However such proofs depend again on the specific measurement model and associated statistical distributions, and their validity in real-world scenarios may be questionable.

2.4 Generalized Likelihood Ratio Tests

In this context, the Neyman-Pearson lemma [11, 12] is an island of certitude upon which test designers often rely. It basically states that the *likelihood ratio test*, or *LRT*, is LMP for the test's size[4] (and may also be UMP). The LRT consists in comparing the *likelihood ratio* (or *likelihood function*) $\Lambda(\mathbf{x})$ to a threshold. The likelihood ratio is the ratio of the conditional probabilities that H_1 and H_0 be true, given the measurement under test:

$$\Lambda(\mathbf{x}) = \frac{p\left(\mathbf{x}|H_1\right)}{p\left(\mathbf{x}|H_0\right)}\,. \tag{4}$$

However, the LRT requires knowledge of the exact values of all parameters upon which $\mathbf{x}$ depends (including, for example, noise power $\sigma^2_{\mathbf{n}}$).

When this knowledge is unavailable, estimating the value of the unknown parameters from one or many measurements using *maximum likelihood estimates* (*MLE*s) results in the *generalized* LRT, or *GLRT*. While the GLRT does not possess the same general optimality properties as the LRT, it provides a solid framework on which to build detection statistics given a measurement model and assumed statistical distributions. In fact, we previously showed that many popular detection statistics are indeed GLRTs for specific measurement models [5]; furthermore, it may be possible in some cases to deduce for which measurement model a particular detection statistic is a GLRT.

Readers interested in a more detailed discussion of hypothesis testing, ROC curves, constant false alarm rates and detection and estimation in general are invited to read

[4]The size α of a statistical test is the maximum probability of incorrectly rejecting the null hypothesis — which is the maximum allowable false alarm probability in the context of this discussion.

the works of Kay and Poor [12, 11] and our previous report on the topic of detection algorithms [5].

2.5 Background models

We conclude this section with a note about background clutter models. This work will deal with three models for modeling background clutter $\mathbf{b}$; other models of course exist but to our knowledge, our selection includes most popular detection statistics. The first model consists simply in assuming that there is no clutter, i.e.

$$\mathbf{b} = \mathbf{0}\,, \tag{5}$$

reducing the additive measurement model to $\mathbf{x} = a\mathbf{s} + \mathbf{n}$. The second model is by far the most common and assumes the clutter to be completely characterized by first- and second-order statistics (namely the clutter mean, $\mathbf{b}_0$, and covariance, $\mathbf{\Sigma_b}$):

$$\mathbf{b} \sim \mathcal{N}(\mathbf{b}_0, \mathbf{\Sigma_b})\,. \tag{6}$$

We will refer to this model as the *statistical background model*. The third model assumes that clutter $\mathbf{b}$ is a linear combination of a limited set of orthogonal components, forming the basis of a *background subspace*. According to this *background subspace model*,

$$\mathbf{b} = \mathbf{B}\mathbf{c}\,, \tag{7}$$

where the columns of $\mathbf{B}$ generate the background subspace and $\mathbf{c}$ is a vector of coefficients that are typically unknown. $\mathbf{B}$ is generally estimated (typically using singular value decomposition, eigenvalue decomposition or linear unmixing techniques) based on samples included in a training set, as are $\mathbf{b}_0$ and $\mathbf{\Sigma_b}$.

Parameters $\mathbf{b}_0$, $\mathbf{\Sigma_b}$ and $\mathbf{B}$ appear in detection statistics when transforming the measurement and target signatures so as to lessen the influence of the background clutter. For example, the matrix square root of the inverse covariance matrix, $\mathbf{\Sigma_b}^{-\frac{1}{2}}$, is commonly called the *whitening* transform because when applied to the measurement $\mathbf{x}$, it reshapes the background clutter term to white noise (so that its covariance matrix $\mathbf{\Sigma_b}$ is transformed into the identity matrix $\mathbf{I}$). Another commonly used transformation is:

$$\mathbf{P}_{\mathbf{B}}^{\perp} = \mathbf{I} - \mathbf{B}\left(\mathbf{B}^{\top}\mathbf{B}\right)^{-1}\mathbf{B}^{\top}, \tag{8}$$

which is an orthogonal projector outside of the background subspace $\mathbf{B}$. In order to remain general, the arbitrary transform $\mathbf{A}$ will be used in this work to denote either of those transformations (or even the identity matrix in the case where clutter is neglected). Again, without loss of generality, and in order to simplify notation,

we will use $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $\mathbf{t} = \mathbf{A}\mathbf{s}$ to denote the transformed measurement and target signature.

Note that when using either the statistical or background subspace model, there is a distinction to be made between the real clutter covariance matrix or projector (if known or assumed a priori) and an estimate thereof computed from a training set containing a number of clutter samples. When an estimate from a training set is used, the detection statistics are often referred to as being *adaptive*.

3 Taxonomy of decision surfaces

As stated in section 2, a binary hypothesis test is entirely defined by its decision surface. In this section, popular detection statistics are sorted into increasingly complex categories according to the shape of their decision surfaces. In the following subsections, each decision surface will be described in the transformed measurement ($\mathbf{y} = \mathbf{Ax}$) space first, as it will be common to (and thus share basic properties with) many detection statistics using different transformations $\mathbf{A}$. The effect of transformation $\mathbf{A}$ on the decision surface in the real measurement space will be discussed afterwards.

3.1 Planar decision surfaces

Planes are arguably the simplest decision surfaces. A plane in the (transformed) measurement space can be defined by specifying the norm (i.e. length) of the projection of $\mathbf{y}$ onto the the plane's normal vector, which is generally chosen to be the transformed target signature $\mathbf{t} = \mathbf{As}$:

$$\|\mathbf{P_t y}\| = \tau\,. \tag{9}$$

In other words, the set of transformed measurements $\mathbf{y}$ for which the norm of their projection unto $\mathbf{t}$ is the same will constitute a plane (see Figure 1). Any monotonic (order-preserving) transformation of $\|\mathbf{P_t y}\|$ and τ constitutes the same constraint and thus will yield the same specific plane. This implies that a large number of detection statistics can be reduced to a test equivalent to the one of Equation (9). Noting first that:

$$\mathbf{P_t} = \mathbf{t}\left(\mathbf{t}^\top\mathbf{t}\right)^{-1}\mathbf{t}^\top, \tag{10}$$

by definition of the projection transformation (see [13] and [5]), the following test also generates a planar decision surface, as it is simply the quantity of Equation (9) squared:

$$\|\mathbf{P_t y}\|^2 = \frac{\left(\mathbf{t}^\top\mathbf{y}\right)^2}{\mathbf{t}^\top\mathbf{t}} \begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases}. \tag{11}$$

This is also true of any test of the form:

$$c\left(\mathbf{t}^\top\mathbf{y}\right)^n \begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases}, \tag{12}$$

since multiplication by a scalar and exponentiation constitute monotonic transformations which do not change the constraint specifying the basic shape of the decision surface. If the same monotonic transformation is also applied on the threshold, then the position of the plane along its normal $\mathbf{t}$ will also be preserved. Note that in Equation (11), the term $\mathbf{t}^\top\mathbf{t}$ on the denominator is not dependent upon the measurement $\mathbf{y}$; and since $\mathbf{t}$ is assumed to be known, or at least is not estimated from the measurement, it reduces to a scalar.

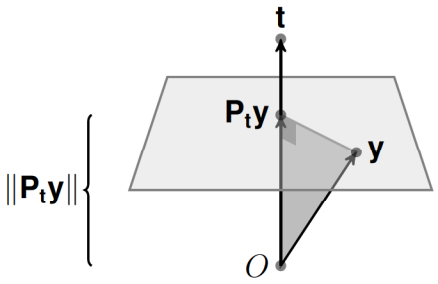

Figure 1: Planar decision surfaces arise from specifying $\|\mathbf{P_t y}\|$.

The linear transformation $\mathbf{A}$ can only scale and rotate the decision boundary. $\mathbf{A}$ is typically applied in order to orient the plane as to better separate the background clutter from measurements containing the target. However, the decision boundary remains a plane whether or not $\mathbf{A} = \mathbf{I}$. An important implication of this is that by choosing the right threshold τ and transformation $\mathbf{A}$, all the tests having a planar decision surface can thus be made to produce the exact same output for any $\mathbf{x}$, and thus have the same ROC and overall performances. In essence, all tests having planar decision surfaces are in fact the same, up to position and orientation parameters.

Detection statistics that produce planar decision surfaces include simple correlators that can be expressed as $\mathbf{t}^\top\mathbf{y}$. When background clutter is neglected ($\mathbf{A} = \mathbf{I}$) this yields the *matched filter* $\mathbf{s}^\top\mathbf{x}$, which is often normalized (without affecting its performances) to yield the quantity:

$$\hat{a} = \frac{\mathbf{s}^\top\mathbf{x}}{\mathbf{s}^\top\mathbf{s}}\,. \tag{13}$$

This quantity is the MLE estimator $\hat{a}$ for the abundance factor a, given the additive measurement model of Equation (1). Note that assuming $\mathbf{s}$ known, the denominator is simply a scalar.

When a statistical clutter model is used and $\mathbf{A} = \mathbf{\Sigma_b}^{-\frac{1}{2}}$, the resulting detection statistics, $\mathbf{s}^\top\mathbf{\Sigma_b}^{-1}\mathbf{x}$, is often called *clutter-matched filter* or *Fischer's linear discriminant* [12, 11], and is also often normalized to produce the corresponding abundance esti-

mator:

$$\hat{a} = \frac{\mathbf{s}^\top \boldsymbol{\Sigma}_\mathbf{b}^{-1} \mathbf{x}}{\mathbf{s}^\top \boldsymbol{\Sigma}_\mathbf{b}^{-1} \mathbf{s}} \,. \tag{14}$$

Using the background subspace model produces the analogous statistic $\mathbf{s}^\top \mathbf{P}_\mathbf{B}^\perp \mathbf{s}$ and abundance estimator:

$$\hat{a} = \frac{\mathbf{s}^\top \mathbf{P}_\mathbf{B}^\perp \mathbf{x}}{\mathbf{s}^\top \mathbf{P}_\mathbf{B}^\perp \mathbf{s}} \,. \tag{15}$$

An interesting case is the *constrained energy minimization* (CEM) detector [14, 15, 16], which has the same form but uses $\mathbf{A} = \mathbf{R}_\mathbf{b}^{-\frac{1}{2}}$, where $\mathbf{R}_\mathbf{b}$ is an estimate of the correlation matrix instead of the covariance matrix. While such a switch changes its position and orientation in measurement space, the decision surface generated by using this detection statistic remains a plane.

Squaring the numerator and using the form of Equation (11), the resulting detection statistics become energy estimators because they are the MLE for $\sigma_\mathbf{n}^2$, given the additive measurement model. When the covariance matrix is known a priori, the *matched subspace detector* (MSD) is thus produced; when an estimate from a training set is used instead, the detection statistics is called *adaptive matched filter* (AMF) or sometimes *adaptive subspace detector* (ASD) [17, 18, 19, 14].

Since variations in $\sigma_\mathbf{n}^2$ will always change the proportion of measurements $\mathbf{x}$ that are on either side of the decision boundary for noise distributions that extend radially around a mean, detection statistics that produce planar decision surfaces cannot possess the CFAR property under this assumption. Indeed, a larger value for $\sigma_\mathbf{n}^2$, which would swell the distribution of $\mathbf{x}$ under hypothesis H_0, would have to be compensated with a larger threshold τ in order to keep the same false alarm rate.[5]

Table 1 presents a summary of various detection statistics that generate planar decision surfaces.

3.2 Cylindrical and closed decision surfaces

It is possible to use the detection statistics described in subsection 3.1 to decide whether any combination of two or more target signatures $(a_0\mathbf{s}_0 + a_1\mathbf{s}_1 + \cdots + a_n\mathbf{s}_n)$

[5]This is true for radial distributions of $\mathbf{x}$ unless the planar decision surface includes the measurement mean $\mathbf{b}_0$ under H_0 and thus splits the distribution under H_0 exactly in half, yielding a constant 50 % false alarm rate when $\sigma_\mathbf{n}^2$ changes. This would happen when setting the threshold to zero and ignoring transformed measurements with the opposite sign to the signature; this situation is thus a trivial case.

Table 1: Examples of detection statistics with planar decision boundaries.

Clutter model	Simple correlators $\mathbf{t}^\top\mathbf{y}$	Abundance estimators $\frac{\mathbf{t}^\top\mathbf{y}}{\mathbf{t}^\top\mathbf{t}}$	Energy estimators $\frac{(\mathbf{t}^\top\mathbf{y})^2}{\mathbf{t}^\top\mathbf{t}}$
$\mathbf{A}=\mathbf{I}$	$\mathbf{s}^\top\mathbf{x}$	$\frac{\mathbf{s}^\top\mathbf{x}}{\mathbf{s}^\top\mathbf{s}}$	$\frac{(\mathbf{s}^\top\mathbf{x})^2}{\mathbf{s}^\top\mathbf{s}}$
$\mathbf{A}=\Sigma^{-\frac{1}{2}}$	$\mathbf{s}^\top\Sigma^{-1}\mathbf{x}$	$\frac{\mathbf{s}^\top\Sigma^{-1}\mathbf{x}}{\mathbf{s}^\top\Sigma^{-1}\mathbf{s}}$	$\frac{(\mathbf{s}^\top\Sigma^{-1}\mathbf{x})^2}{\mathbf{s}^\top\Sigma^{-1}\mathbf{s}}$
$\mathbf{A}=\mathbf{P}_\mathbf{B}^\perp$	$\mathbf{s}^\top\mathbf{P}_\mathbf{B}^\perp\mathbf{x}$	$\frac{\mathbf{s}^\top\mathbf{P}_\mathbf{B}^\perp\mathbf{x}}{\mathbf{s}^\top\mathbf{P}_\mathbf{B}^\perp\mathbf{s}}$	$\frac{(\mathbf{s}^\top\mathbf{P}_\mathbf{B}^\perp\mathbf{x})^2}{\mathbf{s}^\top\mathbf{P}_\mathbf{B}^\perp\mathbf{s}}$

is present using the same additive model as previously used. In that situation, the projector $\mathbf{P_t}$ into the (transformed) target signature is simply replaced by $\mathbf{P_T}$, the projector into the target subspace $\mathbf{T} = [\mathbf{t}_0\,\mathbf{t}_1\cdots\mathbf{t}_n]$. For this reason, such detection statistics are generally called *subspace detectors* [20].

The condition $\|\mathbf{P_T y}\| = \tau$ generates a hyper-cylinder in the transformed measurement space, curving the planar decision in such a waw that it becomes perpendicular to all targets (see Figure 2). The linear transformation $\mathbf{A}$ skews (and rotates) this cylinder, giving it an elliptic cross-section in the original measurement space.

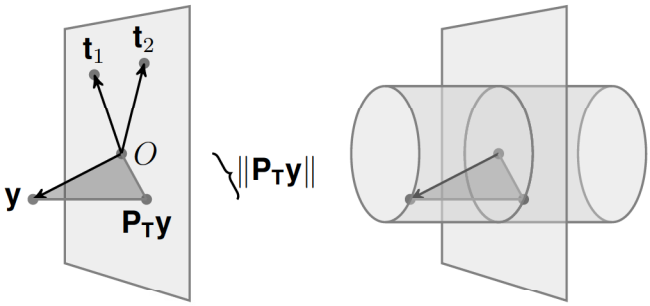

Figure 2: Hyper-cylinders are formed when specifying $\|\mathbf{P_T y}\|$.

As an increasing number of targets are added to the target subspace $\mathbf{T}$, more dimensions are folded into the hyper-cylinder. An interesting phenomenon arises when $\mathbf{P_T} = \mathbf{I}$, meaning that all possible targets are to be detected. This completely wraps the decision boundary around the measurement mean under H_0 so that it becomes perpendicular to all directions. The decision surface is then a closed sphere

in transformed space (see Figure 3), producing a category of detection statistics called *anomaly detectors*, since they are used when no a priori information about the target signature is known. For example, using the statistical background clutter model ($\mathbf{A} = \mathbf{\Sigma_b}^{-\frac{1}{2}}$) with $\mathbf{P_T} = \mathbf{I}$, the energy estimator form of the binary test reduces to:

$$\|\mathbf{y}\|^2 = \mathbf{x}^\top \mathbf{\Sigma_b}^{-1}\mathbf{x} \begin{matrix} > \\ = \tau \\ < \end{matrix} \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases}, \tag{16}$$

which is the *Reed-Xiaoli* (or RX) anomaly detector. This detection statistic works by imposing a threshold on the *Mahalanobis distance* from the measurement mean under H_0, which is $\|\mathbf{y}\|$ assuming mean-removed measurements. The decision surface in the original measurement space is an ellipsoid shell, choosing H_1 over H_0 for all $\mathbf{x}$ outside.

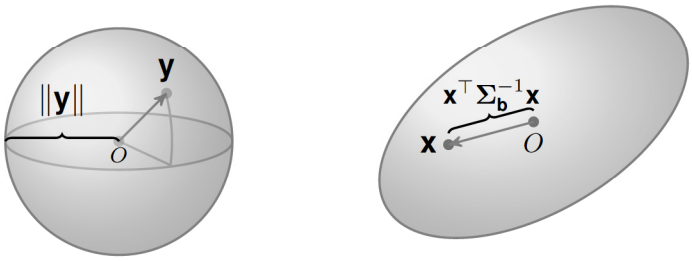

Figure 3: Spheres are formed when specifying $\|\mathbf{y}\|$, which is the case when $\mathbf{P_T} = \mathbf{I}$. In the non-transformed measurement space, this translates into ellipsoids as the Mahalanobis distance is fixed.

When $\mathbf{A} = \mathbf{I}$, the corresponding anomaly detector is simply $\mathbf{x}^\top\mathbf{x} = \|\mathbf{x}\|^2$: any measurement with a norm larger than the threshold (typically some multiple of the noise power) makes the test choose H_1. Using the background subspace model, the corresponding anomaly detector would be $\mathbf{x}^\top\mathbf{P_B^\perp}\mathbf{x}$, but it does not seem to be commonly used as it was not found in the literature.

3.3 Conical decision surfaces

A way to turn a detection statistic with a planar decision surface into a CFAR statistic is to dynamically adjust the threshold τ. This can be accomplished by making τ proportional to the noise power $\sigma_\mathbf{n}^2$. However this is only useful when $\sigma_\mathbf{n}^2$ is variable, in which case it is generally unknown. The MLE estimate for $\sigma_\mathbf{n}^2$ is $\mathbf{y}^\top\mathbf{y}$ (which is true for all background models — meaning any transformation $\mathbf{A}$ — considered under the additive model of Equation (1)). Thus, the decision surface specified by

the constraint:

$$\|\mathbf{P_t y}\|^2 = \mathbf{y}^\top \mathbf{P_t y} = \tau \left(\mathbf{y}^\top \mathbf{y}\right) \tag{17}$$

can be seen as a plane which distance to the origin[6] depends upon the noise power estimate. However, rearranging terms, this constraint can be written as:

$$\frac{\mathbf{y}^\top \mathbf{P_t y}}{\mathbf{y}^\top \mathbf{y}} = \tau \,. \tag{18}$$

The left hand side in the above equation is nothing else than the squared cosine of the angle α between $\mathbf{y}$ and $\mathbf{t}$. A threshold test on this value generates a decision surface that is a cone around $\mathbf{t}$ (see Figure 4) and with its apex at the origin. Transformed measurements $\mathbf{y}$ having a large angular distance with $\mathbf{t}$ are classified as pertaining to H_0, while those inside the cone are deemed to correspond to H_1. It is to be noted that strictly speaking, the decision surface is a double cone. Testing for the sign of $\mathbf{P_t y}$ can be used to reject one of the cones if necessary (e.g. if the measurement model forbids negative values for the abundance factor a).

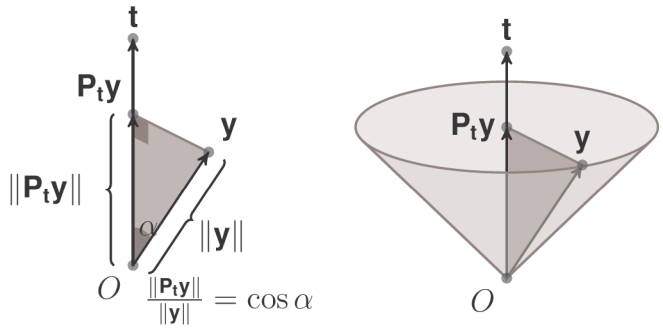

Figure 4: Cones are formed when specifying the angle α between $\mathbf{y}$ and $\mathbf{t}$.

The transformation $\mathbf{A} = \mathbf{\Sigma_b}^{-\frac{1}{2}}$ generates flattened cones having elliptical cross-sections in the original measurement space. The case of $\mathbf{A} = \mathbf{P_B^\perp}$ is a bit more peculiar however. The projector effectively suppresses any constraint in the directions spanned by $\mathbf{B}$, producing "hyper-wedges"' in the original measurement space along the dimensions spanned by $\mathbf{B}$ (see Figure 5). The situation is in fact slightly more complicated as the surface decision is closed (cone-like) around $\mathbf{s}$ along other dimensions (which is especially difficult to illustrate and was not attempted in Figure 5). Cones and wedges become very similar in the original measurement space if the cross-section of the cones are elongated ellipses.

The statistic of Equation (18) can be monotonically transformed in numerous ways, resulting in different expressions that nonetheless generate the exact same decision

[6]The origin corresponds to the measurement mean if this quantity has been removed from all tested measurements prior to processing.

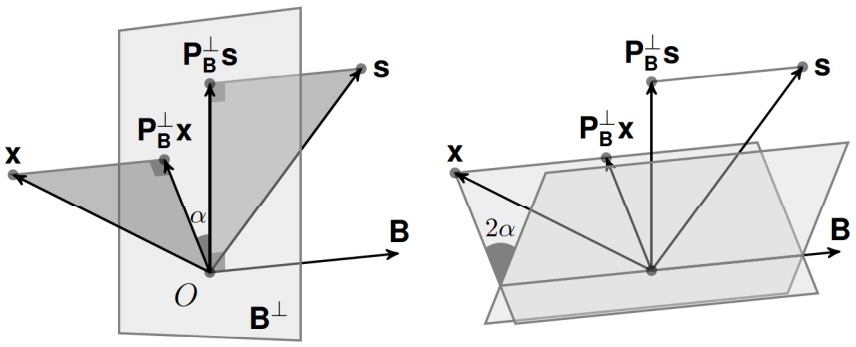

Figure 5: When $\mathbf{A} = \mathbf{P}_\mathbf{B}^\perp$, angular constraints in subspace $\mathbf{B}^\perp$ create wedge-shaped decision surfaces along $\mathbf{B}$ in measurement space. The decision surface actually wraps around $\mathbf{s}$ in directions orthogonal to $\mathbf{s}$ and $\mathbf{B}$, which cannot be easily illustrated in three dimensions.

surface and thus the same ROC. The most frequent forms of detection statistics generating conical decision boundaries are listed in Table 2. The F-form is so called because being a ratio of variance estimates, it constitutes an F-statistic, following a Fischer-Snedecor distribution if $\mathbf{y}$ follows a normal distribution. This ratio of variances can be interpreted as the ratio of the energy of the target contribution alone ($\mathbf{y}^\top\mathbf{P}_\mathbf{t}\mathbf{y} = \|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2$) to the energy of the measurement $\mathbf{y}$ excluding any target contribution ($\mathbf{y}^\top\mathbf{P}_\mathbf{t}^\perp\mathbf{y} = \left\|\mathbf{P}_\mathbf{t}^\perp\mathbf{y}\right\|^2$); thus it constitutes an estimate of the signal-to-noise ratio (or more precisely, the signal-to-clutter ratio). Detection statistics with conical decision surfaces are thus SNR estimators.

Table 2: Common forms and expressions for detection statistics with conical decision surfaces.

α-form	η-form	F-form	Expressions	Range
$\cos^2\alpha$	η	$\dfrac{F}{F+1}$	$\dfrac{\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2}{\|\mathbf{y}\|^2} = \dfrac{\mathbf{y}^\top\mathbf{P}_\mathbf{t}\mathbf{y}}{\mathbf{y}^\top\mathbf{y}}$	$[0,1]$
$\dfrac{1}{\tan^2\alpha}$	$\dfrac{\eta}{1-\eta}$	F	$\dfrac{\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2}{\left\|\mathbf{P}_\mathbf{t}^\perp\mathbf{y}\right\|^2} = \dfrac{\mathbf{y}^\top\mathbf{P}_\mathbf{t}\mathbf{y}}{\mathbf{y}^\top\mathbf{P}_\mathbf{t}^\perp\mathbf{y}}$	$[0,\infty[$
$\dfrac{1}{\sin^2\alpha}$	$\dfrac{1}{1-\eta}$	$F+1$	$\dfrac{\|\mathbf{y}\|^2}{\left\|\mathbf{P}_\mathbf{t}^\perp\mathbf{y}\right\|^2} = \dfrac{\mathbf{y}^\top\mathbf{y}}{\mathbf{y}^\top\mathbf{P}_\mathbf{t}^\perp\mathbf{y}}$	$[1,\infty[$

The sector-shaped decision boundary confers the CFAR property to the associated detection statistics: any radial swelling of the measurement distribution due to vari-

ations of noise power $\sigma_{\mathbf{n}}^2$ will not change the ratio of target-free measurements intercepted by the cones. There are many detection statistics in the literature that generate conical decision surfaces. The *spectral angle mapper* or *SAM* is perhaps the simplest of them[14]. It can be obtained by neglecting background clutter ($\mathbf{A} = \mathbf{I}$), which results in the following statistic:

$$\cos^2 \alpha = \frac{\left(\mathbf{s}^\top \mathbf{x}\right)^2}{\mathbf{x}^\top \mathbf{x}} \,. \tag{19}$$

When the statistical background model is employed, the resulting detection statistics is often called *adaptive coherence* (or *cosine*) *estimator* (ACE), also called *CFAR MSD* (when the covariance matrix is known a priori) or *CFAR ASD* (when an estimate from a training set is used instead) [17, 18, 14]. When $\mathbf{A} = \mathbf{P}_{\mathbf{B}}^{\perp}$, the corresponding detection statistic [20] is bizarrely often called *GLRT*, despite all other statistics cited in this paper being also generalized likelihood ratio tests for different assumptions about the measurement model [5]. This detector will be refered to as *Scharf's GLRT*.

Conical detection statistics can also be used for subspace detection, using $\mathbf{P}_{\mathbf{T}}$ instead of $\mathbf{P}_{\mathbf{t}}$; however the cone opens up as a wedge around the volume formed whenever a dimension is added to $\mathbf{T}$, and thus does not produce an anomaly detector when setting $\mathbf{P}_{\mathbf{T}} = \mathbf{I}$ (which results in a division by zero).

3.4 Hyperboloids and paraboloids

In the previous sections, we found detection statistics having decision surfaces in the shape of planes, cylinders, ellipsoids, cones and wedges. It is then only mildly surprising to discover hyperboloid detection surfaces. However they can arise in different situations. In the case of *Kelly's detector*, in which Kelly [21] used the joint probability distributions of both the measurement under test and the training set to build the likelihood ratio, the GLRT he obtained is the harmonic average of a planar and conical decision surface, producing an hyperboloid of two sheets (with $\mathbf{t}$ as the axis) instead of back-to-back cones. Kelly's detection statistic can be expressed as:

$$d_h(\mathbf{y}) = \frac{\mathbf{y}^\top \mathbf{P}_{\mathbf{t}} \mathbf{y}}{k + \mathbf{y}^\top \mathbf{y}} = \frac{1}{\frac{1}{d_p(\mathbf{y})} + \frac{1}{d_c(\mathbf{y})}} \,, \tag{20}$$

where $d_p(\mathbf{y})$ is a planar detection statistic and $d_c(\mathbf{y})$ a conical detection statistic given by:

$$d_p(\mathbf{y}) = \frac{\mathbf{y}^\top \mathbf{P}_{\mathbf{t}} \mathbf{y}}{k} \quad \text{and} \quad d_c(\mathbf{y}) = \frac{\mathbf{y}^\top \mathbf{P}_{\mathbf{t}} \mathbf{y}}{\mathbf{y}^\top \mathbf{y}} \,. \tag{21}$$

One can show that the statistic of Equation (20) generates an hyperbolic decision surface by comparing it to the equation of an hyperbola in the cartesian (x, y) plane:

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1\,, \tag{22}$$

and by associating the y-axis with the direction of the target signature in the transformed space, $\mathbf{t}$, and the x-axis with the subspace perpendicular to the target subspace. Projecting a transformed measurement $\mathbf{y}$ unto each subspace, the norm of the projection along x becomes $\left\|\mathbf{P}_\mathbf{t}^\perp \mathbf{y}\right\|$ while the norm of the projection along y becomes $\|\mathbf{P}_\mathbf{t}\mathbf{y}\|$ (see Figure 6). By noting that:

$$\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2 + \left\|\mathbf{P}_\mathbf{t}^\perp \mathbf{y}\right\|^2 = \|\mathbf{y}\|^2 \tag{23}$$

and making the following subtitutions into Equation (22):

$$y^2 \to \|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2\,, \tag{24a}$$

$$x^2 \to \left\|\mathbf{P}_\mathbf{t}^\perp \mathbf{y}\right\|^2 = \|\mathbf{y}\|^2 - \|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2\,, \tag{24b}$$

$$b^2 \to k\,, \tag{24c}$$

$$a^2 \to \frac{k\tau}{1-\tau}\,, \tag{24d}$$

one effectively obtains a constraint on the detection statistics $d_h(\mathbf{y})$ of Equation (20):

$$\begin{aligned}
\frac{y^2}{a^2} - \frac{x^2}{b^2} &= 1\,, \\
\frac{1-\tau}{k\tau}\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2 - \frac{\|\mathbf{y}\|^2}{k} + \frac{\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2}{k} &= 1\,, \\
\frac{1-\tau}{\tau}\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2 + \|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2 &= k + \|\mathbf{y}\|^2\,, \\
\frac{\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2}{\tau} &= k + \|\mathbf{y}\|^2\,, \\
\frac{\|\mathbf{P}_\mathbf{t}\mathbf{y}\|^2}{k + \|\mathbf{y}\|^2} = \frac{\mathbf{y}^\top \mathbf{P}_\mathbf{t}\mathbf{y}}{k + \mathbf{y}^\top\mathbf{y}} &= \tau\,.
\end{aligned} \tag{25}$$

In Kelly's formalism, k is the number of elements of the spectral vectors, but together with the threshold, this parameter controls the depth and opening of the hyperboloid. The slope m of the hyperbola's asymptote is entirely defined by the threshold τ:

$$m = \frac{a}{b} = \sqrt{\frac{\tau}{1-\tau}}\,, \tag{26}$$

and the hyperbola's minimum is located at

$$\|\mathbf{P_t y}\| = a = \sqrt{\frac{k\tau}{1-\tau}}\,. \tag{27}$$

Thus, the aperture of the hyperbola (but also the position of the minimum) is controlled by τ through m, and the position of the minimum can be adjusted by k once τ is fixed.

Kelly's detector presents a very interesting compromise between the non-CFAR planar decision boundary and the CFAR conical boundary. The second half of Kelly's original article is dedicated to showing that the statistic of Equation (20) has the CFAR property; but barring a different interpretation of the meaning of this property, it simply cannot be considering the shape of its decision surface. Interestingly, this decision surface becomes an ellipsoid (the RX anomaly detector) when $\mathbf{P_t} = \mathbf{I}$. When $\mathbf{A} = \mathbf{P_B^\perp}$, the detection statistic produces wedges with hyperbolic cross sections; to our knowledge, this has not been described in the literature, but is analogous to both Scharf's GLRT and Kelly's detector.

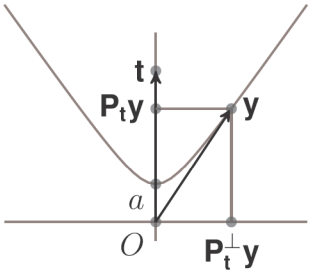

Figure 6: A constraint on Kelly's detection statistic matches the equation of an hyperbola, $\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$, in the plane defined by $x = \left\|\mathbf{P_t^\perp y}\right\|$ and $y = \|\mathbf{P_t y}\|$. The decision surface is thus an hyperboloid of revolution around $\mathbf{t}$.

Another way in which hyperboloids can arise is shown by Schaum [22], in which the *Joint Affine Matched Filter* emerges as a way to discriminate between background clutter and target distributions that lie along different spectral directions, but coincide at the origin when illumination drops[7]. In this scenario, the decision surface is an hyperboloid of one sheet, the asymptotes being half-way between the background and target subspaces.

Paraboloids of revolution appear in measurement models in which a distance metric is held constant between a point and a hyperplane, as in the *Affine Matched Filter* (AMF) [22, 23], which is used for a measurement model where the distribution of the measurement under hypothesis H_1 is a hyperplane whereas it is radially symmetric

[7]This is the case when the clutter is substituted by the target instead of being mixed with it, such as when the target is translucent or sub-pixel in size.

around the measurement mean under H_0. This could also generate a hyperbolic paraboloid (*Pringles*-shaped) decision surface depending on the particular geometry of the problem.

4 Expanding the family

The detection statistics presented in the previous section are described by expressions that often hide the simplicity of their associated decision surfaces. Once an understanding of the mechanism by which decision surfaces are generated is reached, building detection statistics having different decision surfaces becomes easier. This section explores ways to use the mathematical framework of the previous sections to build new detection statitics. This work is not meant to evaluate their performance or compare them, but simply to show that it is possible to precisely craft new decision surfaces if the need arises. We present in this section three vehicles for tweaking the shape of decision surfaces.

4.1 Combinations of detection statistics

Logical operations between the previously visited detection statistics may in turn produce interesting decision surfaces. For example, it might be tempting to limit the false alarm rate of a detection system by raising an alarm only when multiple detection tests have reached the same conclusion. This operation generates the geometrical intersection (logical *and*) of the various decision surfaces involved. For a given measurement $\mathbf{x}$, the more stringent combination of detection statistics and threshold sets the "active" decision surface.

For example, the intersection of a planar an conical decision surface results in a truncated cone. If a measurement is found inside the tip of the cone but under the plane, H_0 is chosen because while it is angularly close enough to the target signature for an alarm to be raised considering the conical statistic alone, it is not far enough from $\mathbf{b}_0$ to warrant an alarm according the the planar statistic. Likewise, a measurement found above the plane but outside of the cone will not trigger an alarm because while far enough from $\mathbf{b}_0$ according to the planar statistic, it is angularly too far from the target signature to warrant an alarm according to the conical statistic. The resulting decision surface is in fact very close to the hyperbolic surface described in subsection 3.4, and as such, may not present a significant performance difference[8].

The *Continuum Fusion* framework proposed by Schaum[24] can be used to build decision statistics by performing the union of an infinite number of decision surfaces generated by continuously varying a parameter of interest in the expression a generating statistic.

The union (logical *or*) of a large number of planar decision surfaces having different normals can also technically be used to create a facetted global decision surface that

[8]Notably, as is the case with the hyperbolic statistic, the truncated cone cannot possess the CFAR property for radial clutter distributions

could have an arbitrary shape. The difficulty of this proposition mostly lies in specifying the exact shape of the surface (i.e. the normals and thresholds), and may be better suited for anomaly detectors or "broad-spectrum" detectors that makes use of a large number of possible target signatures.

4.2 Arbitrary axially-symmetric surfaces

The formalism developed in subsection 3.4 to analyze Kelly's detector makes it easy to build a detection statistic associated with decision surfaces having an axial symmetry around $\mathbf{t}$, but otherwise an arbitrary profile along this axis. The ideas is to specify $y = \|\mathbf{P_t y}\| = \left(\mathbf{y}^\top \mathbf{P_t y}\right)^{\frac{1}{2}}$ as a function of $x = \left\|\mathbf{P_t^\perp y}\right\| = \left(\mathbf{y}^\top \mathbf{P_t^\perp y}\right)^{\frac{1}{2}}$. Since the resulting decision surface is axially-symmetric, its profile only needs to be specified in the first quadrant. For example, this technique can be used to specify the conical statistics discussed in subsection 3.3. The double-cone's profile in the first quadrant is simply given by the equation:

$$y = mx\,, \tag{28}$$

where m is the slope of the cone's profile. The cone is completely open with $m = 0$ and closes as $m \to \infty$. Using the substitutions discussed above, one obtains:

$$\begin{aligned} \|\mathbf{P_t y}\| &= m \left\|\mathbf{P_t^\perp y}\right\|\,, \\ \frac{\|\mathbf{P_t y}\|^2}{\left\|\mathbf{P_t^\perp y}\right\|^2} &= \frac{\mathbf{y}^\top \mathbf{P_t y}}{\mathbf{y}^\top \mathbf{P_t^\perp y}} = m^2\,, \end{aligned} \tag{29}$$

which is the F-form of the conical statistics as listed in Table 2. Thus, a detection test using this statisic is a threshold comparion on the square of the slope of the cone. By noting that $\|\mathbf{y}\|^2 = \|\mathbf{P_t y}\|^2 + \left\|\mathbf{P_t^\perp y}\right\|^2$, that form can be changed to:

$$\begin{aligned} \|\mathbf{P_t y}\|^2 &= m^2 \left(\|\mathbf{y}\|^2 - \|\mathbf{P_t y}\|^2\right)\,, \\ \frac{\|\mathbf{P_t y}\|^2}{\|\mathbf{y}\|^2} &= \frac{\mathbf{y}^\top \mathbf{P_t y}}{\mathbf{y}^\top \mathbf{y}} = \frac{1}{1+m^2}\,, \end{aligned} \tag{30}$$

which is the $\cos^2\alpha$-form of Equation (18). The corresponding detection test is a then a threshold comparison on the right term, which has an inverse ordering compared to m^2: it takes the value 1 when the spectral angle is zero or π and is zero-valued when the angle is $\pi/2$.

We now proceed to use the same technique to build axially-symmetric decision surfaces along $\mathbf{t}$ with profiles that were not described in Section 3.

4.2.1 Double cones with offset tips

Creating a decision surface in the shape of back-to-back cones whose tips are offset (i.e. not set at the origin) becomes a matter of starting with the profile $y = mx + b$, where b is the distance from the origin to the tip of the cones. The presence of the b term prevents the creation of an elegant expression using pure quadratic forms (such as $\mathbf{y}^\top \mathbf{P_t y}$); instead the square root of the quadratic form must be used:

$$\begin{aligned} \|\mathbf{P_t y}\| &= m \left\|\mathbf{P_t^\perp y}\right\| + b\,, \\ \frac{\left(\mathbf{y}^\top \mathbf{P_t y}\right)^{\frac{1}{2}}}{m \left(\mathbf{y}^\top \mathbf{P_t^\perp y}\right)^{\frac{1}{2}} + b} &= 1\,, \\ \frac{\left(\mathbf{y}^\top \mathbf{P_t y}\right)^{\frac{1}{2}}}{\left(\mathbf{y}^\top \mathbf{P_t^\perp y}\right)^{\frac{1}{2}} + \frac{b}{m}} &= m\,, \end{aligned} \tag{31}$$

The offset cone cannot possess the CFAR property but is another way to compromise between a planar and a conical decision surface.

4.2.2 Paraboloid

It is also easy to create a decision surface shaped as a paraboloid of revolution having $\mathbf{t}$ as its axis. By using the equation for a vertical parabola centered on the y-axis:

$$y = ax^2 + c \tag{32}$$

and using the substitutions:

$$y \to \|\mathbf{P_t y}\|\,, \tag{33a}$$

$$x^2 \to \left\|\mathbf{P_t^\perp y}\right\|^2\,, \tag{33b}$$

$$a \to \tau\,, \tag{33c}$$

$$c \to k\tau\,, \tag{33d}$$

one obtains the detection test:

$$d_{\text{offset cone}} = \frac{\|\mathbf{P_t y}\|}{k + \left\|\mathbf{P_t^\perp y}\right\|^2} = \frac{\left(\mathbf{y}^\top \mathbf{P_t y}\right)^{1/2}}{k + \mathbf{y}^\top \mathbf{P_t^\perp y}} \begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases}. \tag{34}$$

This is another compromise between a planar and a conical decision surface; however, contrary to the hyperboloid of Kelly's detector presented at subsection 3.4, the paraboloid is not bounded by asymptotes and thus opens wider around $\mathbf{t}$.

4.2.3 Square root profile

One valid critism of conical decision surfaces is that it fails to prevent false alarms that corresponds to target-free measurements found at the tip of the cones, i.e. close to $\mathbf{b}_0$, where the distribution under H_0 typically shows a high probability density. This is unavoidable if the detection statistics is to have the CFAR property under a radial swelling of the distribution under H_0. Yet imposing a bottom limit on the axial well of the decision surface, as is the case when considering the union of a cone and a plane, or ofsetting the cones's tips, or using either a paraboloid or hyperboloid as a decision surface, may totally preclude the detection of targets at very low SNRs. Another way to limit the false alarm rate without completely preventing detection of low-SNR targets is to pinch the tips of the cone, employing a profile shaped as a square root, for example.

A square root profile would be given by $y = \sqrt{ax}$:

$$\begin{aligned} \|\mathbf{P_t y}\| &= \sqrt{a \left\|\mathbf{P_t^\perp y}\right\|}\,, \\ \|\mathbf{P_t y}\|^2 &= a \left\|\mathbf{P_t^\perp y}\right\|\,, \\ \frac{\mathbf{y}^\top \mathbf{P_t y}}{\left(\mathbf{y}^\top \mathbf{P_t^\perp y}\right)^{\frac{1}{2}}} = a &\begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases} . \end{aligned} \tag{35}$$

In this case the threshold adjusts the aperture of the decision surface. However it opens very fast in the shape of a trumpet bell, meaning that the constraint on angular proximity of the tested measurement with the target is very lax when its norm is large. This may be an interesting behavior but we are not aware of a scenario in which this could provide useful.

4.2.4 Piecewise-continuous profile

Limiting the trumpet-like extension of the decision surface found in the previous example can be done by providing a piecewise-continuous profile, in which the square-root dependance is replaced by a linear dependance at higher measurement norms:

$$y = \begin{cases} \sqrt{ax} & (x < c) \\ mx + b & (x \geq c) \end{cases} . \tag{36}$$

This detection statistic behaves like a conical one at high measurement norm ($x \geq c$) but is more restrictive at low norm. It is however more complex to implement because of the necessary test on the norm of the input measurement. Note that the same decision surface can be obtained by the union of a square root and a conical surface. First-order continuity can be obtained by carefully chosing parameters a, b, c and m.

4.2.5 Another cylindrical decision surface

In the additive measurement model of Equation (1), the measurement distribution under H_1 is translated (by the vector quantity $a\mathbf{s}$) but its covariance is unchanged by the presence of the target. Thus, there is no need for an axially-symmetric decision surface that keeps expanding along the direction of the target signature: at large abundance factors a, a constant threshold on the accepted spectral angle will probably include a measurement subspace that is much larger than the measurement distribution under H_1. One possible fix to that problem is to use a decision surface that is not expanding but which is still axially-symmetric around $\mathbf{t}$. A cylinder having $\mathbf{t}$ as its axis would fit that description and would be given by the constraint $x = r$, leading to the test:

$$\begin{aligned} x^2 &= r^2\,, \\ \mathbf{y}^\top \mathbf{P}_\mathbf{t}^\perp \mathbf{y} = r^2 &\begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases} . \end{aligned} \tag{37}$$

This test would however possess a false alarm rate close to 1, because the bulk of the measurement distribution under H_0 is also contained within the resulting cylinder. In order to prevent this, the cylindrical decision surface should be combined with another constraint that would act as a bottom on the cylinder. This could be planar decision boundaries, but also (and perhaps more simply), an anomaly detector: the intersection (logical *and*) between the H_1 region of an anomaly detector (i.e. the exterior of its shell) and the H_1 region of the cylindrical decision surface above (its interior) would produce cylinders capped with curved surfaces following the H_0 distribution profile. Note that the cylindrical decision surfaces presented in subsection 3.2 are not axially-symmetric along $\mathbf{t}$; instead the cylinder wall of planar subspace detectors is perpendicular to all targets signatures.

4.2.6 Tangent profile

A variation on the capped cylindrical decision surface presented above would be to use a tangent-shaped profile:

$$y = b \tan\left(\frac{\pi x}{a} - \frac{\pi}{2}\right) + c\,. \tag{38}$$

The $\frac{\pi}{2}$ translation is meant to give a profile that is almost cylindrical when y is large but pinched when it is small, reducing the false alarm rate more drastically than a conical decision surface. Parameter a adjusts the radius of the limiting cylinder (which is unity when $a = 1$). Parameter c raises the position of the tangent profile's knee, which should probably be adjusted at the distance from $\mathbf{b}_0$ where one would traditionnaly use a planar decision surface. Parameter b is a stretch factor that adjusts the slope of the knee.

4.2.7 General profile

While it is unclear if any of the previous decision surfaces would provide an advantage over conical or Kelly's statistics, with which they share their axial symmetry, they illustrate how any profile can be used with the proposed technique. In general, arbitrary profiles specified by $y = \tau f(x)$ will yield detection tests of the form

$$d = \frac{y}{f(x)} \begin{matrix} > \\ = \\ < \end{matrix} \tau \begin{cases} \text{choose } H_1 \\ H_0 \text{ or } H_1 \\ \text{choose } H_0 \end{cases} . \tag{39}$$

Without loss of generality, if $f(x)$ is expanded into its power series $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots$, one can see that τ pre-multiplies all of the coefficients when considering the decision surface given by $y = \tau f(x)$. Thus the threshold influences both the distance of the decision surface in the direction of $\mathbf{t}$, as is the case with planar surfaces; the slope of the linear profile (and thus its opening angle), as is the case with conical surfaces; and all further terms. In the non-transformed measurement space, the axially-symmetric profiles adopt squished (elliptical) cross-sections instead of circular ones.

It could be interesting to build decision surfaces that are tailored to the background clutter distribution by carefully choosing $f(x)$, perhaps by compiling an histogram of the actual values of y from a training set. The decision surface profile would then be obtained by interpolating the results and then computing an inverse profile within the range of the histogram, and extrapolating the profile outside the range of the histogram with either a cylindrical or conical shape. This idea has yet to be tested but might provide a way to adapt an axially-symmetric decision surface to non-elliptical measurement distributions.

4.3 Other transformations

We explored the three cases $\mathbf{A} = \mathbf{I}$, $\mathbf{A} = \mathbf{\Sigma_b}^{-\frac{1}{2}}$ and $\mathbf{A} = \mathbf{P_B^\perp}$, which are different ways of dealing with background clutter. The formalism of specifying an arbitrary transform $\mathbf{A}$ suggests that other background clutter models could be taken into account, or in general, new properties can be injected into otherwise known detection statistics just by tweaking the transformation matrix $\mathbf{A}$.

One example of this is rendering detection statistics invariant to biases in $\mathbf{y}$ or $\mathbf{t}$, which could be useful for measurement models in which an unknown bias appears[9].

[9] We are however not aware of such a model; errors in atmospheric compensation or signature characterization may induce such biases.

Intuitively, this could be achieved my removing the average levels $\bar{x}$ and $\bar{s}$ of $\mathbf{x}$ and $\mathbf{s}$, respectively, before computing the detection statistics. This operation is a projection transformation out of the bias space $\mathbf{1} = [1\,1\,\cdots\,1]^\top$:

$$\mathbf{P}_{\mathbf{1}}^{\perp} = \mathbf{I} - \mathbf{1}\left(\mathbf{1}^\top\mathbf{1}\right)^{-1}\mathbf{1}^\top. \tag{40}$$

The expression above is simply the same as in Equation (8), but with the background subspace $\mathbf{B}$ replaced with the bias subspace $\mathbf{1}$. When using $\mathbf{A} = \mathbf{P}_{\mathbf{1}}^{\perp}$, for example, the conical detection statistics becomes Pearson's coefficient of determination, r^2. It is to be noted that r^2 can be proven to be the GLRT for the additive model with unknown biases [5].

It is interesting to see that this property can also be blended with projection outside of a background clutter subspace. The projector:

$$\mathbf{P}_{\mathbf{B},\mathbf{1}}^{\perp} = \mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{P}_{\mathbf{z}}^{\perp}\mathbf{P}_{\mathbf{B}}^{\perp}\,, \tag{41}$$

with $\mathbf{z} = \mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{1}$, is a projector that is simultaneously orthogonal to $\mathbf{1}$ and $\mathbf{B}$. Using $\mathbf{A} = \mathbf{P}_{\mathbf{B},\mathbf{1}}^{\perp}$, one can add bias invariance to all detection statistics using the background subspace model. Bias invariance could also be included in detections statistics based on the statistical background model by using $\mathbf{A} = \mathbf{\Sigma}_{\mathbf{b}}^{-1}\mathbf{P}_{\mathbf{1}}^{\perp}$.

The linear transformation $\mathbf{A}$ described in this paper could also conceivably be replaced by a non-linear transformation of both $\mathbf{x}$ and $\mathbf{s}$. This could be used to adapt the test's decision surface to the shape of the clutter's distribution. Producing decision surface tightly tailored for a clutter distribution is however aptly done using the so-called *kernel trick*-based detection statistics [25], in which all the samples in a training set are involved (at the cost of more computing effort).

5 Conclusion

By examining the shape of the decision surfaces generated by threshold-based detection tests, we established that a large number of popular detections statistics for hyperspectral target detection can be bundled in a few categories that share the same basic decision surface geometry, despite differences in nomenclature and form. Indeed, detection statistics with different names and described by different equations can be made equivalent by using the appropriate parameters. Equivalently, many detection statistics found in the literature can be implemented as a parametrization of a single expression. This also suggests that comparing the performances of detection statistics of a same family is futile, since they can be made to have the exact same decision surface with the use of the appropriate parameters.

An important class of detection statistics possess a planar decision boundary, the position and orientation of which can be adjusted by the threshold value and a linear transformation of the measurement and target vectors. These statistics can be seen as (possibly monotone transformations of) abundance or energy estimators. They cannot possess the CFAR property for radially symmetric background clutter, but can be turned into anomaly detectors — which possess closed decision surfaces — by choosing a target subspace that encompasses all measurement space.

Making the threshold of a planar statistic proportional to an estimate of the noise power turns the planar decision surface into a cone around the target vector, and enables CFAR operation for radially symmetric background clutter. In the case of the background subspace model, in which the clutter is modeled as a subspace instead of a normal distribution, the decision surface is wedge-shaped along the clutter dimensions. The threshold controls the aperture of the cone or wedge in such statistics; the linear transformation applied to measurements and targets produces flattened cones. They can be expressed in a variety of forms which can complicate comparisons, but they all can be seen as estimate of signal to noise ratio.

Kelly's detector stands out in this tableau with its hyperboloid-shaped detection surface, which share properties of both the planar and the conical statistics, at the cost of using a supplementary control parameter. We propose in this work a way to generate axially symmetric decision surfaces with arbitrary profiles using a simple formalism used to analyze Kelly's detector.

Finally, changing the linear transformation applied to the measurements and targets allowed us to build detection statistics that, to our knowledge, are not yet described in the literature, such as anomaly detectors and Kelly-like detectors that use the background subspace model, or bias-insensitive statistics.

We conclude this work by nothing that while many detection statistics can be made

equivalent, this emphasizes the role of choosing the right parameters over the right statistic. Indeed, choosing the right threshold or the right linear transformation (which is linked to background clutter estimation, and thus clutter sampling and training set management) can have a much more important impact on performance than choosing a particular statistic. It may be argued that conical statistics should be preferred because of the CFAR property. However, CFAR operation for conical statistics is only demonstrated for radially symmetric clutter distributions, which may not fit experimental distributions. Thus, quasi-CFAR detection statistics, such as the truncated cone, Kelly's detector, or the paraboloid and offset cones proposed in this work, may provide a compromise between planar and conical decision surfaces. The cost of using these statistics is setting the value of a second parameter that will both affect false alarm and detection rates.

References

[1] Thériault, J.-M., Bradette, C., and Gilbert, J. (1996), Atmospheric remote sensing with a ground-based spectrometer system, In *Proceedings of the SPIE*, Vol. 2744, p. 664.

[2] Thériault, J.-M. (2001), Passive standoff detection of chemical vapors by differential FTIR radiometry, (Technical Report DREV-TR-2000-156) DRDC Valcartier.

[3] Lachance, R. L. (1997), GASEM: Gaseous emanations detection, (Technical Report SP-BOM-001/97) Bomem.

[4] Thériault, J.-M., Lacasse, P., Lavoie, H., Bouffard, F., Puckrin, E., Turcotte, C. S., and Dubé, D. (2010), The CATSI EDM System: Final Report, (Technical Report TR-2010-115) Defence R&D Canada — Valcartier.

[5] Bouffard, F. (2008), Detection and identification algorithms for CATSI EDM: Theoretical basis, (Technical Report TR 2008-165) Defence R&D Canada — Valcartier.

[6] Thériault, J.-M., Lacasse, P., Fortin, G., Lavoie, H., Bouffard, F., Dubé, D., Bergeron, E., and Nadeau, G. (2012), MoDDIFS long-range gas detection configuration: Test report November 2010 to September 2011, (Technical Report TR-2012-263) Defence R&D Canada — Valcartier. PROTECTED A.

[7] Thériault, J.-M., Puckrin, E., Bouffard, F., Déry, B., Moreau, L., and Jensen, J. O. (2004), Passive remote monitoring of chemical vapors by differential FTIR radiometry: Results at a range of 1.5 km, *Applied Optics*, 43, 1425.

[8] Manolakis, D., Lockwood, R., Cooley, T., and Jacobson, J. (2009), Is there a best hyperspectral detection algorithm?, *SPIE Newsroom*, p. 3.

[9] Fawcett, T. (2003), ROC Graphs: Notes and practical considerations for data mining researchers, (Tech Report HPL-2003-4) Hewlett-Packard Laboratories Palo Alto.

[10] Fawcett, T. (2006), An introduction to ROC analysis, *Pattern Recognition Letters*, 27, 861.

[11] Poor, H. V. (1998), An introduction to signal detection and estimation, second edition ed, New York: Springer.

[12] Kay, S. M. (1998), Fundamentals of statistical signal processing volume II: detection theory, Upper Saddle River, NJ: Prentice-Hall.

[13] Trefethen, L. N. and Bau, D., III (1997), Numerical linear algebra, Philadelphia: Society for Industrial and Applied Mathematics.

[14] Manolakis, D., Marden, D., and Shaw, G. A. (2003), Hyperspectral image processing for automatic target detection applications, *Lincoln Laboratory Journal*, 14, 79.

[15] Farrand, W. and Harsanyi, J. (1997), Mapping the Distribution of Mine Tailings in the Coeur d'Alene River Valley, Idaho, through the Use of a Constrained Energy Minimization Technique, *Remote Sensing of the Environment*, 59, 64–76.

[16] Settle, J. (2002), On constrained energy minimization and the partial unmixing of multispectral images, *IEEE Transactions on Geoscience and Remote Sensing*, 40(3), 718–721.

[17] Kraut, S. and Scharf, L. L. (1999), The CFAR adaptive subspace detector is a scale-invariant GLRT, *IEEE Transactions on Signal Processing*, 47, 2538.

[18] Kraut, S., Scharf, L. L., and McWorther, L. T. (2001), Adaptive subspace detectors, *IEEE Transactions on Signal Processing*, 49, 1.

[19] Manolakis, D., Siracusa, C., and Shaw, G. (2001), Hyperspectral subpixel target detection using the linear mixing model, *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1392.

[20] Scharf, L. L. and Friedlander, B. (1994), Matched subspace detectors, *IEEE Transactions on Signal Processing*, 42, 2146.

[21] Kelly, E. J. (1986), An adaptive detection algorithm, *IEEE Transactions on Aerospace and Electronics Systems*, 22, 115.

[22] Schaum, A. (2009), Remote spectral detection using a laboratory signature, In *IEEE WHISPERS 2009*.

[23] Schaum, A. (2010), Methods of hyperspectral detection based on a single signature sample, *IEEE Sensors Journal*, 10, 518–523.

[24] Schaum, A. (2010), Continuum fusion: a theory of inference, with applications to hyperspectral detection, *Optics Express*, 18, 8171–8181.

[25] Nasrabadi, N. M. and Kwon, H. (2005), Kernel spectral matched filter for hyperspectral target detection, Vol. 4, p. 665.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence R&D Canada – Valcartier
2459 Pie-XI Blvd. North, Québec QC G3J 1X5, Canada

2a. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED

2b. CONTROLLED GOODS

(NON-CONTROLLED GOODS)
DMC A
REVIEW: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Decision surfaces of binary tests in hyperspectral detection

4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.)

Bouffard, F.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
April 2013	48	25

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – Valcartier
2459 Pie-XI Blvd. North, Québec QC G3J 1X5, Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC Valcartier TR 2013-121	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Defence departments and defence contractors; further distribution only as approved
( ) Defence departments and Canadian defence contractors; further distribution only as approved
( ) Government departments and agencies; further distribution only as approved
( ) Defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

This report is first and foremost an analysis of the decision surfaces associated with common detection statistics used for hyperspectral target detection. The intention is to clarify the process leading to detection using those statistics and show where their differences and similarities lie. The large number of available detection statistics, along with their numerous parameters, makes formal comparisons complicated and sometimes produce ambiguous or uncertain conclusions. This document tries to establish a solid framework on which to base this analysis and then demonstrates how detection statistics can be grouped in distinct classes based on their decision surface's geometry. In doing so, possibly new detection statistics (or methods that can be used to build them) are exposed.

Ce rapport se veut principalement une analyse des surfaces de décision associées aux statistiques de détection utilisées couramment en détection hyperspectrale de cibles. L'intention première est la clarification du procédé menant `a la détection en utilisant ces statistiques, ainsi que l'exposition de leur différence et similarités. Le grand nombre de statistiques de détection disponibles, couplées avec leurs nombreux paramètres, rend les comparaisons formelles complexes qui parfois produisent des conclusions ambigues ou incertaines. Ce document tente d'établir un formalisme solid sur lequel cette analyse peut se baser et démontre ensuite comment les statistiques de détection peuvent être groupées en classes distinctes en se basant sur la géométrie de leurs surfaces de décision. Ce faisant, des statistiques de détection possiblement nouvelles, ainsi que des méthodes pouvant être utilisées pour les bâtir, seront exposées

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Decision test, decision surface, hyperspectral target detection

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

www.drdc-rddc.gc.ca