Estimation of Subspace Occupancy

by

Kaitlyn Beaudet

A Thesis Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Science

Approved November 2014 by the
Graduate Supervisory Committee:

Douglas Cochran, Chair
Pavan Turaga
Visar Berisha

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

The ability to identify unoccupied resources in the radio spectrum is a key capability for opportunistic users in a cognitive radio environment. This paper draws upon and extends geometrically based ideas in statistical signal processing to develop estimators for the rank and the occupied subspace in a multi-user environment from multiple temporal samples of the signal received at a single antenna. These estimators enable identification of resources, such as the orthogonal complement of the occupied subspace, that may be exploitable by an opportunistic user. This concept is supported by simulations showing the estimation of the number of users in a simple CDMA system using a maximum *a posteriori* (MAP) estimate for the rank. It was found that with suitable parameters, such as high SNR, sufficient number of time epochs and codes of appropriate length, the number of users could be correctly estimated using the MAP estimator even when the noise variance is unknown. Additionally, the process of identifying the maximum likelihood estimate of the orthogonal projector onto the unoccupied subspace is discussed.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

The problem of determining whether a signal is present in two or more channels of sensor data has applications in many different fields. The application context where it has been most studied is in defense and security systems, such as radar and sonar, where it pertains to detecting and localizing a target from data collected at multiple geographically distributed sensors. However, due to the need to improve the utilization of spectrum resources, detection methods of this kind have been applied over the past decade in spectrum sensing for cognitive radio in order to determine the presence of a primary user.

Tests for determining the presence of a common but unknown signal in two or more noisy channels have been studied extensively in connection with passive localization of emitters. Such detectors include those based on the magnitude-square coherence (MSC) estimate and generalized coherence (GC) estimate, which are functions of the determinant of a Gram matrix formed from the collected data. The rise of multiple input, multiple output (MIMO) systems in sensing and communications has led to a renewed interest in multiple-channel detection. Motivated in part by MIMO applications, a variety of statistical hypothesis tests including generalized likelihood ratio tests (GLRTs), Bayesian tests, locally most powerful invariance tests (LMPITs), and maximum *a posteriori* (MAP) tests have been recently been derived for various multiple-channel sensing problems, in many cases yielding results that are functions of the eigenvalues of the Gram matrix.

This thesis explores how recently derived detectors and their properties can be applied to the spectrum sensing problem. Chapter 2 looks at the previous work ac-

complished. This includes a summary of spectrum sensing in cognitive radio and various relevant detection methods. Additionally, the Gram matrix is examined and various established detectors and their features are explored. Chapter 3 shows the similarities between the problem formulations for application in spectrum sensing and more classical problems motivated by radar/sonar. Chapter 4 discusses the formulation of the MAP estimate for signal rank. Chapter 5 shows the use of the MAP estimate for the rank of a signal to estimate the number of users in a CDMA system, and also discusses estimation of the occupied subspace. Finally, Chapter 6 discusses potential problems in real world utilization.

Chapter 2

BACKGROUND

Cognitive radio aims to improve the utilization of the radio frequency (RF) spectrum by allowing unlicensed secondary users to transmit on the spectrum when it is not being used by a licensed primary user. As stated in [1], usage of allocated spectrum has been reported to have utilization ranging from 15% to 85%. Spectrum sensing is a fundamental part of cognitive radio, as it helps to identify available spectrum, thus maintaining low interference for the primary user, while enabling access of secondary "opportunistic" users. As discussed in [2], spectrum sensing can be used to obtain the spectrum usage characteristics across multiple dimensions including time, space, frequency and code.

2.1 Spectrum Sensing Methods

There are many different techniques used for spectrum sensing, such as the methods described in [3] which include energy detection, cyclostationary detection, matched filter detection and detection using multiple antennas.

One of the most common methods of spectrum sensing is the energy detector, which has low computational and implementation complexities. Additionally, the receivers require no knowledge of the primary user's signal. The test statistic for the energy detector was given in [2] as

$$T_{ED} = \sum_{n=0}^{N} |y(n)|^2$$

where N is the size of the observation vector and $y(n)$ is the received signal. However, the energy detector has poor performance under low-SNR conditions and does not

work efficiently for detecting spread spectrum signals. Additionally, it was discussed in [3] that the energy detector is the optimal detector if the only the noise power is known. However, it is unable to distinguish between different types of signals which can increase the probability of false alarm. Additionally, the influence of uncertain noise power can make it challenging to determine the detection threshold.

Cyclostationary-based sensing exploits the cyclostationarity features of the received signals using feature detection for detecting primary user transmissions. Such features are caused by the periodicity in the signal statistics such as mean and autocovariace. Detection is typically based on an estimate of the cyclic spectral density, given in [2] for a discrete-time process y as

$$S(f,\alpha) = \sum_{\tau=-\infty}^{\infty} R_y^{\alpha}(\tau) e^{-j2\pi f\tau}$$

where

$$R_y^{\alpha}(\tau) = E\left[y(n+\tau)y^*(n-\tau)e^{j2\pi\alpha n}\right]$$

is the cyclic autocorrelation function and α is the cyclic frequency. The cyclic spectral density outputs peak values when the cyclic frequency and the fundamental frequencies of the transmitted signal are equivalent. As discussed in [3], the main advantage of cyclostationary spectrum sensing is the ability to differentiate the primary signal from noise as well as interference with different cyclic frequencies, which might arise from the presence of interfering signals with different modulation types or parameters. Additionally, the SNR does not affect the cyclostationary feature, which enables this method of detection to be successful even at very low SNRs. Various specific detectors for cyclostationary features have been developed in [4, 5, 6, 7], some quite recently and with impetus from spectrum sensing [8].

Matched filtering is also discussed in [2], and is the optimal detector for one receiver when the transmitted signal is known. However, it requires the cognitive

radio to be able to demodulate the received signals, as perfect knowledge of the primary users' signaling features is required. The matched filter statistic is given in [9] for a real signal as

$$T_{MF}(\mathbf{x}) = \sum_{n=0}^{N-1} \mathbf{s}^T(n)\mathbf{x}(n)$$

where $s(n)$ is the source signal and is deterministic and known, and $x(n)$ is the received data.

In [10] a multitaper spectral estimation method using Slepian tapers was used to define a decision statistic for detecting the transition into a spectrum hole. A spectrum hole is defined as a band of frequencies assigned to a primary user that is not being utilized at a particular time and geographical location. The detector is based on the statistic

$$D(t) = \sum_{l=0}^{L-1} \sum_{v=0}^{M-1} \left|\sigma_l \left(f_{\text{low}} + v \cdot \Delta f; t\right)\right|^2 \Delta f$$

where L represents the number of largest eigenvalues of the matrix $\mathbf{A}(f)^\dagger \mathbf{A}(f)$ that are considered to play important roles in estimating the interference temperature, where $\mathbf{A}(f)$ is a spatiotemporal complex-valued matrix whose columns are produced using stimuli sensed at different gridpoints, and whose rows are computed using different Slepian tapers with variable weights accounting for relative areas of gridpoints. The l^{th} largest eigenvalue produced by the burst of RF stimuli received at time t is denoted by $\left|\sigma_l(f,t)\right|^2$, and M denotes the number of frequency resolutions of width Δf which occupy the occupied space under scrutiny, f_{low} is the lowest end of an occupied space, and $f = f_{\text{low}} + v \cdot \Delta f, v = 0, 1, \ldots, M-1$. This method is shown to be almost optimal for wideband signals, however has a high computational complexity [11]. This method was expanded in [12] to include an adaptive multitaper method for when the number of tapers is increased towards the limiting value of $2NW$ where NW is the time-bandwidth product. The adaptive method is computed through an iterative process.

Additional methods discussed in [2] include radio identification based sensing which identifies the transmission technology (e.g. modulation type) used by the primary users by extracting features from the received signal that are used for selecting the more probable primary user technology and then using energy detector based methods for detection. Also mentioned are wavelet transform based estimation, Hough transform and time-frequency analysis.

However, as expressed in [13], the detection of a source with a sensor array is of particular interest in cognitive radio where it can be used with a mutlisensor cognitive device or a collaborative network. As stated in [3], spectrum sensing using multiple antennas can be accomplished using eigenvalue-based detection. Such detectors tend to be robust to noise power uncertainty as the noise variance and signal power are estimated simultaneously. A review of eigenvalue-based detectors will be given in section 2.5.

Although the focus of this thesis is on cognitive radio, it is worth noting that there has been renewed interest in multiple channel detection and estimation for radar purposes. This is due in part to the rise in interest in passive radar, which uses one or more high-SNR direct path signals from illuminators of opportunity. For example, [14] illustrates the use of geosynchronous satellites and a terrestrial TV transmitter as illuminators of opportunity. Interest in such applications has resulted in new findings for multiple channel detection in radar settings over the past couple of years [15, 16, 17, 18] .

With this in mind, previous and recent detectors and estimators developed with application focusing on radar/sonar and spectrum sensing will be explored. These detectors generally turn out to be functions of the eigenvalues of the Gram matrix.

2.2 Gram Matrix

Consider M complex N-vectors $\mathbf{x}_1, \ldots, \mathbf{x}_M$ with $M < N$ corresponding to segments of time-series data from a collection of M channels. The channels may be from distinct receivers or formed from time segments of a single receiver. Denoting the associated matrix as

$$\mathbf{X} = \begin{bmatrix} - & \mathbf{x}_1 & - \\ & \vdots & \\ - & \mathbf{x}_M & - \end{bmatrix}$$

From [19, Pg 177], the positive semidefinite Gram matrix of $\mathbf{X}$ is an $M \times M$ matrix of the inner products denoted as

$$\mathbf{G} = \mathbf{X}\mathbf{X}^\dagger = \begin{bmatrix} \langle \mathbf{x}_1, \mathbf{x}_1 \rangle & \langle \mathbf{x}_1, \mathbf{x}_2 \rangle & \ldots & \langle \mathbf{x}_1, \mathbf{x}_M \rangle \\ \vdots & & & \vdots \\ \langle \mathbf{x}_M, \mathbf{x}_1 \rangle & \langle \mathbf{x}_M, \mathbf{x}_2 \rangle & \ldots & \langle \mathbf{x}_M, \mathbf{x}_M \rangle \end{bmatrix} \tag{2.1}$$

where $\mathbf{X}^\dagger$ denotes the Hermitian transpose of $\mathbf{X}$. And the inner product $\langle \mathbf{x}_i, \mathbf{x}_j \rangle, i, j = 1, \ldots, M$ is defined by

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = \sum_{k=1}^{N} x_{i,k} x_{j,k}^*$$

where $*$ denotes complex conjugation.

If the vectors $\mathbf{x}_m$ are normalized to unit length (i.e. $\mathbf{x}_M$ is replaced by $\frac{\mathbf{x}_M}{||\mathbf{x}_M||}$), the Gram matrix takes the form

$$\mathbf{G} = \begin{bmatrix} 1 & \left\langle \frac{\mathbf{x}_1}{||\mathbf{x}_1||}, \frac{\mathbf{x}_2}{||\mathbf{x}_2||} \right\rangle & \ldots & \left\langle \frac{\mathbf{x}_1}{||\mathbf{x}_1||}, \frac{\mathbf{x}_M}{||\mathbf{x}_M||} \right\rangle \\ \vdots & \ddots & & \vdots \\ \left\langle \frac{\mathbf{x}_M}{||\mathbf{x}_M||}, \frac{\mathbf{x}_1}{||\mathbf{x}_1||} \right\rangle & \left\langle \frac{\mathbf{x}_M}{||\mathbf{x}_M||}, \frac{\mathbf{x}_2}{||\mathbf{x}_2||} \right\rangle & \ldots & 1 \end{bmatrix}$$

2.2.1 Determinant of the Gram Matrix

The determinant of the Gram matrix is denoted as

$$g = g(\mathbf{x}_1, \ldots, \mathbf{x}_M) = |\mathbf{G}|$$

From [19, Pg 178], the determinant is bounded by

$$0 \leq g(\mathbf{x}_1, \ldots, \mathbf{x}_M) \leq \prod_{i=1}^{M} ||\mathbf{x}_i||^2 \tag{2.2}$$

If the elements of the matrix have been normalized, then

$$0 \leq g(\mathbf{x}_1, \ldots, \mathbf{x}_M) \leq 1$$

Where $g = 0$ occurs if and only if vectors $\mathbf{x}_m,\ m = 1, \ldots, M$ are linearly dependent (i.e. $\mathbf{X}$ is not full rank), and the upper extreme occurs if and only if the vectors are orthogonal.

The determinant of the Gram matrix has the following properties [19, Pg 184]

(a) g is a symmetric function of its arguments.

(b) $g(\mathbf{x}_1, \ldots, \sigma\mathbf{x}_j, \ldots, \mathbf{x}_n) = |\sigma|^2 g(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

(c) $g(\mathbf{x}_1, \ldots, \mathbf{x}_j + \sigma\mathbf{x}_k, \ldots, \mathbf{x}_n) = g(\mathbf{x}_1, \ldots, \mathbf{x}_n), j \neq k$.

(d) $g^{\frac{1}{2}}(\mathbf{x}_1' + \mathbf{x}_1'', \mathbf{x}_2, \ldots, \mathbf{x}_n) \leq g^{\frac{1}{2}}(\mathbf{x}_1', \mathbf{x}_2, \ldots, \mathbf{x}_n) + g^{\frac{1}{2}}(\mathbf{x}_1'', \mathbf{x}_2, \ldots, \mathbf{x}_n)$.

(e) $g(\mathbf{x}_1, \ldots, \mathbf{x}_n) \leq g(\mathbf{x}_1, \ldots, \mathbf{x}_p) g(\mathbf{x}_{p+1}, \ldots, \mathbf{x}_n), 1 \leq p < n$.

These properties add desirable attributes to detectors, such as the generalized coherence detector, and are important in deriving the distributions of eigenvalue-based detection statistics under suitable null hypotheses.

2.2.2 Distribution and Invariance of the Normalized Gram Matrix

In [20], a geometric perspective is used to determine the distribution of the normalized Gram matrix under suitable H_0 conditions. Additionally, it is shown that the null distribution of the normalized Gram matrix does not depend on the distribution of one channel, and thus is invariant to the statistics of one channel. This invariance result carries over to the spectrum of the Gram matrix.

The invariance to the statistics of one channel is an important tool especially in the case of passive and active radar when the transmitted signal can be obtained via a high SNR received signal, or an exact replica respectively. These benefits extend to spectrum sensing in cases when the primary user is known. The invariance property means that the replica or high SNR signal can be contained in one channel of data without affecting the thresholds set using a desired false alarm probability rate.

2.2.3 Sample Covariance Matrix

The sample covariance matrix is the maximum likelihood estimate for the covariance matrix of a $M \times N$ random matrix $\mathbf{X}$ and assuming zero mean is denoted by

$$\hat{\mathbf{R}} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\dagger} \tag{2.3}$$

$\hat{\mathbf{R}}$ is an $M \times M$ matrix. The $N \times N$ matrix

$$\mathbf{W} = \frac{1}{N}\mathbf{X}^{\dagger}\mathbf{X} \tag{2.4}$$

has the same non-zero eigenvalues as the sample covariance matrix.

2.2.4 Eigenvalues of the Gram Matrix

Recently the individual eigenvalues of the Gram matrix have been seen to be useful in various detection tests. The invariance of the distribution of the eigenvalues

to the statistics of one channel was proven in [21].

2.2.4.1 *Distribution of the Eigenvalues of Complex Wishart Distributed Matricies*

Detection thresholds corresponding to false alarm probabilities are set using the known distribution of the detection statistic under the signal absent null hypothesis. When the null hypothesis conditions are assumed, that is that the M channels are independent and contain only independent complex Gaussian noise, the Gram matrix belongs to the class of complex Wishart distributed matrices.

In [22], it is shown that Wishart distributed matrices can be factored into $\mathbf{T}^{\dagger}\mathbf{T}$, where $\mathbf{T}$ is a upper triangular matrix with real values on the main diagonal. Such factoring is useful in determining the distribution of functions of the elements of a complex Wishart matrix, such as the multiple coherence between the last variable in a M-tuble, and the remaining $M-1$ variables.

In [23], the distribution of the largest eigenvalue was found by integrating the joint PDF of $\boldsymbol{\lambda}$ over $\lambda_2, \ldots, \lambda_M$ where λ_2 and λ_M are the second largest, and smallest eigenvalues respectively. Thus, the PDF of the largest eigenvalue can be written as

$$f_{\lambda_1}(x_1) = \int_0^{x_1} \int_0^{x_2} \cdots \int_0^{x_{M-1}} f_{\lambda}(\mathbf{x}) dx_M \ldots dx_3 dx_2$$

where

$$f_{\lambda}(\mathbf{x}) = K|\Phi(\mathbf{x})| \cdot |\Psi(\mathbf{x})| \prod_{l=1}^{M} (\xi(x_l))$$

and $\mathbf{x} = [x_1, \cdots, x_M]^T$ and $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_M]^T$ is a vector of ordered eigenvalues ($\lambda_1 \geq \cdots \geq \lambda_M$). The values of the normalizing constant K, $\Phi(\mathbf{x})$, $\Psi(\mathbf{x})$ and $\xi(x)$ are given in Table 1 of [23] for uncorrelated central, uncorrelated non-central, and correlated central Wishart matrices. Additionally, the PDF for the smallest eigenvalue, and the m^{th} ordered eigenvalue was found.

In [24], the exact joint density function for the M eigenvalues is found for complex

Wishart matrices. Additionally, the PDFs for the largest and smallest eigenvalues of real and complex Wishart matrices are discussed, with exact distributions being given for some cases.

In [25], the exact distribution was derived for the scaled largest eigenvalue given by

$$X := \frac{\lambda_1}{\frac{1}{M}\sum_{i=1}^{M}\lambda_i} = \frac{\lambda_1}{T}$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_M > 0$ are the eigenvalues of the matrix $\mathbf{G}$ as given in equation (2.1).

Similarly, in [26], the exact density of the condition number of a complex Wishart matrix was found and used to calculate the exact probability of false alarm for the ratio of the maximum and minimum eigenvalues.

Additionally, some thresholds for detectors based on the eigenvalues of $\hat{\mathbf{R}}$ can be approximated since as $N, M \to \infty$, the distribution of some eigenvalue-baed detection statistics asymptotically follow a second order Tracy-Widom distribution. An example of this was shown in [27].

Thus, there are cases for which false alarm thresholds can be analytically determined using the known, if rather obstinate, distributions of the eigenvalues of complex Wishart-distributed matrices.

2.3 Magnitude-Squared Coherence Estimator

The magnitude-squared coherence (MSC) estimate has seen wide application in situations involving two channels for over five decades. The MSC is a function of the inner product of two channels shown in [28] to be

$$\gamma^2(\mathbf{x}_1, \mathbf{x}_2) = \frac{\left|\langle \mathbf{x}_1, \mathbf{x}_2 \rangle\right|^2}{\left|\left|\mathbf{x}_1\right|\right|^2 \left|\left|\mathbf{x}_2\right|\right|^2} \tag{2.5}$$

The geometry of the MSC was shown in [29] by denoting $\mathbf{x}_1 = \{\tau_n + i\eta_n\}_1^N$ and

$\mathbf{x}_2 = \{u_n + iv_n\}_1^N$, $\boldsymbol{\alpha}, \boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are unit vectors in $\mathbf{R}^{2N}$ defined by

$$\begin{aligned}
\boldsymbol{\alpha} &\triangleq \frac{(\tau_1, \cdots, \tau_N, \eta_1, \cdots, \eta_N)}{\sum_{n=1}^N \sqrt{\tau_n^2 + \eta_n^2}} \\
\boldsymbol{\beta}_1 &\triangleq \frac{(u_1, \cdots, u_N, v_1, \cdots, v_N)}{\sum_{n=1}^N \sqrt{u_n^2 + v_n^2}} \\
\boldsymbol{\beta}_1 &\triangleq \frac{(-v_1, \cdots, v_N, u_1, \cdots, u_N)}{\sum_{n=1}^N \sqrt{u_n^2 + v_n^2}}
\end{aligned}$$

so that (2.5) reduces to

$$\gamma^2(\mathbf{x}_1, \mathbf{x}_2) = \langle \boldsymbol{\alpha}, \boldsymbol{\beta}_1 \rangle^2 + \langle \boldsymbol{\alpha}, \boldsymbol{\beta}_2 \rangle^2$$

and the distribution can be seen to be the square of the length of the projection of $\boldsymbol{\alpha}$ onto the plane defined by $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are orthonormal by construction.

In [29], the geometry of the MSC was used to determine the cumulative distribution function (CDF) given as

$$\Pr\{\gamma^2(\mathbf{x}_1, \mathbf{x}_2) \leq R\} = 1 - (1 - R^2)^{N-1}$$

This was accomplished by determining the fraction of the surface of the unit sphere in $\mathbf{R}^{2N}$ that projects onto the annular region in $\mathbf{R}^2$ written in polar coordinates as $\{(r, \theta) : |R| < r \leq 1 \text{ and } 0 \leq \theta < 2\pi\}$. From this, the fraction of the sphere that projects on the disk of radius R centered at the origin in $\mathbf{R}^2$ was implicitly determined. The distribution can be utilized to set thresholds for desired false alarm probabilities.

The invariance of the MSC to the distribution of the data on one channel was shown in [28] under the conditions that the two channels are independent and the second channel contains only white Gaussian noise. This was accomplished by evaluating the conditional CDF of γ^2 holding $\mathbf{x}_2$ fixed, and showing that there is no dependence on the values of $\mathbf{x}_2$. This invariance was given a geometric interpretation in [29].

2.4 Generalized Coherence Estimator

The generalized coherence (GC) estimate was introduced in 1988 [30] for detection use in $M \geq 2$ channels. To form the GC estimate, the MSC estimate was written as

$$\gamma^2(\mathbf{x}_1, \mathbf{x}_2) = 1 - \frac{g(\mathbf{x}_1, \mathbf{x}_2)}{||\mathbf{x}_1||^2||\mathbf{x}_2||^2}$$

By generalizing the MSC estimate to the case of M non-zero sequences, the GC estimate is defined as

$$\gamma^2(\mathbf{x}_1, \ldots, \mathbf{x}_M) = 1 - \frac{g(\mathbf{x}_1, \ldots, \mathbf{x}_M)}{||\mathbf{x}_1||^2 \cdots ||\mathbf{x}_m||^2} \tag{2.6}$$

Due to the limits of $g(\mathbf{x}_1, \ldots, \mathbf{x}_M)$ provided in (2.2), the generalized coherence estimate will have values between zero and one.

Recently in [31], the GC estimate was derived from the Bayesian perspective, when each vector $\mathbf{x}_m$ contains independent samples of independent zero-mean Gaussian noise and the covariance matrix of $\mathbf{x}_1, \ldots, \mathbf{x}_M$ is diagonal under the signal-absent hypothesis and non-diagonal under H_1. Non-informative priors were determined for the covariance matrices under H_0 and H_1 and used to establish the likelihood ratio $\frac{p(\mathbf{X}|H_1)}{p(\mathbf{X}|H_0)}$. It was determined that the result of the likelihood function is a monotonic function of the GC estimate.

The geometry of the GC estimate is discussed in [32]. The determinant of the Gram matrix may be regarded as the squared volume of a parallelepiped in a complex N-dimensional space formed by sample vectors, $\mathbf{x}_1, \ldots, \mathbf{x}_M$. Normalizing the product of the squared lengths of the vectors yields a number between zero and one that is subtracted from unity to give the GC estimate.

The distribution of the GC estimate under H_0 can be found by using a Gram-Schmidt procedure to factor (2.6) into $\gamma^2(\mathbf{x}_1, \ldots, \mathbf{x}_M) = 1 - \prod_{j=2}^{M} z_j$ and in [30], it was found that the z_j are independent and that each z_j has a beta distribution with

$2N-2(j-1)$ and $2(j-1)$ degrees of freedom under H_0. In [33] a recursion formula was found and applied to generate values for the GC for two and three-channel estimates according to a range of false alarm probabilities and various sample sequence lengths N. Due to the difficulty of evaluating the probability distribution function under the signal absent null hypothesis, and the fact that the distribution is unknown for the signal present case, a more tractable asymptotic (in M) analysis of the GC estimate is shown in [34]. The false alarm probabilities found using the asymptotic method were shown to closely match the theoretical false alarm probabilities for the three-channel GC estimate, and the theoretical predictions of detection performance are shown to closely match empirical results obtained by Monte Carlo simulations.

The invariance property of the GC was proven in [35] by using a Gram-Schmidt procedure to express $\mathbf{x}_1, \ldots, \mathbf{x}_M$ in terms of orthogonal vectors, and expressing the GC in terms of an MSC estimate, which could be seen to have the desired invariance as discussed in 2.3. As a note, the invariance of distribution of the Gram matrix discussed in 2.2.2 supersedes the invariance of the distributions of the GC, the MSC and the individual eigenvalues.

2.5 Estimators Based on Individual Eigenvalues

Recently, detectors based on individual eigenvalues of the Gram matrix have been derived for both spectrum sensing applications and radar/sonar applications using various principles (e.g. GLRT, ML, MAP). In the derivations of these detectors, under the null hypothesis the received data contains only white noise whose strength is the same on each channel and may not be known. The alternative hypothesis varies from case to case. In some settings, the received data in each channel contains a common signal in additive noise. In others, the covariance matrix of the channels is assumed non-diagonal under H_1.

In [36] detection statistics were derived for multiple channel detection of signals with a known rank K in M independent channels using generalized likelihood ratio tests (GLRTs) and a Bayesian approach. For known noise variance, the GLR was given as

$$GLR = e^{\frac{N}{\sigma^2}\sum_{i=1}^{K}\lambda_i}$$

where N is the number of samples, σ^2 is the noise variance and λ_i denotes the i^{th} eigenvalue $\lambda_1 > \lambda_2 > \ldots > \lambda_N$ of the matrix $\mathbf{W}$. Similarly, the GLR for unknown noise variance was given as

$$GLR = \left(1 - \frac{\sum_{i=1}^{K}\lambda_i}{\sum_{i=1}^{N}\lambda_i}\right)^{-MN}$$

Using the Bayesian approach, the decision statistic in the case of known noise variance was found to be

$$\frac{p(X|H_1)}{p(X|H_0)} = Qe^{\frac{N\alpha}{\sigma^2}\sum_{i=1}^{K}\lambda_i}\prod_{i=1}^{N-K}\prod_{j=1}^{K}\frac{1}{\lambda_j - \lambda_{K+i} + \frac{\sigma^2}{\alpha}}$$

where $Q = \left(\frac{\pi\sigma^2}{N\alpha}\right)^{K(N-K)}\frac{1}{(1+\beta^2)^{MK}\text{vol}(G_{K,N})}$ and $\text{vol}(G_{K,N})$ is the volume of the complex Grassmanian manifold, $\alpha = \frac{\beta^2}{1+\beta^2}$ and $\beta^2 = \frac{\sigma_a^2}{\sigma^2}$. Similarly, when the noise variance is unknown, the decision statistic can be seen to be

$$\frac{p(X|H_1)}{p(X|H_0)} = Q\left(1 = \alpha\frac{\sum_{i=1}^{K}\lambda_i}{\sum_{i=1}^{N}\lambda_i}\right)^{K(N-K)-p}\prod_{i=1}^{N-K}\prod_{j=1}^{K}\frac{1}{\lambda_j - \lambda_{K+i} + \delta}$$

where $Q = \frac{\frac{\pi}{p\alpha}^{K(N-K)}}{((1+\beta^2)^{MK}\text{vol}(G_{K,N}))}$ and $p = MN + 1$.

In [37], the GLR was derived for a system of $M \geq 2$ antennas, with a primary signal known to have known spatial rank K. The log-GLR was found to be a function of the ratio between the geometric and arithmetic means of all eigenvalues and the difference of the number of antennas and known spacial rank smallest eigenvalues of

the known covariance matrix given by

$$\ln(\mathcal{L}) = MN \log \left[\frac{\left(\sum_{i=1}^{M} \lambda_i \right)^{\frac{1}{M}}}{\frac{1}{M} \sum_{i=1}^{M} \lambda_i} \right] - N(M-K) \log \left[\frac{\left(\sum_{i=K+1}^{M} \lambda_i \right)^{\frac{1}{M-K}}}{\frac{1}{M-K} \sum_{i=K+1}^{M} \lambda_i} \right]$$

For a rank one signal, this ends up being a function of the largest eigenvalue.

The locally most powerful invariant test for the signal-plus-noise model for linearly independent K signals impinging on an antenna monitoring system is given in [38] by

$$\mathcal{L} \propto \sum_{k=1}^{M} \lambda_k^2 = \left\| \hat{\bar{\mathbf{R}}} \right\|_F^2$$

where $\hat{\bar{\mathbf{R}}} = \frac{\hat{\mathbf{R}}}{\mathrm{Tr}(\hat{\mathbf{R}})}$ and $\hat{\mathbf{R}}$ is the sample covariance matrix.

In [27], likelihood ratio analysis was performed to find the detection statistics when the noise was known and unknown. When the noise was known, the statistic was seen to reduce to Roy's largest root test which can be given by

$$T_{RLRT} \triangleq \frac{\lambda_1}{\sigma_v^2}$$

where σ_v^2 is the noise variance and λ_1 is the largest eigenvalue of the sample covariance matrix. Additionally, a GLRT was calculated when the noise variance was unknown, the detection statistic for which is given by

$$T_{GLRT} \triangleq \frac{\lambda_1}{\frac{1}{M}\mathrm{Tr}(\hat{\mathbf{R}})}$$

In [39] a maximum-minimum eigenvalue detector is derived by considering the effects of the presence of a signal on the eigenvalues of the sample covariance matrix. The detection statistic was seen to be

$$T_{MME} = \frac{\lambda_{\max}}{\lambda_{\min}}$$

The covariance of the matrix formed by the transmitted signal being sent through the linear channel $\mathbb{H}$ is denoted as $\mathbb{H}\mathbf{R}_s\mathbb{H}$ with eigenvalues $\hat{\rho}$. The detector was formed

by considering the eigenvalues of $\hat{\mathbf{R}}$ to be the sum of the $\hat{\rho}$ and the noise variance, thus, $\hat{\lambda}_{\max} = \hat{\rho}_{\max} + \sigma_n^2$. A similar expression was derived for $\lambda_{\min}$. It can be seen that $\hat{\rho}_{\max} = \hat{\rho}_{\min}$ only when $\mathbb{H}\mathbf{R}_s\mathbb{H} = \delta\mathbb{I}_{ML}$ with $\delta > 0$. However, when a signal is present, this scenario is very unlikely to occur. Thus, if there is no signal, the detector evaluates to one, and when a signal is present, the ratio is greater than one.

In [40], a detector was determined by optimally combining the received samples in space and time on the principles of maximizing SNR. The blindly combined energy detector is designed to work solely on the received data and is given by

$$T_{BCED}(N) = \frac{1}{N}\sum_{n=0}^{N-1} |\tilde{z}(n)|^2 = \hat{\lambda}_{\max}(N)$$

where $\tilde{z}(n) = \tilde{\beta}^T x(n)$ and $\tilde{\beta}$ is chosen to maximize the SNR cost function, and is found to be the normalized eigenvector corresponding the largest eigenvalue of $\hat{\mathbf{R}}$.

In the case of an unknown primary signal covariance and noise variance, the GLR was found to be the ratio of the arithmetic mean to the geometric mean of the eigenvalues, and is given in [41] as

$$T_{AGM}(\boldsymbol{\lambda}) = \frac{\frac{1}{M}\sum_m \lambda_m}{(\prod_m \lambda_m)^{\frac{1}{M}}}$$

The detector depends solely on the eigenvalues of the sample covariance matrix, $\hat{\mathbf{R}}$, and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_M]$.

The GLR for the scenario when the noise variance is known and the primary signal covariance is unknown is shown in [41] to be

$$T_{SSE}(\boldsymbol{\lambda}) = \frac{Nm'}{2}\left[\frac{\mathrm{AM}\left(\boldsymbol{\lambda}^s\right)}{\sigma^2} - \ln\left(\frac{\mathrm{GM}\left(\boldsymbol{\lambda}^s\right)}{\sigma^2}\right) - 1\right] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where AM and GM denote the arithmetic and geometric means over the elements in a vector $\mathbf{x}$ respectively, m' corresponds to the largest m such that $\lambda_m > \sigma^2$, and $\boldsymbol{\lambda}^s = [\lambda_1, \ldots, \lambda_{m'}]$ denotes the vector of signal-subspace eigenvalues of $\hat{\mathbf{R}}$ in decreasing order.

Also of interest are methods estimating of the rank of a signal. The maximum likelihood (ML) estimate for rank was given in [42] to be

$$\hat{K} = \underset{K\in\{0,\ldots,M-1\}}{\arg\min}\; MN \log\left(1 - \frac{\sum_{j=1}^{K}\lambda_j}{\mathrm{Tr}(\mathbf{W})}\right)$$

where again, M is the number of sensors, N is the number of samples and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ are the eigenvalues of $\mathbf{W}$. However, as K increases, the hypothesis corresponding to the rank of that K becomes increasingly likely compared to H_0.

In [43], the Bayesian information criterion (BIC) was used to mitigate this shortcoming of the ML estimator by penalizing the more complex models corresponding to larger K. This gives an estimate of the form

$$\hat{K} = \underset{K\in\{0,\ldots,M-1\}}{\arg\min}\; \frac{N\mathrm{Tr}(\mathbf{W})}{\sigma^2}\left(1 - \sum_{j=1}^{K}\frac{\lambda_j}{\mathrm{Tr}(\mathbf{W})}\right) + (K(N+N) - K^2)\log N$$

when the noise variance is known. For an unknown noise variance the BIC-penalized ML rank estimator takes the form

$$\hat{K} = \underset{K\in\{0,\ldots,M-1\}}{\arg\min}\; MN \log\left(1 - \sum_{j=1}^{K}\frac{\lambda_j}{\mathrm{Tr}(\mathbf{W})}\right) + (K(M+N) - K^2 + \frac{1}{2})\log(N)$$

and the penalty function for the Bayesian information criterion is

$$L(\nu(K), N) = \frac{\nu(K)}{2}\log N$$

where $\nu(K)$ is the number of parameters.

Additionally the maximum *a posteriori* estimate when the noise variance is known is the value of K that gives the maximum value for the posterior density given by

$$p(K|X,\sigma^2) = \frac{C}{\mathrm{vol}(G_{K,N})}\left(\frac{\pi\sigma^2}{N}\right)^{K(N-K)} e^{\frac{N}{\sigma^2}\sum_{i=1}^{K}\lambda_i}\prod_{i=1}^{K}\prod_{j=1}^{N-K}(\lambda_i - \lambda_{K+j} + \sigma^2)^{-1}$$

where $C = p(K = 0 \,|\, X, \sigma^2)$. The MAP estimate for unknown noise variance was also derived and is shown in detail in Chapter 4.

Chapter 3

SCOPE

Determining detectors for radar purposes has seen a renewed interest in the past few years due to the rise in the use of MIMO systems as well as growing interest in applications such as passive radar. Similarly the past decade has seen the rise of interest in spectrum sensing due to the limits of the physical spectrum available for use. The similarities between sensing to determine if a primary user signal is present and multiple channel detection and estimation for radar/sonar applications can be seen by considering sample problem statements for the different applications.

In the detection problem [40], for example, the signal model for a multi-antenna spectrum sensing scenario is given as

$$\begin{aligned} H_0 &: x_m(n) = \eta_m(n) \\ H_1 &: x_m(n) = s_m(n) + \eta_m(n), \\ & m = 1, \ldots, M \end{aligned} \tag{3.1}$$

where $M \geq 1$ represents the number of antennas at the receiver, η_m is the noise and $s_m(n) = \sum_{k=1}^{K} \sum_{l=0}^{q_{mk}} h_{mk}(l)\tilde{s}_k(n-l)$ is the signal received by antenna m with K being the number of primary user/antenna signals, $\tilde{s}_k(n)$ being the transmitted signal from primary user/antenna k, $h_{mk}(l)$ the impulse response of the propagation channel from user k to receiver m and q_{mk} the channel order. The noise samples $\eta_m(n)$ are assumed iid across both n and m.

In [36], the system model to detect the presence of a rank-K emitter with $M > K$

spatially distributed sensors is given as

$$\begin{aligned} H_0 &: \mathbf{X} = \boldsymbol{\nu} \\ H_1 &: \mathbf{X} = \mathbf{AS} + \boldsymbol{\nu}, \end{aligned} \tag{3.2}$$

where a K-dimensional signal subspace is defined by an unknown $N \times K$ complex matrix $\mathbf{S}$ whose columns are orthonormal vectors in $\mathbb{C}^N$, and the element a_{km} of the unknown $K \times M$ complex matrix $\mathbf{A}$ is the complex amplitude of the component of the signal received at sensor m and in the subspace corresponding to the k^{th} column of $\mathbf{S}$. The noise $\boldsymbol{\nu}$ is assumed to be normally distributed with zero mean and both spatially and temporally white with covariance matrix $\sigma^2 \mathbb{I}_{NM}$.

It can be seen that both applications entail very similar multiple channel detection problems, and as such detectors and their properties may be applicable to either application. For instance, the geometrical nature of tests developed in [36, 43] and related work leads to estimators of pertinent signal structure, such as rank, and the occupied subspace. With this insight, one can consider estimating the unoccupied subspace in settings where multiple access is not based on frequency division, opening the possibility of opportunistic use of unoccupied communication resources that are not necessarily defined by spectral bands.

Chapter 4

MATHEMATICAL FORMULATION

To estimate the number of users K in a CDMA system, the detection scheme is adapted from [43], which draws on work completed in [31], [36] and [44]. To formulate the maximum *a posteriori* estimator for rank the problem definition becomes a multi-hypothesis test modified from (3.2)

$$\begin{aligned} H_0 &: \mathbf{X} = \boldsymbol{\nu} \\ H_K &: \mathbf{X} = \mathbf{A}_K \mathbf{S}_K + \boldsymbol{\nu}, \end{aligned} \tag{4.1}$$

for $K = 1, \ldots, M-1$. $\mathbf{S}_K$ denotes the complex $N \times K$ matrix that defines the K-dimensional signal subspace, A_K is the $K \times M$ matrix of complex amplitudes, and the $N \times M$ noise matrix $\boldsymbol{\nu}$ is assumed to be normally distributed with covariance matrix $\sigma^2 \mathbb{I}_{NM}$. The joint probability density function (pdf) of $\mathbf{X}$ conditioned on σ^2 under H_0 is given by

$$p\left(\mathbf{X}|H_0, \sigma^2\right) = \left(\pi\sigma^2\right)^{-MN} e^{\frac{N}{\sigma^2}\mathrm{Tr}(\mathbf{W})} \tag{4.2}$$

The joint pdf of $\mathbf{X}$ conditioned on $\mathbf{A}_K$, $\mathbf{S}_K$, and σ^2 is

$$p\left(\mathbf{X}|H_1, \mathbf{S}_K, \mathbf{A}_K, \sigma^2\right) = \left(\pi\sigma^2\right)^{-MN} e^{-\frac{N}{\sigma^2}\mathrm{Tr}(\mathbf{W})} e^{-\frac{1}{\sigma^2}\mathrm{Tr}\left(\left(\mathbf{A}_K\mathbf{A}_K^\dagger - \mathbf{A}_K\mathbf{S}_K\mathbf{X}^\dagger - \mathbf{X}\mathbf{S}_K^\dagger\mathbf{A}_K^\dagger\right)\right)} \tag{4.3}$$

with $K > 0$.

The matrix $\mathbf{P} = \mathbf{S}_K^\dagger \mathbf{S}_K$ and is an $N \times N$, rank-K orthogonal projection matrix into subspace V which is spanned by the rows of $\mathbf{S}_K$. This definition is used in order to obtain a non-redundant model parameterization, and it is possible to associate a unique choice of $\mathbf{S}_K$ with each $\mathbf{P}$. Since $\mathbf{P}$ is a projection matrix, it satisfies the

conditions $\mathbf{P} = \mathbf{P}^\dagger$, $\mathbf{P}^2 = \mathbf{P}$ and $\mathrm{Tr}(\mathbf{P}) = K$. The collection of all orthogonal projection matrices constitute the complex Grassmannian $G_{K,N}$, which has complex dimension $K(N-K)$.

The likelihood function (4.3) is invariant under the transformations

$$\mathbf{X} \to \mu\mathbf{U}\mathbf{X}\mathbf{L},\ \mathbf{A} \to \mu\mathbf{U}\mathbf{A},\ \mathbf{S} \to \mathbf{S}\mathbf{L},\ \text{and } \sigma \to \mu\sigma \tag{4.4}$$

where $\mathbf{U}$ and $\mathbf{L}$ are unitary matrices of dimensions $M \times M$ and $N \times N$ respectively, and $\mu > 0$.

4.0.1 Prior Distributions

In order to determine the estimate, priors for $\mathbf{A}$, $\mathbf{P}$ and σ^2 are taken to be as non-informative as possible, and invariant under the transforms given in (4.4). The invariant non-informative prior measure on the space of unknown parameters is $d\mathbf{A}d\sigma^{-2}d\mu\mathbf{P}$, where $d\mu(\mathbf{P})$ is the normalized invariant measure on $G_{K,N}$, and $d\mathbf{A}$ is the Lebesgue measure in $\mathbb{C}^{MK}$. However, this prior is not proper. To this end, proper priors are found by approaching the non-informative prior in an appropriate limiting sense. Thus the prior is taken to have the form

$$p(K, \mathbf{A}, \sigma^2)d\mathbf{A}d\sigma^{-2}d\mu(\mathbf{P}) = p(K|\beta^2)p(\mathbf{A}|K, \sigma^2, \beta^2)p(\sigma^{-2}|\tau)d\mathbf{A}d\sigma^{-2}d\mu(\mathbf{P})$$

and the components are assigned as follows. The prior for $\mathbf{A}$ is chosen to be

$$p\left(\mathbf{A}|K, \sigma^2, \beta^2\right) = \left(\pi\beta^2\sigma^2\right)^{-MK} e^{\frac{1}{\beta^2\sigma^2}\mathrm{Tr}(\mathbf{A}\mathbf{A}^\dagger)}$$

which becomes less informative as $\beta^2 \to \infty$. The prior for σ^2 is taken to be the maximum entropy prior

$$p\left(\sigma^{-2}|\tau\right) = \tau^M e^{-\tau M\sigma^{-2}}$$

which becomes less informative as $\tau \to 0$. Additionally, the prior used for K is

$$p(K|\beta^2) = \frac{(1+\beta^2)^{MK}}{\sum_{K=0}^{M-1}(1+\beta^2)^{MK}}$$

which ensures that as the prior for $\mathbf{A}$ becomes less informative, the posterior ratio for any two ranks K and K' approaches a finite non-zero limit. Otherwise, the hypothesis H_K with the smallest value of K would dominate regardless of the data.

To determine the invariant measure on $G_{K,N}$ it is necessary to parameterize $\mathbf{P}$ in local coordinates on $G_{K,N}$. As there can be many matrices $\mathbf{S}_K$ that can represent the same point on $G_{K,N}$, the orthogonal projector $\mathbf{P}$ is used. Thus, $\mathbf{P}$ is constructed such that $V \to \mathbf{P}$ is one-to-one and onto. This is achieved by writing $\mathbf{S}_K$ as

$$\mathbf{S}_K = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 \end{bmatrix}$$

where $\mathbf{S}_1$ is $K \times K$, and it is assumed that a basis is chosen for which $\mathbf{S}_1$ is invertible. $\mathbf{S}_2$ is then a $K \times (N-K)$ matrix. $\mathbf{S}_K$ can then be written as

$$\begin{aligned} \mathbf{S}_K &= \mathbf{S}_1 \begin{bmatrix} \mathbb{I}_K & \mathbf{S}_1^{-1}\mathbf{S}_2 \end{bmatrix} \\ &= \mathbf{S}_1 \mathbf{T} \end{aligned}$$

where $\mathbf{T}$ is standard form and completely characterized by $\mathbf{S}_1^{-1}\mathbf{S}_2$. The subspace V is specified uniquely by the $K \times (N-K)$ matrix $\mathbf{S}_1^{-1}\mathbf{S}_2 = \mathbf{Z}^\dagger$. Denoting

$$\mathbf{S}_z = (\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z})^{-\frac{1}{2}} \begin{bmatrix} \mathbb{I}_K & \mathbf{Z}^\dagger \end{bmatrix}$$

the projection can be written as the block matrix

$$\begin{aligned} \mathbf{P} &= \mathbf{S}_z^\dagger \mathbf{S}_z \\ &= \begin{bmatrix} \mathbb{I}_k \\ \mathbf{Z} \end{bmatrix} (\mathbb{I}_K + (\mathbf{Z}^\dagger \mathbf{Z})^{-\frac{1}{2}} (\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z})^{-\frac{1}{2}} \begin{bmatrix} \mathbb{I}_k & \mathbf{Z}^\dagger \end{bmatrix} \\ &= \begin{pmatrix} \left(\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} & \left(\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} \mathbf{Z}^\dagger \\ \mathbf{Z}\left(\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} & \mathbf{Z}\left(\mathbb{I}_K + \mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} \mathbf{Z}^\dagger \end{pmatrix} \end{aligned}$$

Thus, there can be one-to-one mapping between the entities

$$\text{Matrices } \mathbf{Z}^\dagger_{K\times(N-K)} \leftrightarrow \text{Subspaces } V \leftrightarrow \text{Projectors } \mathbf{P} \leftrightarrow \text{Elements of } G_{K,N}$$

Through the use of differential forms following [45], the normalized invariant measure on the Grassmanian $G_{K,N}$ is shown in [36] to be

$$d\mu(\mathbf{P}) = \frac{1}{\text{vol}(G_{K,N})}\det\left(\mathbb{I}_K - \mathbf{Z}^\dagger\mathbf{Z}\right)^{-N} d\mathbf{Z}$$

where

$$d\mathbf{Z} = \prod_{i=1}^{N-K}\prod_{j=1}^{K} d\text{Re}(z_{ij})d\text{Im}(z_{ij)}$$

and the volume of the Grassmanian is

$$\text{vol}(G_{K,N}) = \frac{\prod_{n=N-K+1}^{N} A_{2n-1}}{\prod_{n=1}^{K} A_{2n-1}}$$

where A_n is the area of the unit sphere in $\mathbb{R}^n$

$$A_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}$$

and Γ denotes the gamma function.

4.0.2 MAP Estimate

Assuming σ^2 is unknown, the marginalized likelihoods are

$$p(\mathbf{X}|K=0) = \frac{\tau^M\Gamma(p)\text{Tr}\left(\tilde{\mathbf{W}}\right)^{-l}}{N^l\pi^{MN}}$$

and

$$p\left(\mathbf{X}|K,\beta^2,\tau\right) = \frac{p\left(\mathbf{X}|K=0\right)}{\left(1+\beta^2\right)^{MK}}\int_{G_{K_N}}\left(1-\alpha\frac{\text{Tr}\left(\mathbf{WP}\right)}{\text{Tr}\left(\tilde{\mathbf{W}}\right)}\right)^{-l}\text{d}\mu(\mathbf{P})$$

where $l = MN+1$ and $\tilde{\mathbf{W}} = \mathbf{W} + \frac{M\tau}{N^2}\mathbb{I}_N$.

Using the results obtained in section 4.0.1, the integral can be rewritten as

$$\begin{aligned} p(\mathbf{X}|K,\beta^2,\tau) &= \frac{p(\mathbf{X}|K=0)}{(1+\beta^2)^{MK}\text{vol}(G_{K,N})}\int_{\mathbf{Z}\in\mathbb{C}^{(N-K)\times K}}\left(1-\alpha\frac{\text{Tr}\left(\mathbf{WP}\right)}{\text{Tr}\left(\tilde{\mathbf{W}}\right)}\right)^{-l}\det\left(\mathbb{I}_K-\mathbf{Z}^\dagger\mathbf{Z}\right)^{-N}\text{d}\mathbf{Z} \\ &= \frac{p(\mathbf{X}|K=0)}{(1+\beta^2)^{MK}\text{vol}(G_{K,N})}\int_{\mathbf{Z}} e^{-l\log\left(1-\alpha\frac{\text{Tr}(\mathbf{WP})}{\text{Tr}(\tilde{\mathbf{W}})}\right)}e^{-N\log\det\left(\mathbb{I}_K-\mathbf{Z}^\dagger\mathbf{Z}\right)}\text{d}\mathbf{Z} \end{aligned}$$

As $\beta^2 \to \infty$ and $\tau \to 0$, the integrals can be approximated using Laplace approximation and some matrix identities. The posteriors are found to be

$$p(K = 0|\mathbf{X}) = C$$

and

$$p(K|\mathbf{X}) = C\frac{1}{\text{vol}(G_{K,N})}\left(\frac{\pi}{p}\right)^{K(N-K)}\gamma^{K(N-K)-p}\prod_{i=1}^{K}\prod_{j=1}^{N-K}\left(\tilde{\lambda}_i - \tilde{\lambda}_{K+j} + \frac{N\gamma}{p}\right)^{-1} \tag{4.5}$$

where $p = M(N-K)+1$, $\gamma = (1 - \sum_{i=1}^{K}\tilde{\lambda}_i)$ where in the limit $\tilde{\lambda}_i = \frac{\lambda_i}{\text{Tr}(\mathbf{W})}$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ are the eigenvalues of $\mathbf{W}$.

For this problem, assuming the noise variance σ^2 is unknown, the MAP estimate of K is

$$\hat{K} = \arg\max_{K} p(K|\mathbf{X}) \tag{4.6}$$

for which the computation of C is unnecessary as it is constant for all values of K.

Chapter 5

RESULTS

To evaluate the performance of the MAP described in Chapter 4, a simple CDMA system was simulated and the MAP detector (4.6) was applied to estimate the number of users in the CDMA system. Additionally, an estimate for the orthogonal projector of the occupied subspace is shown.

5.1 Simulation Parameters

A K user CDMA system was simulated by generating K QPSK signals and encoding them on K PN codes of length N. M denotes the number of time epochs that are taken, and $K < M < N$. The K code vectors are normalized to unit length and are the rows of $\mathbf{S}_{K\times N}$ in equation (4.5). The m^{th} row of matrix $\mathbf{A}_{M\times K}$ corresponds to the m^{th} measurement epoch, and its elements are determined by the QPSK sequence elements q_{mk} belonging to each of the K users during that epoch, with $q_{mk} \in \{1 \ -1 \ i \ -i\}$. The diversity needed to obtain full rank is provided by the presence of the same codes in different linear combinations in the respective data segments. The noise is assumed to be zero mean white complex Gaussian, $\boldsymbol{\nu}_{M\times N}$.

5.2 Estimation of Rank

Using the setup described above, the MAP estimate (4.5) was calculated for $k = 1, \ldots, M-1$ without exploiting knowledge of $\mathbf{S}, \mathbf{A}$, or $\boldsymbol{\nu}$ beyond what was described. The value of k yielding the maximum posterior probability was taken as the estimate of the number of users in the CDMA system. The simulation was performed for various combinations of actual rank K, number of time epochs M, length of the code

N, and SNR values at the receiver, which were identical and held constant for each user.

A system with PN codes of length $N = 100$ and utilizing $M = 95$ time epochs was simulated for $K = 30, 50, 70, 90$ users. It can be seen in figure 5.1 that when the actual value of K is well below the limit of $M - 1$, the MAP estimator does an excellent job correctly estimating the correct rank. However, as the value of the actual K approached the maximum value, the performance degrades markedly.

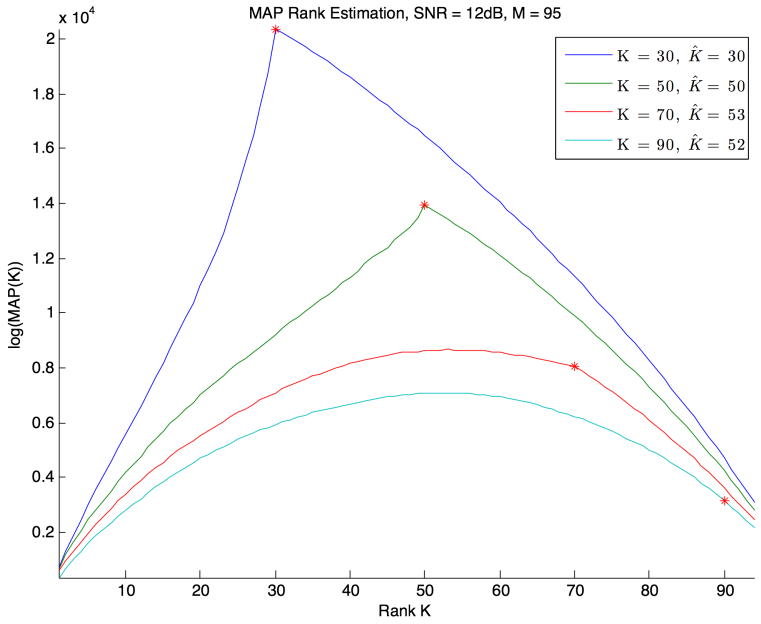

Figure 5.1: MAP Estimate Results for $M = 95$, $N = 100$, $K = 30, 50, 70, 90$ for the Blue, Green, Red and Cyan Lines Respectively, with an SNR at the Receiver of 12dB. The Red Asterisk Represents the Point Where the True Rank Falls on the MAP.

By increasing the SNR, the rank can be correctly estimated closer to the limit of M. This can be seen by keeping the same values for K, M and N as before, but increasing the SNR at the receiver to 24dB, it can be seen in figure 5.2 that the estimate is improved closer to the limit, however, at a certain point, the detector cannot estimate the number of users correctly regardless of the SNR. At this point, it is only possible to obtain a correct estimation by increasing the number of time segments,

and accordingly the length of the codes. While this can increase the computation time, a correct estimate can then be obtained with a lower SNR, as seen in figure 5.3.

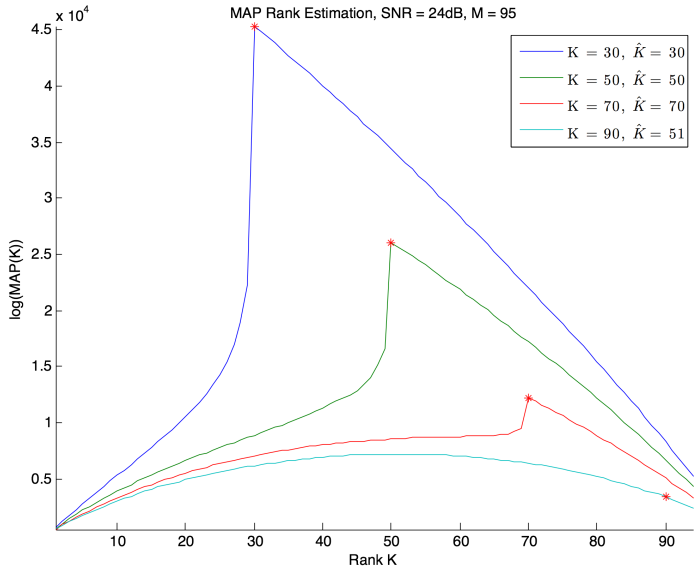

Figure 5.2: MAP Estimate Results for $M = 95$, $N = 100$, $K = 30, 50, 70, 90$ for the Blue, Green, Red and Cyan Lines Respectively, with an SNR at the Receiver of 24dB. The Red Asterisk Represents the Point Where the True Rank Falls on the MAP.

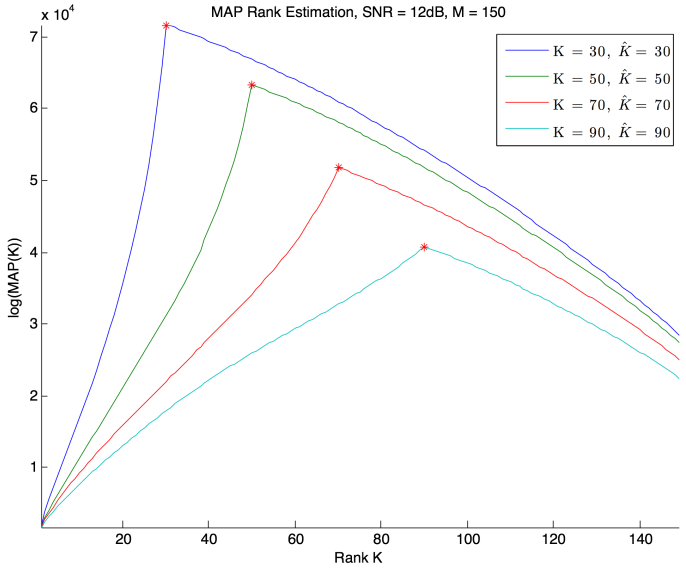

Figure 5.3: MAP Estimate Results for $M = 150$, $N = 200$, $K = 30, 50, 70, 90$ for the Blue, Green, Red and Cyan Lines Respectively, with an SNR at the Receiver of 12dB. The Red Asterisk Represents the Point Where the True Rank Falls on the MAP.

However, if the number of time epochs M is not approaching its limit of $N-1$, then desired performance can be provided solely with an increase in SNR, as shown below in figure 5.4. This is compared to the effect of changing the SNR when the value of M is close to the limit, shown in figure 5.5, where even extremely high SNRs cannot resolve the estimate.

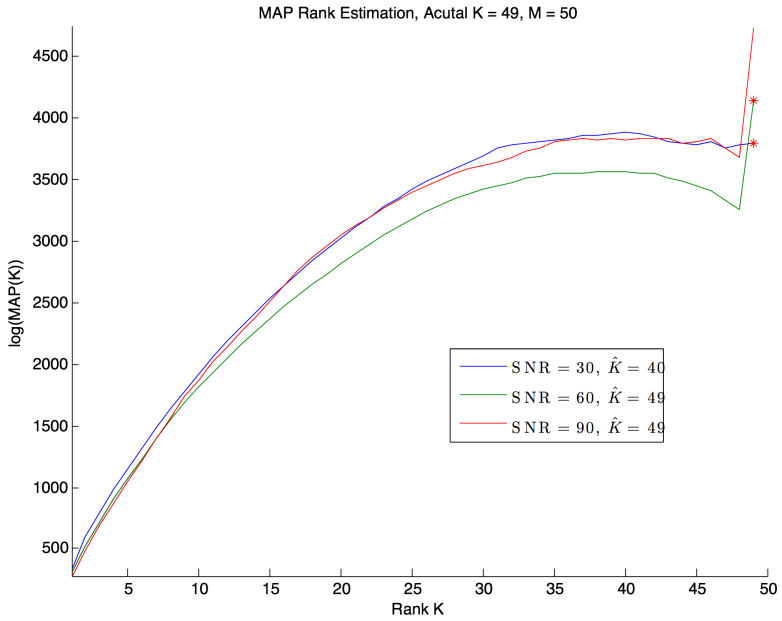

Figure 5.4: MAP Estimate Results for $M = 50$, $N = 100$, $K = 49$ and the SNR at the Receiver is Varied to Be 12dB, 24dB, 36dB and 60dB for the Blue, Green, Red and Cyan Lines Respectively. The Red Asterisk Represents the Point Where the True Rank Falls on the MAP.

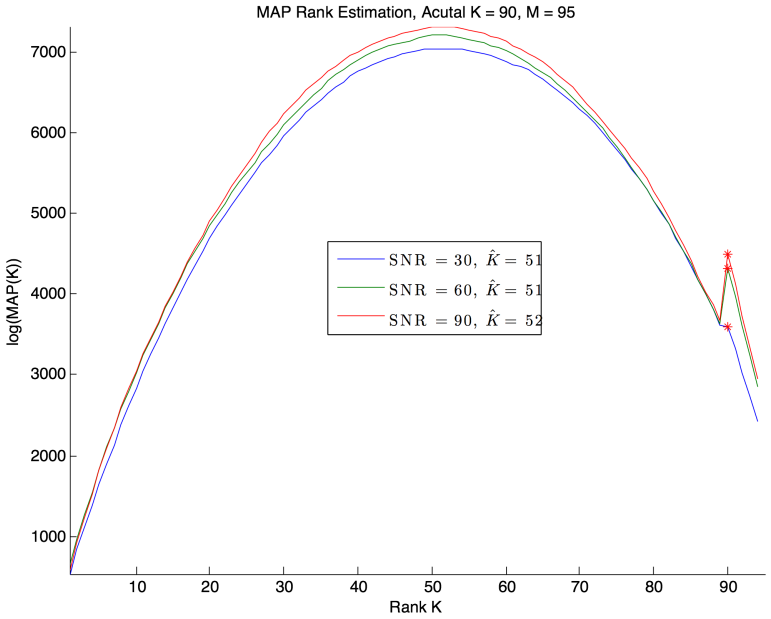

Figure 5.5: MAP Estimate Results for $M = 95$, $N = 100$, $K = 90$ and the SNR at the Receiver is Varied to be 12dB, 24dB, 36dB and 60dB for the Blue, Green, Red and Cyan Lines Respectively. The Red Asterisk Represents the Point Where the True Rank Falls on the MAP.

5.3 Estimation of the Subspace

The likelihood function corresponding to the H_1 signal model (4.1) with known noise variance σ^2 and signal rank K is

$$p(\mathbf{X}|H_1, \mathbf{A}, \mathbf{S}, \sigma^2) = \left(\pi\sigma^2\right)^{-MN} e^{-\frac{N}{\sigma^2}\mathrm{Tr}(\mathbf{W})} e^{-\frac{1}{\sigma^2}\mathrm{Tr}\left((\mathbf{A}-\mathbf{X}\mathbf{S}^\dagger)(\mathbf{A}-\mathbf{X}\mathbf{S}^\dagger)^\dagger - \mathbf{X}\mathbf{S}^\dagger\mathbf{S}\mathbf{X}^\dagger\right)} \tag{5.1}$$

Maximizing the function with respect to $\mathbf{A}$ can be achieved by minimizing $\left(\mathbf{A} - \mathbf{X}\mathbf{S}^\dagger\right)\left(\mathbf{A} - \mathbf{X}\mathbf{S}^\dagger\right)^\dagger$, thus the estimate is obtained to be $\hat{\mathbf{A}} = \mathbf{X}\mathbf{S}^\dagger$. The estimate for the projector is found by substituting $\hat{\mathbf{A}}$ for $\mathbf{A}$ and is the value of $\mathbf{P}$ that maximizes the function in (5.1). The Schur-Horn theorem [46] is given in [36] to be

$$d_H = \sum_\rho a_\rho \rho(\Lambda), 0 \leq a_\rho \leq 1, \sum_\rho a_\rho = 1$$

where d_H is the vector of diagonal elements of a Hermitian matrix $\mathbf{H}$ that lies in the convex hull of all permutations $\rho(\Lambda)$ of $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)$ which is the matrix

of non-increasing eigenvalues of $\mathbf{H}$. By using the Schur-Horn theorem to maximize $\mathbf{P} = \mathbf{S}_K^\dagger \mathbf{S}_K$ over the Grassmanian $G_{K,N}$ as in [44], the estimate for the orthogonal projector is found to be

$$\hat{\mathbf{P}} = \sum_{k=1}^{K} \mathbf{v}_k \mathbf{v}_k^\dagger \tag{5.2}$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_K$ are the normalized eigenvectors of $\mathbf{W}$ corresponding to its K largest eigenvalues. The estimated occupied subspace is uniquely specified by the orthogonal projector $\hat{\mathbf{P}}$ and from this, the estimation of the $(N-K)$-dimensional orthogonal complement can be found.

Note that the K value in the estimate is the value of the actual rank. In practice, the estimate for K would have to be used provided the real K is unknown. In such a case, the accuracy of the estimate of the orthogonal projector of subspace would be dependent upon the accuracy of the estimate of the rank.

The ability to obtain this subspace is useful, as the codes for the secondary users can be generated to be in the unoccupied subspace, and thus will be orthogonal to the codes of the primary users, theoretically resulting in less interference.

Chapter 6

NOTES ON IMPLEMENTATION

While this paper shows how the MAP can be applied to a spectrum sensing problem, the MAP estimate does not yield correct estimates when the SNR is not sufficiently high. However, in spectrum sensing applications, the SNR can be very low, with it being necessary to detect signals when the SNR is as lower than -20dB as described in [39].

Additionally, it would be beneficial to run simulations showing the difference in transmission error of transmission when additional users are added to the CDMA system using random PN codes of length N as opposed to using PN codes that are known to be based in the unoccupied subspace. However, when determining the occupied subspace, the synchronicity of the system must be considered. In order to obtain the correct location of the occupied subspace (and by extension its orthogonal complement) the opportunistic user may need to know the location of the transmitters of the primary users.

Chapter 7

CONCLUSIONS

Determining the presence of an unknown signal can be useful in a variety of applications. To this end, a summary of the current work in spectrum sensing and radar/sonar applications was provided, with focus on detectors based on the eigenvalues of the Gram matrix. Additionally, it was shown that the detection problem posed by spectrum sensing and radar/sonar is similar, and thus detectors and properties found for either application can be successfully applied in the other.

To illustrate this concept, an example was completed using the MAP estimator for rank to predict the number of users in a simple CDMA system. It was seen that the detector worked well when estimating the number of users under certain parameters, mainly that the signal had a high SNR and that the values were well within their limits. However, as the number of users increases, the number of time epochs and correspondingly, the length of the codes must increase as well to obtain adequate performance. This can result in added computational time.

Additionally, an estimator for the orthogonal projector of the occupied subspace was shown. From this projector, it is possible to determine the associated occupied subspace, and by extension the orthogonal complement of the occupied subspace. By selecting codes known to be in the unoccupied subspace, and thus orthogonal to the codes in use, it is theoretically possible to reduce the interference caused by secondary users.

REFERENCES

[1] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, Nov 2004, pp. 772–776 Vol.1.

[2] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, First Quarter 2009.

[3] Y.-C. Liang, K.-C. Chen, G. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3386–3407, Sept 2011.

[4] A. Dandawate and G. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2355–2369, Sep 1994.

[5] ——, "Nonparametric polyspectral estimators for kth-order (almost) cyclostationary processes," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 67–84, Jan 1994.

[6] S. Enserink and D. Cochran, "A cyclostationary feature detector," in *Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 2, Oct 1994, pp. 806–810 vol.2.

[7] ——, "On detection of cyclostationary signals," in *International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, May 1995, pp. 2004–2007 vol.3.

[8] D. Ramírez, L. Scharf, J. Vía, I. Santamaría, and P. Schreier, "An asymptotic GLRT for the detection of cyclostationary signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2014, pp. 3415–3419.

[9] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, Jan 2010.

[10] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[11] J. Ma, G. Li, and B.-H. Juang, "Signal processing in cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[12] S. Haykin, D. Thomson, and J. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[13] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2400–2419, April 2011.

[14] J. Palmer, S. Palumbo, A. Summers, D. Merretta, S. Searle, and S. Howard, "An overview of an illuminator of opportunity passive radar research project and its signal processing research directions," *Digital Signal Processing*, vol. 21, no. 5, pp. 593–599, September 2011.

[15] K. Bialkowski, I. Clarkson, and S. Howard, "Generalized canonical correlation for passive multistatic radar detection," in *IEEE Statistical Signal Processing Workshop*, June 2011, pp. 417–420.

[16] D. Hack, L. Patton, B. Himed, and M. Saville, "Centralized passive MIMO radar detection without direct-path reference signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 3013–3023, June 2014.

[17] ——, "Detection in passive MIMO radar networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2999–3012, June 2014.

[18] D. Hack, C. Rossler, and L. Patton, "Multichannel detection of an unknown rank-N signal using uncalibrated receivers," *IEEE Signal Processing Letters*, vol. 21, no. 8, pp. 998–1002, Aug 2014.

[19] P. J. Davis, *Interpolation and Approximation*. Dover Publications, 1963.

[20] S. Howard, S. Sirianunpiboon, and D. Cochran, "Invariance of the distributions of normalized gram matrices," in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, June 2014, pp. 352–355.

[21] K. Beaudet and D. Cochran, "Multiple-channel detection in active sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 3910–3914.

[22] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. pp. 152–177, March 1963.

[23] A. Zanella, M. Chiani, and M. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Transactions on Communications*, vol. 57, no. 4, pp. 1050–1060, April 2009.

[24] A. Edelman, "Eigenvalues and condition numbers of random matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 9, no. 4, pp. 543–560, 1988.

[25] L. Wei, O. Tirkkonen, K. D. P. Dharmawansa, and M. R. McKay, "On the exact distribution of the scaled largest eigenvalue," *CoRR*, 2012.

[26] A. Kortun, T. Ratnarajah, M. Sellathurai, C. Zhong, and C. Papadias, "On the performance of eigenvalue-based cooperative spectrum sensing for cognitive radio," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 49–55, Feb 2011.

[27] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise level," in *IEEE International Conference on Communications*, June 2011, pp. 1–5.

[28] A. Nuttall, "Invariance of distribution of coherence estimate to second-channel statistics," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 1, pp. 120–122, Feb 1981.

[29] H. Gish and D. Cochran, "Invariance of the magnitude-squared coherence estimate with respect to second-channel statistics," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, no. 12, pp. 1774–1776, Dec 1987.

[30] ——, "Generalized coherence," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Apr 1988, pp. 2745–2748 vol.5.

[31] S. Sirianunpiboon, S. Howard, and D. Cochran, "A Bayesian derivation of generalized coherence detectors," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2012, pp. 3253–3256.

[32] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2049–2057, Sep 1995.

[33] D. Cochran and H. Gish, "Multiple-channel detection using generalized coherence," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Apr 1990, pp. 2883–2886 vol.5.

[34] A. Clausen and D. Cochran, "Asymptotic analysis of the generalized coherence estimate," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 45–53, Jan 2001.

[35] ——, "An invariance property of the generalized coherence estimate," *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 1065–1067, Apr 1997.

[36] S. Sirianunpiboon, S. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6536–6540.

[37] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3764–3774, Aug 2011.

[38] D. Ramírez, J. Iscar, J. Vía, I. Santamaría, and L. Scharf, "The locally most powerful invariant test for detecting a rank-P Gaussian signal in white noise," in *Sensor Array and Multichannel Signal Processing Workshop, 2012*, June 2012, pp. 493–496.

[39] Y. Zeng and Y.-C. Liang, "Maximum-minimum eigenvalue detection for cognitive radio," in *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2007, pp. 1–5.

[40] Y. Zeng, Y.-C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Processing Letters*, vol. 15, pp. 649–652, 2008.

[41] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 84–88, January 2010.

[42] S. Sirianunpiboon, D. Cochran, and S. Howard, "Invariant detection and estimation for MIMO radar signals," in *IEEE Radar Conference*, May 2014, pp. 1203–1208.

[43] S. Sirianunpiboon, S. Howard, and D. Cochran, "Maximum a posteriori estimation of signal rank," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 5676–5680.

[44] S. Howard, S. Sirianunpiboon, and D. Cochran, "Detection and characterization of MIMO radar signals," in *International Conference on Radar*, Sept 2013, pp. 330–334.

[45] A. T. James, "Normal multivariate analysis and the orthogonal group," *Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, March 1954.

[46] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix," *American Journal of Mathematics*, vol. 76, no. 3, pp. 620–630, Jul 1954.