Theoretical Receiver Operating Characteristics of Two-Stage Change Detector for Synthetic Aperture Radar Images

by

Akshay Sunil Bondre

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

Approved April 2020 by the
Graduate Supervisory Committee:

Christ Richmond, Chair
Antonia Papandreou-Suppappola
Daniel Bliss

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

Detecting areas of change between two synthetic aperture radar (SAR) images of the same scene, taken at different times is generally performed using two approaches. Non-coherent change detection is performed using the sample variance ratio detector, and displays a good performance in detecting areas of significant changes. Coherent change detection can be implemented using the classical coherence estimator, which does better at detecting subtle changes, like vehicle tracks. A two-stage detector was proposed by Cha *et al.*, where the sample variance ratio forms the first stage, and the second stage comprises of Berger's alternative coherence estimator.

A modification to the first stage of the two-stage detector is proposed in this study, which significantly simplifies the analysis of the this detector. Cha *et al.* have used a heuristic approach to determine the thresholds for this two-stage detector. In this study, the probability density function for the modified two-stage detector is derived, and using this probability density function, an approach for determining the thresholds for this two-dimensional detection problem has been proposed. The proposed method of threshold selection reveals an interesting behavior shown by the two-stage detector. With the help of theoretical receiver operating characteristic analysis, it is shown that the two-stage detector gives a better detection performance as compared to the other three detectors. However, the Berger's estimator proves to be a simpler alternative, since it gives only a slightly poorer performance as compared to the two-stage detector. All the four detectors have also been implemented on a SAR data set, and it is shown that the two-stage detector and the Berger's estimator generate images where the areas showing change are easily visible.

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor Dr. Christ Richmond, for mentoring me and providing invaluable insight throughout my research. I would like to thank my roommates and my friends in Tempe for their support throughout the past two years. I would also like to thank my parents Urmila Bondre and Sunil Bondre, and my grandmother Lata Bondre, who have always supported me unconditionally in every endeavor.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Chapter 1

INTRODUCTION

1.1 Motivation

Synthetic aperture radar (SAR) is the concept of an airborne or space-based radar generating high resolution images of the scene underneath, by processing the complex valued radar return signals from the scene. If we have two SAR return data sets of the same landscape, but taken at different times, an important application which arises is the detection of regions where the underlying scene has undergone a change. Change detection can be done by processing the two SAR data sets and generating an image where a black pixel denotes change and a white pixel denotes no change. This processing can be done by two methods, coherent change detection and non-coherent change detection. Non-coherent change detection compares the sample variance of the two corresponding locations, essentially comparing the intensity (magnitude squared) of the two locations, by forming the ratio of the two sample variances, and decides that a change has occurred if this sample variance ratio is not close to unity [1]. Coherent change detection determines change by considering an estimate of the coherence (which is the magnitude of the correlation coefficient) between the two radar returns corresponding to the same location, and declares change if the value of this estimate is significantly less than one. The sample coherence (also referred to as the classical coherence estimator) is generally used. The non-coherent change detector generally performs better at detecting significant changes like the displacement of a vehicle, while coherent change detection shows a better performance at detecting subtle changes in the scene [2],[3]. The ability of the classical coherent change

detector (classical CCD) to detect vehicle tracks has been demonstrated in [4]. In [3], theoretical receiver operating characteristics (ROC) for the sample variance ratio and the sample coherence detectors have been shown. A two-stage change detection method has been introduced in [2], wherein the first stage incorporates the sample variance ratio change detector, and the second stage implements a coherent change detector. The second stage uses the Berger's coherence estimator given in [5] instead of the classical coherence estimate, as it has been shown in [2] that the Berger's estimator performs better than the classical estimator if the true variances of the two corresponding points are almost equal. Since the sample variance ratio (first stage) compares the variances of the two locations, only those pairs of pixels which exhibit almost equal variance are passed to the second stage. Thus, the first stage processes the two SAR data sets to check for large scale changes, the pixel pairs which show no large scale change are passed to the second stage, and finally the second stage checks for subtle small scale changes between the two images. The two-stage detector determines that a location has undergone a change if either one of the stages detects a change.

All of the change detection methods mentioned above determine whether a pixel pair has undergone a change by comparing some quantity to a threshold value. Computing these thresholds in a way such that the error in detection performance is minimum is an important aspect of any detection algorithm. The method for computing thresholds for the sample variance ratio and the classical CCD have been demonstrated in [3]. For the two-stage change detector, a heuristic approach to determine the thresholds has been proposed in [2]. Also, using this heuristic approach, receiver operating characteristics have been generated empirically in [2], which suggest that the two-stage detector gives a better performance as compared to the sample variance ratio and the classical CCD methods.

1.2 Proposed Thesis Work

The main goal of this thesis is to generate the theoretical ROC curves for the two-stage detector in order to evaluate it's detection performance and to compare it with other change detection methods. In order to do this, an approach for selecting the thresholds using the probability density function (pdf) of the two-stage detector has been developed. Prior to dealing with these aspects of the detection performance of the two-stage detector, a modification to the structure of the two-stage detector has been proposed. The complex correlation coefficient between the two pixels acts as a nuisance parameter for the sample variance ratio detector [6]. When this correlation coefficient is assumed to be zero, this detector reduces to a simple F test. In this study, an argument is made that considering a non-zero correlation coefficient under the no change scenario can improve the detection performance for the sample variance ratio detector. The cumulative distribution function for this detector has also been derived in order to simplify the analysis. Furthermore, it has been shown that replacing the sample variance ratio detector in the first stage by the symmetric sample variance ratio significantly simplifies the analysis of the two-stage detector. This symmetric detector has been introduced in [7] in the context of edge detection. [1] and [3] have used this detector for change detection, but by assuming the correlation coefficient between the two pixels to be zero even under the no change scenario. The probability density function and the cumulative distribution function for this symmetric first stage detector with a non-zero correlation coefficient have also been derived in this study.

Considering the modifications mentioned above, the joint probability density function for this modified two-stage change detector has been derived. Based on this distribution, a procedure has been developed for computing the thresholds for the

two-stage change detector. This procedure introduces a new parameter α, on which the detection performance of the two-stage detector is dependent. This new parameter dictates the behavior of the two-stage detector, and for two specific values of α, the two-stage detector reduces to the Berger's coherence estimator and the sample variance ratio respectively. Detailed analysis of the dependence of the detection performance on α has been shown. The theoretical ROC curves for the two-stage detector are generated using the proposed method, and comparison of its detection performance with the sample variance ratio, classical CCD, and Berger's estimator have been shown. Additionally, these detectors are also compared to the log-likelihood ratio detector given in [3], which achieves the highest possible performance for any detector. Finally, the four detectors are applied to a SAR data set, and the images showing areas of change are generated for each of these detectors.

1.3 Thesis Organization

Chapter 2 contains a high-level overview of the concept of synthetic aperture radar. Some important concepts of detection theory have also been discussed, along with the data model for SAR change detection. The four change detectors discussed in this thesis have been been discussed in Chapter 3. In Chapter 4, the modified change detector is proposed, which has a simpler form. The new threshold selection method for the two-stage detector is proposed in Chapter 5. Using this threshold selection method, theoretical ROC curves are generated for the two-stage detector, and are compared with the ROC curves of the other three detectors in Chapter 6. Finally, in Chapter 7, all the detectors are applied to a SAR data set and the results are compared.

Chapter 2

BACKGROUND

2.1 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a technique which is used to generate high-resolution images of a landscape, and can operate in any weather conditions. It is a coherent processing technique which means that it measures the amplitude as well as the phase of the received signals, i.e. the received signals are complex valued. As shown in Figure 2.1, SAR involves a radar system mounted on a moving airborne or space-based platform, where radar signals are transmitted and received at distinct locations along the platform motion path.

As seen in Figure 2.1, two sets of coordinates, (x, y) and (u, r) have been used. (x, y) is the coordinate system for the complex reflectivity of the scene, and (u, r) is the coordinate system used to form the SAR raw data. At a given platform location, the radar signal is transmitted, which encounters scatterers situated in the scene and these scatterers reflect the signal incident on them. A part of the reflected signal is in the direction of the radar, and forms the received signal at the radar. Based on the time delay $t = t_d$ at which this signal is received, the distance of the scatterer from the platform location can be computed by:

$$R = \frac{ct_d}{2} \tag{2.1}$$

where c is the speed of propagation of the electromagnetic signal, i.e. the speed of light. The transmitted signal is generally a pulse repeated periodically, over a time interval known as the coherent processing interval (CPI). The pulse width τ

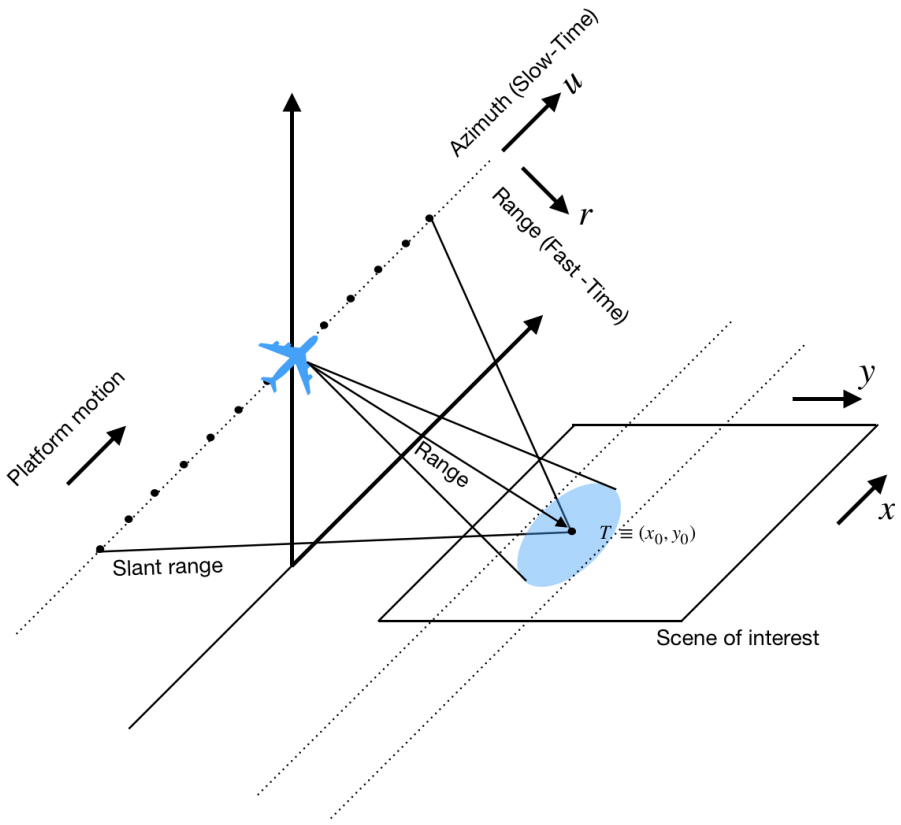


Figure 2.1: Overview of SAR.

is approximately equal to the inverse of the signal's bandwidth. Two targets can be resolved as separate entities only if their returns are separated by a time delay greater than τ. In other words, the smallest separation in the range direction at which two targets can be resolved is $\Delta r = c\tau/2$. This distance is called the range resolution. The range dimension is divided into separate range bins, each range bin is of length Δr. For each of the pulses over the CPI, a range profile is created, which is a one-dimensional vector containing the value of the radar return for every range bin. Combining these range profiles for all the pulses, the received data takes the form of a complex matrix storing the radar return values, where the dimensions correspond to the range (also known as fast time) and pulse number (known as slow time). The resolution in the range direction is improved by using pulse compression waveforms.

Linear frequency modulated (LFM) chirp signals are one of the most popular signals used, which are given by

$$f(t) = \exp\left[j(\omega t + \frac{\mu t^2}{2})\right] \qquad \frac{-\tau}{2} \leq t \leq \frac{\tau}{2} \tag{2.2}$$

where $\omega = 2\pi f_c$, f_c is the carrier frequency. The platform motion creates a *synthetic aperture* of a large size, which results in improved resolution in the direction along the flight path of the platform. This resolution is inversely proportional to the distance over which the data is collected, i.e. the length of the synthetic aperture (where smaller resolution means a better resolution).

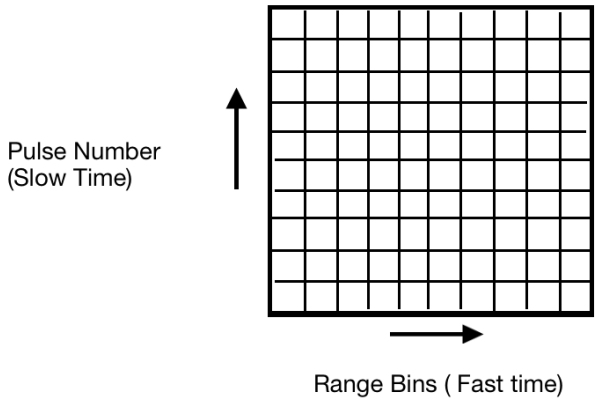


Figure 2.2: SAR Data Matrix.

In order to generate the image from the received data, a reference function is first generated. This reference function is generated by assuming the presence of a single point target at location (x_0 ,y_0). Based on the transmitted waveform and the platform motion characteristics, an estimate of the radar data matrix due to the point target at (x_0 ,y_0) can be generated, and this estimate is the reference function. It is a function of u and r for a given point target at x_0 and y_0, and can be denoted as $h(u, r; x_0, y_0)$. This reference function thus can be thought of as the shifted *impulse response* for

the SAR system. In order to generate the complex pixel value at $(x_0\ ,y_0)$, a 2D matched filter response is computed by performing the 2D convolution between the received data matrix $x(u,r)$ and $h^*(-u,-r;x_0,y_0)$ [8]. In order to generate an image, a matched filter bank is created with the reference functions for every (x_0,y_0), and the received data matrix $x(u,r)$ passed through this filter bank gives the final image. The range-doppler, range-migration, chirp-scaling, and backprojection algorithms are few of the of algorithms which are used to generate SAR images, where all of these algorithms try to implement this 2D matched filter in different ways.

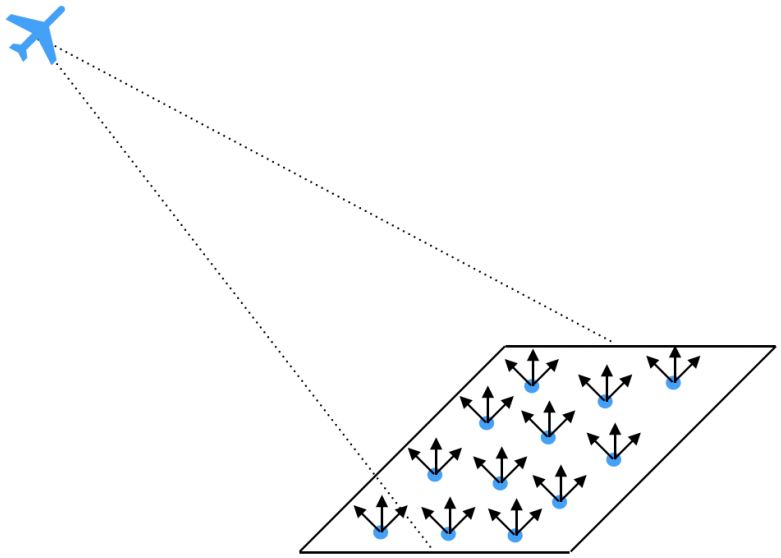


Figure 2.3: Large Number of Scatterers Illuminated by the Antenna Beam.

In general, the SAR return for a single resolution cell consists of a coherent sum of a number of independent scatterers [9]. If this number is large, the central limit theorem can be applied, and the radar return for each resolution cell can be modeled as a zero mean circularly symmetric complex Gaussian random variable. The probability density function (pdf) for this random variable is given by [10]:

$$p(x) = \frac{1}{\pi\sigma^2}\exp\left[-\frac{|x|^2}{\sigma^2}\right]. \tag{2.3}$$

2.2 Detection Theory Concepts

In a decision theory problem, we have a set of possible hypotheses, and under each of these hypotheses, the random data vector $\mathbf{x}$ ($\in \mathbb{R}^n$ or $\mathbb{C}^n$) which is observed follows a different probability distribution. The simplest form of a decision problem is the *binary hypothesis testing* problem, where we have only two hypotheses, the null hypothesis (H_0) and the alternative hypothesis (H_1). If the probability density function (pdf) of $\mathbf{x}$ depends on parameter θ, then the hypothesis testing problem can be written as:

$$\begin{aligned} H_0 &: \mathbf{x} \sim p(\mathbf{x}; H_0) = p(\mathbf{x}; \theta_0) \\ H_1 &: \mathbf{x} \sim p(\mathbf{x}; H_1) = p(\mathbf{x}; \theta_1) \end{aligned} \tag{2.4}$$

where θ_i denotes the value of θ under H_i. The goal is to come up with a decision rule which can accurately map any realization of the random vector to the probability distribution which most likely generated it. In order to form a decision rule, a function of the random vector $\mathbf{x}$ is created. The range (known as the decision space) of this function is partitioned into the critical region (Z_1) and the non-critical region (Z_0). Z_i is the region where the decision rule decides that H_i is true. Such a function is known as a test statistic, denoted by $T(\mathbf{x})$, and this test statistic along with the decision rule is referred to as a detector. In order to characterize the performance of any detector, two quantities are generally used, the probability of false alarm (P_{FA}) and the probability of detection (P_D). P_{FA} is the probability that the detector will reject H_0, when the underlying true hypothesis is H_0, and P_D is defined as the probability

that detector rejects H_0 when actually H_1 is true. These probabilities are given by:

$$P_{FA} = \Pr\{T(\mathbf{x}) \in Z_1; H_0\} = \int_{Z_1} p(T; H_0) dT \tag{2.5}$$

$$P_D = \Pr\{T(\mathbf{x}) \in Z_1; H_1\} = \int_{Z_1} p(T; H_1) dT \tag{2.6}$$

where $p(T; H_0)$ and $p(T; H_1)$ denote the pdf of $T(\mathbf{x})$ under H_0 and H_1 respectively. We can see that an ideal detector should have $P_D = 1$ and $P_{FA} = 0$, but it turns out that we cannot increase P_D and decrease P_{FA} simultaneously. However, we can find an optimal detector which maximizes P_D for a given P_{FA} constraint. The optimal test statistic (known as the *likelihood ratio*) is given by the *Neyman-Pearson rule* [11], and it decides that H_1 is true if:

$$T(\mathbf{x}) = \frac{p(\mathbf{x}; H_1)}{p(\mathbf{x}; H_0)} > \eta \tag{2.7}$$

where the threshold η is obtained by solving:

$$P_{FA} = \int_{\mathbf{x}:T(\mathbf{x})>\eta} p(\mathbf{x}; H_0) d\mathbf{x} = \int_{\eta}^{\infty} p(T; H_0) dT. \tag{2.8}$$

It is evident that in order to implement the likelihood ratio, complete knowledge of the pdfs of $\mathbf{x}$ under both hypotheses (i.e knowledge of θ_0 and θ_1) is necessary. Since this is rarely the case, this optimal detector generally cannot be implemented. As a result, different detectors are used, which although do not give the optimal performance, are physically realizable. In the next section, the distribution of the data for the problem of change detection in SAR images is introduced, and a number of detectors which have been proposed for this problem are summarized in the next chapter.

2.3 Data Model for Change Detection in SAR Images

In a SAR change detection problem, we are given two data sets $\{f_k\}$ and $\{g_k\}$ of the same size, where $f_k,\ g_k \in \mathbb{C}$. At a time, change is detected between a single pair of pixels (say $k = i$) f_i and g_i, which are the radar returns corresponding to the same location, but measured at two different times. Since it is clear that change is detected for a fixed pair of pixels, for simplicity we can now drop the subscript i. We use the statistical properties of the complex 2×1 random vector $[f,\ g]^T$, in order to characterize change. For the purpose of estimating these statistical properties, we assume that all the radar returns within a spatial window of size N around f follow a distribution identical to f, and are independent of f. A similar argument is made for g. Each data window of size N is modeled as a collection of independent samples from a zero mean circularly symmetric complex Gaussian distribution. However, since f and g correspond to the same location, they are not uncorrelated/independent in general. Thus, each of the 2×1 vectors $[f_k,\ g_k]^T$ for k=1, . . . , N (which form a spatial window around $[f,\ g]^T$) follow a zero mean bivariate complex Gaussian distribution, with the covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_f^2 & \rho\sigma_f\sigma_g \\ \rho^*\sigma_f\sigma_g & \sigma_g^2 \end{bmatrix} \tag{2.9}$$

where

$$\sigma_f^2 = \mathrm{Var}(f) = \mathbb{E}(|f|^2), \quad \sigma_g^2 = \mathrm{Var}(g) = \mathbb{E}(|g|^2) \tag{2.10}$$

and ρ is the complex correlation coefficient between f and g given by

$$\rho = \frac{\mathbb{E}(fg^*)}{\sqrt{\mathbb{E}(|f|^2)\mathbb{E}(|g|^2)}} \tag{2.11}$$

The magnitude of the correlation coefficient ρ, i.e., $|\rho|$, is known as the coherence between f and g. Thus, the parameters which characterize the joint distribution of

f and g are σ_f^2, σ_g^2, and ρ. Based on the values of these parameters under the change and no-change scenarios, hypothesis tests are formulated where the null hypothesis represents no change and the alternative hypothesis represents a change between the pixels. These different hypothesis test formulations lead to different change detectors. Some of these detectors are summarized in the next chapter.

Chapter 3

CHANGE DETECTION STATISTICS

Change detection statistics can be classified into two types - non-coherent and coherent change detection. Non-coherent change detection (NCCD) involves exploiting only the magnitude information of the complex-valued radar returns, whereas coherent change detection (CCD) exploits the phase information as well. The problem of detecting change is formulated as a hypothesis testing problem, in each of the cases. NCCD uses the sample variance ratio to test the equality of the two variances σ_f^2 and σ_g^2, and declares no change if they are equal. The classical CCD uses the sample coherence statistic in order to determine the coherence $|\rho|$ between the two pixels, and declares change if they are uncorrelated. An alternative coherence estimator, given by Berger in [5] can also be used for estimating coherence, however it provides accurate estimates only if the underlying variances are equal. A two-stage detector introduced by [2] combines the two techniques, where the first stage is the sample variance ratio and the second stage is Berger's coherence estimator. The first stage tests if the two variances are equal, the pixel pairs which show equal variances (and thus are classified as no change) will pass to the second stage where the Berger's estimator will test for the subtle changes.

3.1 Sample Variance Ratio

Non-coherent change detection compares the two variances by forming the ratio $R = \sigma_f^2/\sigma_g^2 \in (0, \infty)$. For a given pair of pixels $[f, g]^T$, if the value of R is close to one, i.e. when the variances of $\sigma_f^2 = \sigma_g^2$ are equal, then this pair is labeled as no change, whereas if R deviates further from unity, the detector declares change.

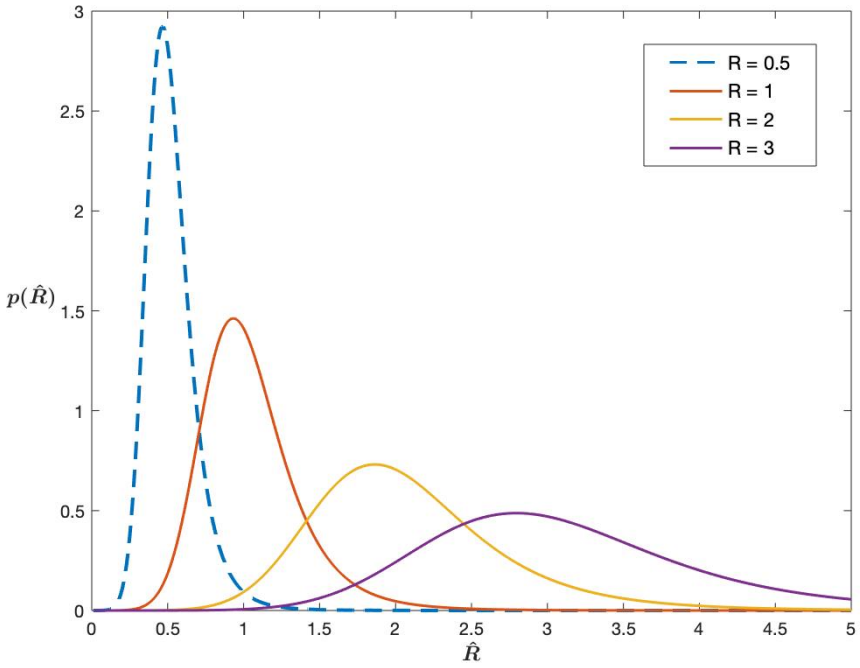


Figure 3.1: Probability Density Function $p(\hat{R}; R, |\rho|, N)$ for $N = 5$, $|\rho| = 0.9$, and different values of R.

The hypothesis test problem for NCCD can be written as follows:

$$H_0 : R \approx 1$$
$$H_1 : R \neq 1 \tag{3.1}$$

The test statistic used for this hypothesis test is the sample variance ratio ($\hat{R}$), given by:

$$\hat{R} = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_g^2} = \frac{\sum_{k=1}^{N} |f_k|^2}{\sum_{k=1}^{N} |g_k|^2} \quad \text{where} \quad \hat{\sigma}_f^2 = \frac{1}{N}\sum_{k=1}^{N} |f_k|^2, \quad \hat{\sigma}_g^2 = \frac{1}{N}\sum_{k=1}^{N} |g_k|^2 \tag{3.2}$$

It is evident from the above expression that $\hat{R}$ uses only the magnitude information of f_k and g_k, and not the phase information. The pdf of $\hat{R}$ is given by [12]:

$$p(\hat{R}; R, |\rho|, N) = \frac{\Gamma(2N)}{\Gamma(N)^2} \frac{(1-|\rho|^2)^N (\hat{R}+R) R^N \hat{R}^{N-1}}{[(\hat{R}+R)^2 - 4\hat{R}R|\rho|^2]^{N+\frac{1}{2}}} \qquad \text{for } 0 < \hat{R} < \infty \tag{3.3}$$

where $R = \sigma_f^2/\sigma_g^2$ is the ratio of the true variances, ρ is the complex correlation coefficient between the two complex random radar returns and $\Gamma(.)$ denotes the gamma function. We can see that the pdf of $\hat{R}$ depends on the true variances σ_f^2 and σ_g^2 only through the true variance ratio R, coherence $|\rho|$, and the window size N. Figure 3.1 shows the pdf of $\hat{R}$ plotted for different values of R and fixed N and $|\rho|$. We can

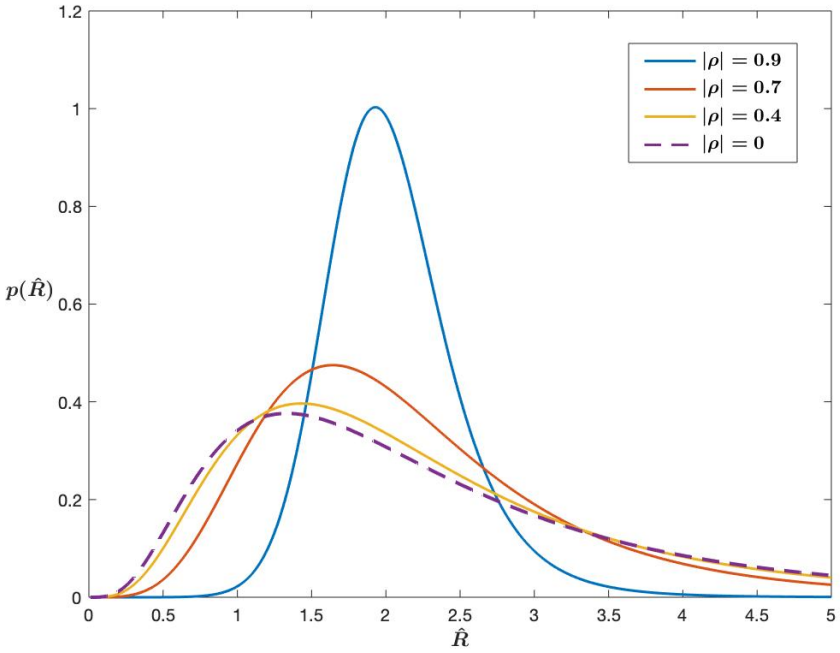


Figure 3.2: Probability Density Function $p(\hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 2$, and Different Values of $|\rho|$.

see that for this particular choice of N and $|\rho|$, the pdf is approximately centered at the true R. However, as we decrease the value of the coherence $|\rho|$ for fixed R and N, the bias of $\hat{R}$ increases, which can be seen from Figure 3.2. We can also see that as $|\rho|$ reduces, the variance of $\hat{R}$ increases. $\hat{R}$ decides that a pixel pair has not undergone change if $\eta_1 \leq \hat{R} \leq \eta_2$ (where $\eta_1 < 1 < \eta_2$), else, the detector declares change. In other words, the region Z_1 where this detector declares change is $Z_1 = \{x \in \mathbb{R}_+ | x \leq \eta_1 \text{ or } x \geq \eta_2\}$ for some η_1 and η_2 such that $\eta_1 < 1 < \eta_2$, as shown in Figure 3.3.

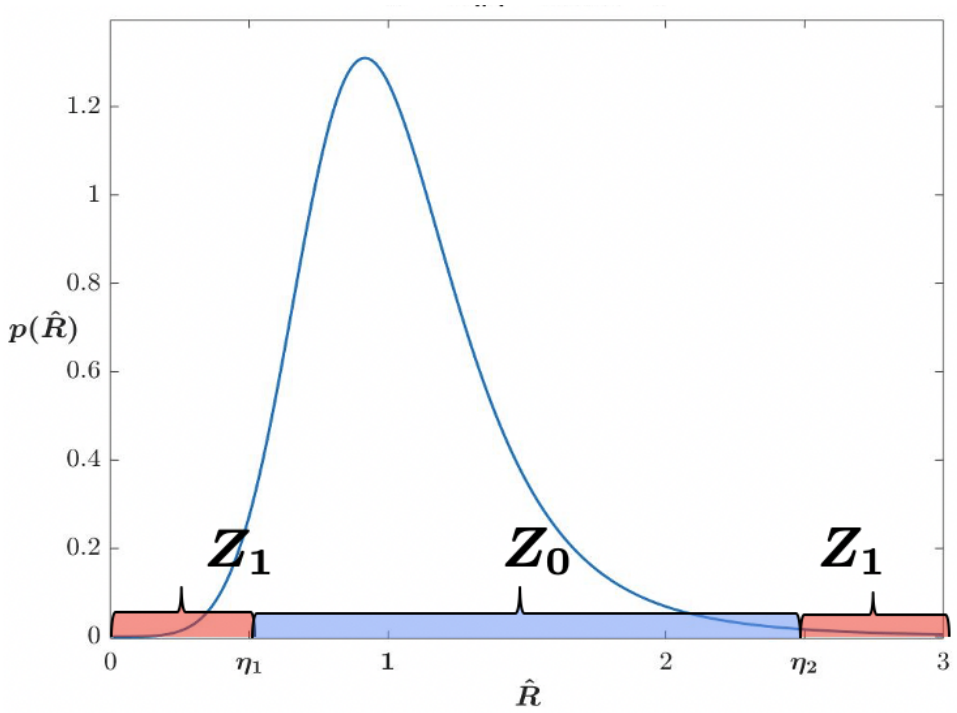


Figure 3.3: Probability Density Function $p(\hat{R}; R, |\rho|, N)$, Showing the Region Z_1 where $\hat{R}$ Declares Change.

3.2 Sample Coherence/Classical Coherence Estimator

Coherent change detection uses phase information (in addition to the magnitude information) of the complex-valued radar returns in order to detect subtle changes. The parameter used to define change is the coherence $|\rho|$, i.e. pixel pairs which are uncorrelated are classified as changed pixels and those showing some degree of correlation are classified as change. Since coherence is sensitive to a number of factors which govern the SAR data acquisition, practically speaking, even pixels undergoing no change do not exhibit a value of coherence equal to one [2]. The hypothesis testing problem can be written as:

$$\begin{aligned} H_0 &: |\rho| \approx 1 \\ H_1 &: |\rho| \approx 0 \end{aligned} \tag{3.4}$$

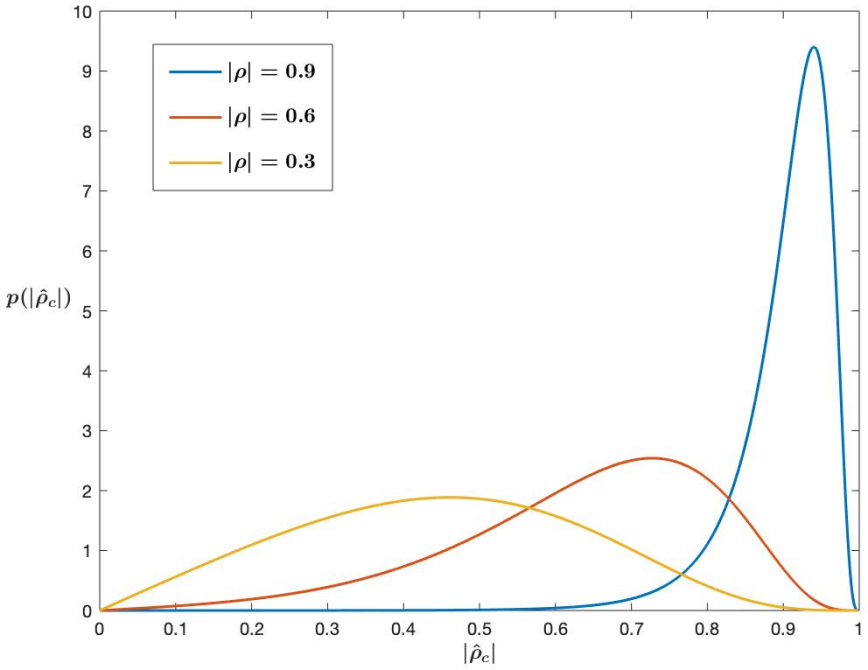


Figure 3.4: Probability Density Function $p(|\hat{\rho}_c|; |\rho|, N)$ for $N = 6$ and Different Values of $|\rho|$.

The sample coherence statistic is the magnitude of $\hat{\rho}_c$, which is an estimator of the cross-correlation coefficient ρ. The sample coherence is given by:

$$|\hat{\rho}_c| = \frac{\left|\sum_{k=1}^{N} f_k g_k^*\right|}{\sqrt{\sum_{k=1}^{N} |f_k|^2 \sum_{k=1}^{N} |g_k|^2}} \tag{3.5}$$

The pdf of $|\hat{\rho}_c|$ is given by [13]:

$$p(|\hat{\rho}_c|; |\rho|, N) = 2(N-1)\left(1-|\rho|^2\right)^N |\hat{\rho}_c| \left(1-|\hat{\rho}_c|^2\right)^{N-2} {}_2F_1(N, N; 1; |\hat{\rho}_c|^2|\rho|^2) \quad \text{for } 0 \leq |\hat{\rho}_c| < 1 \tag{3.6}$$

where ${}_2F_1(.,.;.;.)$ is the Gauss hypergeometric function [14]. It is evident that this pdf depends only on N and $|\rho|$. It is worth noting that $p(|\hat{\rho}_c|)$ is independent of the true variances σ_f^2 and σ_g^2. The pdf is plotted in Figure 3.4 for a fixed N and

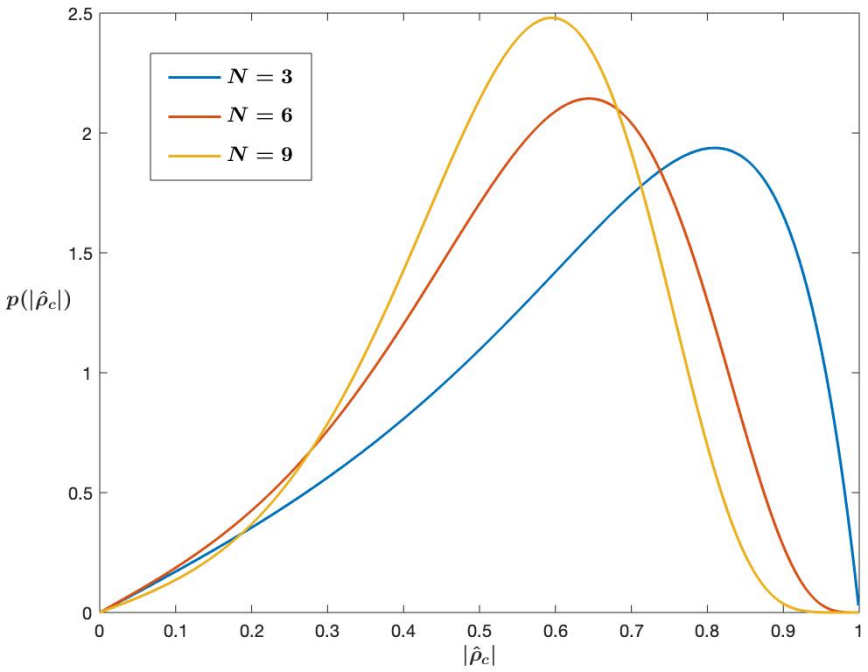


Figure 3.5: Probability Density Function $p(|\hat{\rho}_c|; |\rho|, N)$ for $|\rho| = 0.5$ and Different Values of N.

different values of $|\rho|$. We can see that the bias and variance increase as the true $|\rho|$ decreases for a constant N. In Figure 3.5, the pdf of $|\hat{\rho}_c|$ is plotted for different values of N, keeping $|\rho|$ constant. In this case, as N increases, the bias and variance of $|\hat{\rho}_c|$ decrease. The region Z_1 for this statistic is of the form $Z_1 = \{|\hat{\rho}_c| \in \mathbb{R}_+ : |\hat{\rho}_c| \leq \eta\}$ for some $0 < \eta < 1$.

3.3 Berger's Alternative Coherence Estimator

An alternative coherence estimator was given by Berger in [5], for the case when the underlying variances σ_f^2 and σ_g^2 are equal. This alternative estimator is denoted by $|\hat{\rho}_a|$, and is given by

$$|\hat{\rho}_a| = \frac{\left|\sum_{k=1}^{N} f_k g_k^*\right|}{\frac{1}{2}\left(\sum_{k=1}^{N} |f_k|^2 + \sum_{k=1}^{N} |g_k|^2\right)}. \tag{3.7}$$

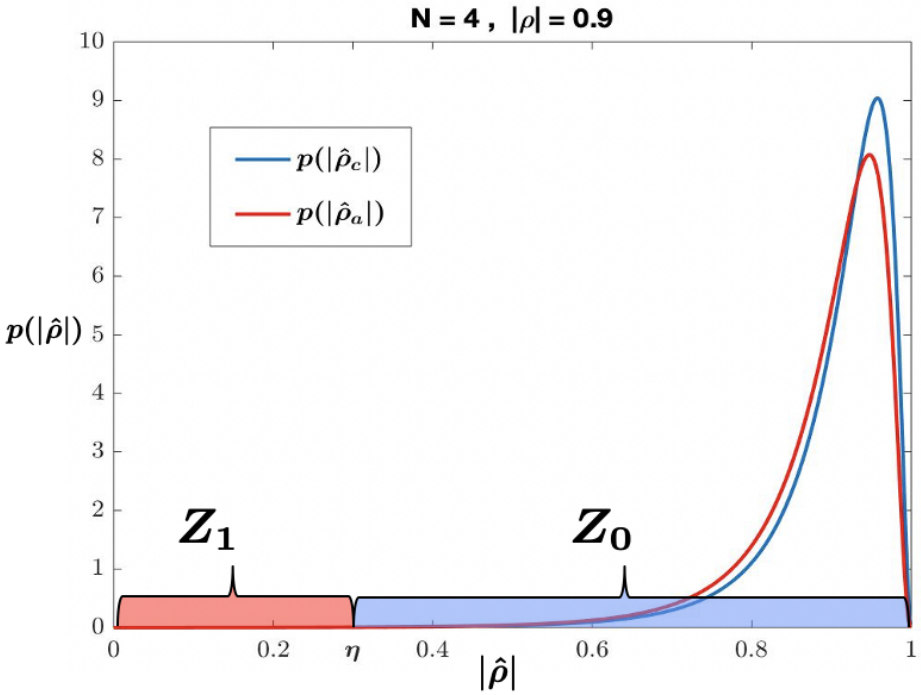


Figure 3.6: Probability Density Functions $p(|\hat{\rho}_c|; |\rho|, N)$ and $p(|\hat{\rho}_a|; |\rho|, N)$, Showing the Region Z_1 where the Two Detectors Declare Change.

Comparing (3.7) and (3.5), we can see that $|\hat{\rho}_a|$ and $|\hat{\rho}_c|$ have the same numerator, but the denominators are different. The sample coherence statistic has the geometric mean of two quantities (sample variances multiplied by N) as the denominator, while Berger's estimate has the arithmetic mean of the same two quantities as the denominator. Only if the two sample variances are approximately equal, the arithmetic and the geometric means are equivalent and thus in this case $|\hat{\rho}_a|$ becomes an accurate estimator of the coherence. The condition that the underlying variances are equal is equivalent to $R = 1$. The pdf of $|\hat{\rho}_a|$ under the assumption $R = 1$, is given by [5]:

$$\begin{aligned} p(|\hat{\rho}_a|; |\rho|, N) =& (2N-1)\left(1-|\rho|^2\right)^N |\hat{\rho}_a| \left(1-|\hat{\rho}_a|^2\right)^{N-\frac{3}{2}} \\ & {}_2F_1(N, N+\frac{1}{2}; 1; |\hat{\rho}_a|^2|\rho|^2) \qquad \text{for } 0 \le |\hat{\rho}_a| < 1 \end{aligned} \tag{3.8}$$

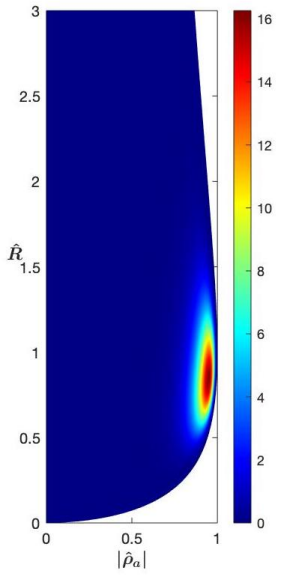
A detailed comparison of estimation characteristics between $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ has been shown in [2]. Z_1 for $|\hat{\rho}_a|$ takes a similar form to that of $|\hat{\rho}_c|$.

3.4 Two-Stage Change Detector

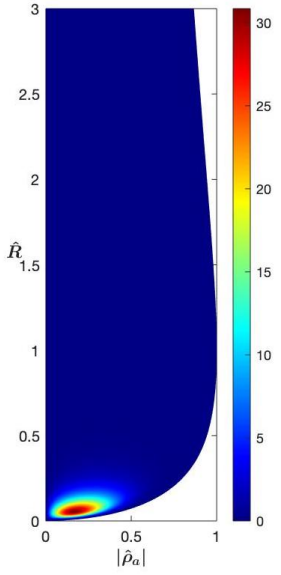
Many studies including [2] have shown that the sample variance ratio change detector generally performs better at detecting significant changes like the appearance of a large object like a car , while coherent change detection shows a better performance at detecting subtle changes in the scene. The two-stage change detector, was introduced by [2], and it combines the coherent change detection and the variance ratio information in order to detect changes with higher accuracy at any scale. However, instead of the sample coherence detector, it uses the alternative coherence estimator given by Berger. With the help of mean squared error analysis (MSE), it has been shown in [2] that when the underlying true variances σ_f^2 and σ_g^2 are approximately equal, $|\hat{\rho}_a|$ is a better estimator of $|\rho|$ as compared to $|\hat{\rho}_c|$. Thus, the first stage of the two-stage detector is $\hat{R}$, which tests for equality of the two variances. If the variances are not equal, it decides that a change has occurred. However, if $\sigma_f^2 \approx \sigma_g^2$, then those pixel pairs are passed on to the second stage, which is $|\hat{\rho}_a|$. Since the pixel pairs passed on to the second stage have almost equal variance, $|\hat{\rho}_a|$ gives an accurate estimation of $|\rho|$, and thus changes can be detected on a smaller scale as well. Based on the hypotheses testing formulations for R and $|\rho|$, the detection problem for the two-stage change detector can be written hypothesis testing problem as follows :

$$\begin{aligned} H_0 &: R \approx 1 \quad \text{and} \quad |\rho| \approx 1 \\ H_1 &: R \neq 1 \quad \text{or} \quad |\rho| \approx 0 \end{aligned} \tag{3.9}$$

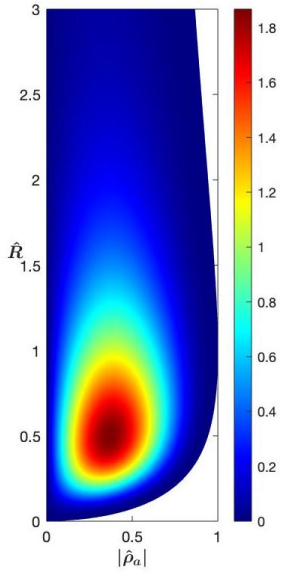
The two-stage change detector works as a two dimensional detector, i.e. the decision space is two dimensional, where the dimensions correspond to the $\hat{R}$ and $|\hat{\rho}_a|$ statistics. In order to evaluate the performance of this 2D detector, analysis of the joint pdf of $\hat{R}$ and $|\hat{\rho}_a|$ is required. The joint density function for the two test statistics is derived



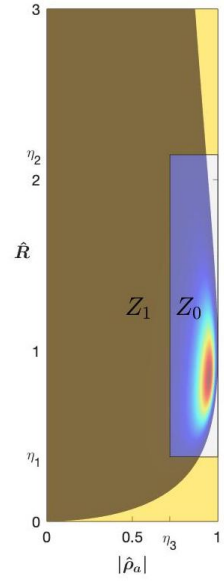
(a) Joint pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 0.9$, and $|\rho| = 0.9$.



(b) Joint pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 0.1$, and $|\rho| = 0$.



(c) Joint pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 0.8$, and $|\rho| = 0.1$.



(d) Typical form of Z_0 and Z_1 for Two-stage Detector.

Figure 3.7: Joint Pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for Two-stage Detector.

in [2], and is given by:

$$\begin{aligned} p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N) =& \frac{(1-|\rho|^2)^N \Gamma(2N)}{\Gamma(N)\Gamma(N-1)} \frac{|\hat{\rho}_a|}{2(\hat{R}+1)^2} \\ & \cdot \left[\frac{\hat{R}}{(\hat{R}+1)^2} - \frac{|\hat{\rho}|^2}{4} \right]^{N-2} \\ & \cdot \left[|\hat{\rho}_a||\rho| + \frac{\hat{R}+R}{(\hat{R}+1)\sqrt{R}} \right]^{-2N} \\ & \cdot {}_2F_1 \left(\frac{1}{2}, 2N; 1; \frac{2|\hat{\rho}_a||\rho|}{|\hat{\rho}_a||\rho| + \frac{\hat{R}+R}{(\hat{R}+1)\sqrt{R}}} \right) \end{aligned} \tag{3.10}$$

for $0 \leq \hat{R} < \infty$, $|\hat{\rho}_a| \leq [4\hat{R}/(\hat{R}+1)^2]^{1/2}$ and zero elsewhere, where ${}_2F_1(.,.;.;.)$ is the Gauss hypergeometric function [14]. The pdf depends on the true variance ratio $R = \sigma_f^2/\sigma_g^2$, the true coherence $|\rho|$, and the spatial window size N. The joint pdf is plotted for different values of R, $|\rho|$ and N in Figures 3.7a, 3.7b, and 3.7c.

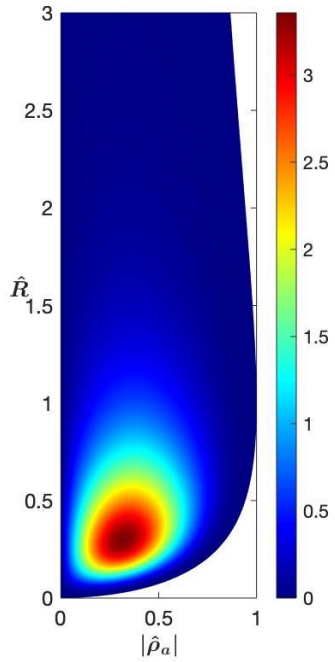
Considering the hypothesis test in (3.9), and the forms of region Z_1 for the Berger's estimate (which is similar to that for $|\hat{\rho}_c|$) and the sample variance ratio statistics, we can see that the partition of the decision space for the two-stage detector has the form shown in Figure 3.7d, given by $Z_0 = \{(|\hat{\rho}_a|, \hat{R}) \in (0,1) \times \mathbb{R}_+ : |\hat{\rho}_a| > \eta_3 \text{ and } \eta_1 < \hat{R} < \eta_2\}$.

Chapter 4

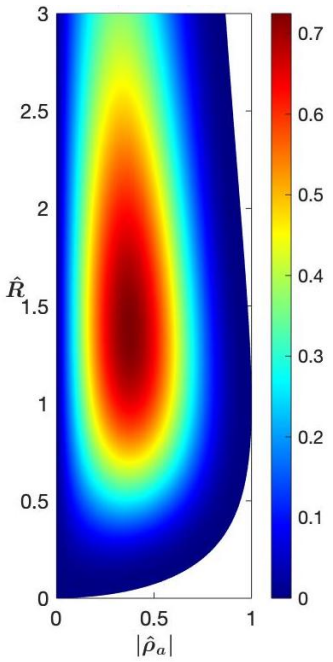
MODIFIED TWO-STAGE CHANGE DETECTOR

The probability density function of the two-stage detector is defined for all values of $\hat{R}$ which are non-negative. Considering the form of the partitions Z_0 and Z_1 in Figure 3.7d, we can see that η_2 can take any value greater than one. Also, an important point worth noting is that the value of R depends on the assignment of the two data sets as f and g (if we interchange f and g, R will become $1/R$). Since our goal is to detect change irrespective of whether the change in variance is an increase or a decrease, the detector should give exactly the same performance no matter which image we consider as f or g, i.e. for both the cases when $R = R_1$ and when $R = 1/R_1$, where R_1 is any arbitrary value for the variance ratio. However, $\hat{R}$ is not symmetric in this sense, i.e. it is not invariant to the assignment of images as f and g. In order to see this, we should consider Figure 3.1, which shows the pdfs of $\hat{R}$ for $R = 1/2$ and $R = 2$. We can see that the two functions are completely different, and thus might not give the same detection performance in both the cases. Furthermore, Figures 4.1a and 4.1b show the joint pdfs for the two dimensional detector for $R = 1/2$ and $R = 2$ respectively, for fixed $N = 5$ and $|\rho| = 0$. It is evident that these two pdfs are not the same, and thus might give different performances, in spite of the fact that both these cases depict the same scenario in terms of change detection. In order to avoid this problem, we must use the symmetric ratio detector $\hat{r}$ given by [7] as the first stage, in place of $\hat{R}$. This detector is given by:

$$\hat{r} = \begin{cases} \hat{R} & \text{when } \hat{R} \leq 1 \\ \hat{R}^{-1} & \text{when } \hat{R} > 1. \end{cases} \tag{4.1}$$



(a) Joint pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 1/2$, and $|\rho| = 0$.



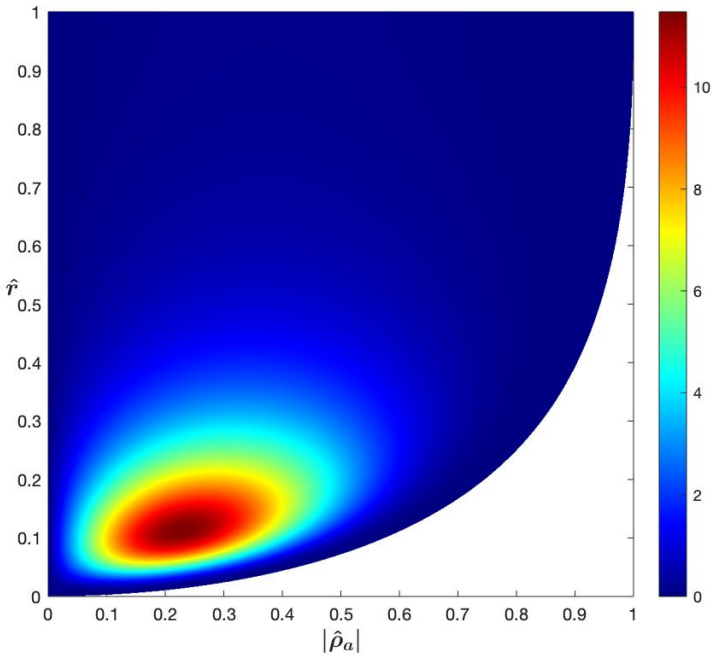
(b) Joint Pdf $p(|\hat{\rho}_a|, \hat{R}; R, |\rho|, N)$ for $N = 5$, $R = 2$, and $|\rho| = 0$.

Figure 4.1: Joint Pdf of Two-Stage Detector for Reciprocal Values of R.

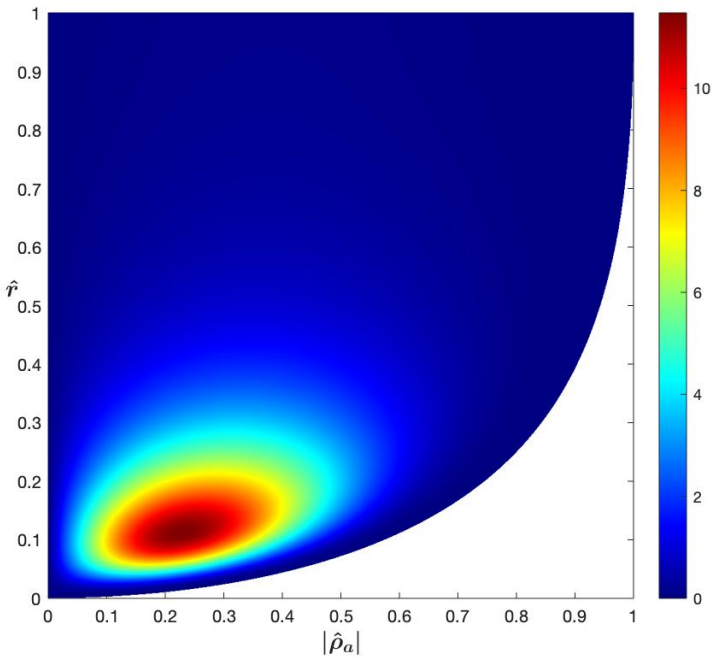
The pdf for $\hat{r}$ has been derived in [7] for the case when $|\rho| = 0$. The density function for a non-zero coherence $|\rho|$ has been derived in Appenidx A, and is given by:

$$p(\hat{r}; R, |\rho|, N) = \frac{\Gamma(2N)}{\Gamma(N)^2} \hat{r}^{N-1}(1-|\rho|^2)^N \cdot \left(\frac{R^N(\hat{r}+R)}{[(\hat{r}+R)^2 - 4\hat{r}R|\rho|^2]^{N+\frac{1}{2}}} + \frac{R^{-N}(\hat{r}+R^{-1})}{[(\hat{r}+R^{-1})^2 - 4\hat{r}R^{-1}|\rho|^2]^{N+\frac{1}{2}}} \right) \quad (4.2)$$

for $0 \leq \hat{r} \leq 1$ and zero elsewhere. Looking at the expression for the pdf of $\hat{r}$, we can easily see that i.e., $\hat{r}$ has the same pdf for $R = R_1$ and $R = R_1^{-1}$ for fixed $|\rho|$ and N, and this is exactly the desirable symmetry property. Furthermore, since $\hat{r}$ takes values only between zero and one, using $\hat{r}$ instead of $\hat{R}$ for the first stage reduces the size of the decision space for the two-stage detector from $(0 \leq |\hat{\rho}_a| < 1, 0 \leq \hat{R} < \infty)$



(a) $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$ for $N = 5$, $R = 1/5$, and $|\rho| = 0$.



(b) $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$ for $N = 5$, $R = 5$, and $|\rho| = 0$.

Figure 4.2: Joint Pdf $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$ for Modified Two-Stage Detector Showing Symmetry.

to $(0 \leq |\hat{\rho}_a| < 1, 0 \leq \hat{r} \leq 1)$. The joint pdf of $\hat{r}$ and $|\hat{\rho}_a|$ is derived in Appendix B, and is given by:

$$p_{|\hat{\rho}_a|,\hat{r}}(x, y; R, |\rho|, N) = p_{|\hat{\rho}_a|,\hat{R}}(x, y; R, |\rho|, N) + p_{|\hat{\rho}_a|,\hat{R}}(x, y; \frac{1}{R}, |\rho|, N) \tag{4.3}$$

for $0 \leq y \leq 1$, $x \leq [4y/(y+1)^2]^{1/2}$ and zero elsewhere. Note that the subscript indicates the random variable to which the distribution belongs, and this notation is used to interpret (4.3) easily. $p_{|\hat{\rho}_a|,\hat{R}}(x, y; R, |\rho|, N)$ is the joint pdf of $|\hat{\rho}_a|$ and $\hat{R}$, as given in (3.10). The joint pdf for $|\hat{\rho}_a|$ and $\hat{r}$ is plotted in Figures 4.2a and 4.2b for $R = 5$ and $R = 1/5$ respectively, for the same values of $|\rho|$ and N, and we can see that the distributions are exactly the same. Thus, using $\hat{r}$ as the first stage makes the detector invariant to image assignment as f and g. Figures 4.3 and 4.4 show the joint pdf for $|\rho| = 0.9$ and $|\rho| = 0$ respectively, for fixed $N = 5$ and $R = 1$.

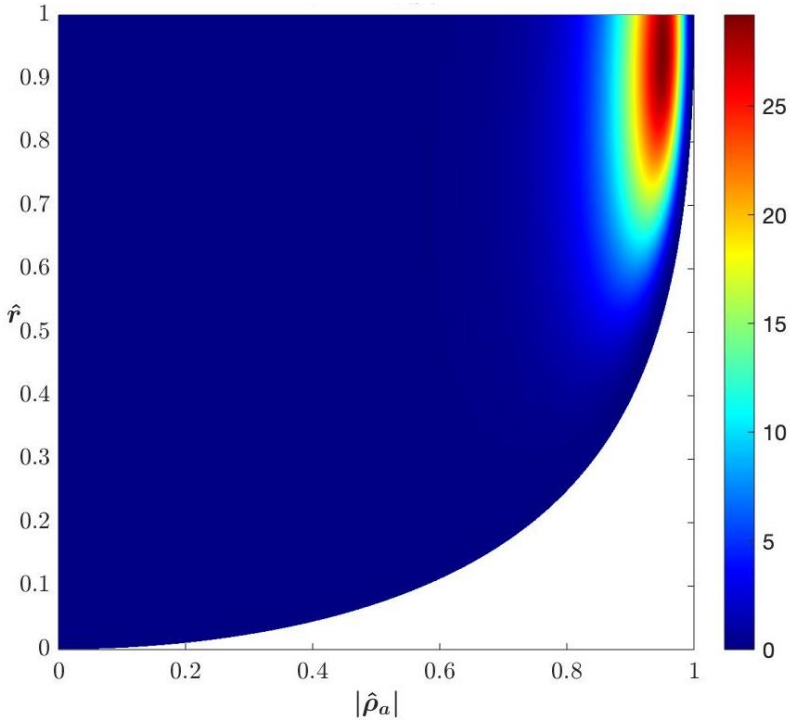


Figure 4.3: Joint Pdf $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$ for $N = 5$, $R = 1$, and $|\rho| = 0.9$.

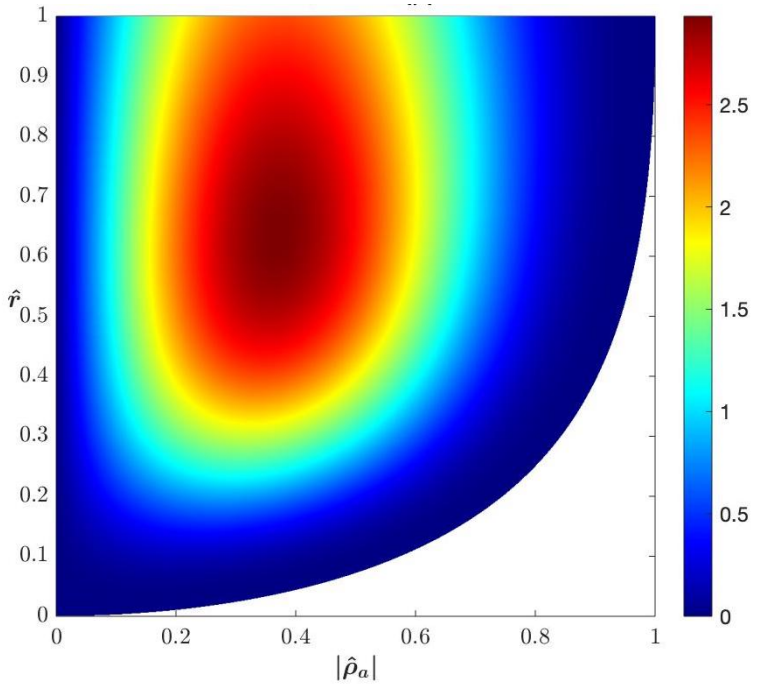


Figure 4.4: Joint Pdf $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$ for $N = 5$, $R = 1$, and $|\rho| = 0$.

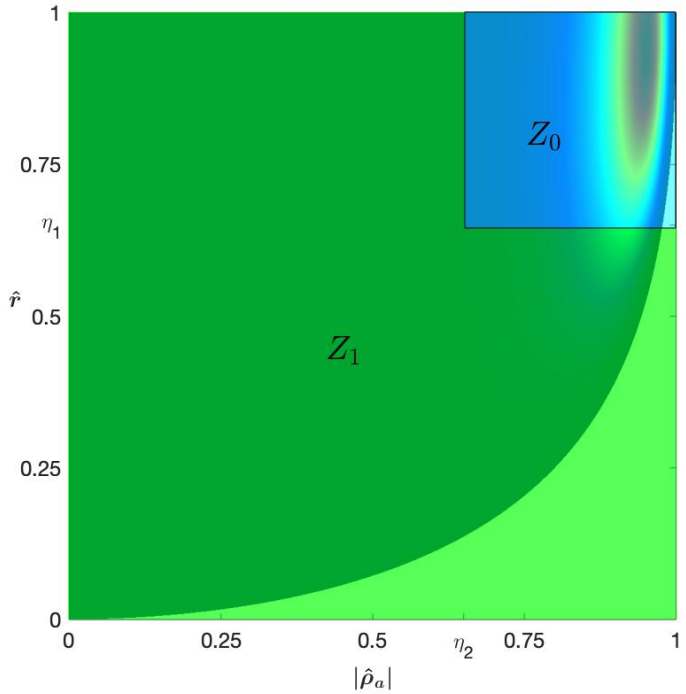


Figure 4.5: Typical Form of Partitions Z_0 and Z_1 of Decision Space for Modified Two-Stage Detector.

Furthermore, as shown in Figure 4.5, the partitions Z_0 and Z_1 now assume a much simpler form given by $Z_0 = \{(|\hat{\rho}_a|, \hat{r}) \in (0,1) \times (0,1] : |\hat{\rho}_a| > \eta_2 \text{ and } \hat{r} > \eta_1\}$, which makes the analysis of this method significantly simpler. We can see this by comparing Figures 4.5 and 3.7d. Instead of computing three thresholds, we now are required to compute just two thresholds η_1 and η_2.

Thus, it has been shown that replacing $\hat{R}$ by the symmetric statistic $\hat{r}$ as the first stage results in a form of the two-stage detector which is significantly simpler.

Chapter 5

THRESHOLD SELECTION FOR MODIFIED TWO-STAGE CHANGE DETECTOR

In the previous chapter, the typical forms of Z_0 and Z_1 for the modified two-stage detector were introduced, and it is evident that these regions are uniquely determined by the value of thresholds η_1 and η_2. A heuristic approach for computing the three thresholds η_1, η_2, and η_3 for the original two-stage detector has been demonstrated in [2]. In this chapter, a new approach for determining these thresholds has been proposed, which uses the joint probability density function of the modified two-stage detector, i.e. $p(|\hat{\rho}_a|, \hat{r}; R, |\rho|, N)$. Before introducing the proposed method of threshold selection, the computation of thresholds for single-dimensional test statistics (like $\hat{r}$ or $|\hat{\rho}_c|$) are discussed.

5.1 Computing Thresholds for $\hat{r}$ and $|\hat{\rho}_c|$

The performance of different detectors can be compared by comparing the value of probability of detection (P_D) which they can achieve, for a given probability of false alarm (P_{FA}). We can determine Z_1 by using the given P_{FA} constraint as shown in (2.5), and then find the P_D by integrating the pdf of the test statistic under H_1 over Z_1. In Section 3.2, we have seen that for $|\hat{\rho}_c|$, Z_1 is the interval $0 \leq |\hat{\rho}_c| \leq \eta$. Therefore, the threshold η can be found by solving

$$P_{FA} = \int_{Z_1} p(|\hat{\rho}_c|; H_0) d|\hat{\rho}_c| = \int_0^{\eta} p(|\hat{\rho}_c|; H_0) d|\hat{\rho}_c| \tag{5.1}$$

where $p(|\hat{\rho}_c|; N, |\rho|)$ is given in (3.6). η is computed by finding the zero of the function

$$h_1(\eta) = \int_0^{\eta} p(|\hat{\rho}_c|; H_0) d|\hat{\rho}_c| - P_{FA}. \tag{5.2}$$

The region Z_1 for $\hat{r}$ is identical to that of $|\hat{\rho}_c|$, i.e. it is of the form $0 \leq \hat{r} \leq \eta$. Thus, similar to (5.1), η is computed by solving

$$P_{FA} = \int_0^{\eta} p(\hat{r}; H_0) d\hat{r} = F_{\hat{r}}(\eta; R, |\rho|, N) \tag{5.3}$$

where $F_{\hat{r}}(\eta; R, |\rho|, N)$ is the cumulative distribution function (cdf) of $\hat{r}$, and the expression for this cdf has been derived in Appendix A, which is given by:

$$F_{\hat{r}}(\eta; R, |\rho|, N) = F_{\hat{R}}(\eta; R, |\rho|, N) + F_{\hat{R}}(\eta; R^{-1}, |\rho|, N) \tag{5.4}$$

for $0 < \eta \leq 1$, where the expression for $F_{\hat{R}}(\eta; R, |\rho|, N)$ is derived in [15], given by

$$\begin{aligned} F_{\hat{R}}(\eta; R, |\rho|, N) =& \Pr(\hat{R} \leq \eta; R, |\rho|, N) \\ =& 0.5\{1 - \text{sign}[\lambda_1(\eta)]\} + \text{sign}[\lambda_1(\eta)].F_G(l(\eta), N). \end{aligned} \tag{5.5}$$

In the above expression, G is a complex central F random variable with both degrees of freedom equal. $F_G(.\,, N)$ denotes the cdf of this random variable with both degrees of freedom equal to N, and is given by [16]:

$$F_G(x, N) = \frac{x^N}{(1+x)^{2N-1}} \sum_{k=0}^{N-1} \binom{2N-1}{k+N} x^k \tag{5.6}$$

The remaining terms are given as follows:

$$\begin{aligned} \lambda_1(\eta) &= (R - \eta) - \sqrt{(R+\eta)^2 - 4\eta|\rho|^2 R} \\ \lambda_2(\eta) &= (R - \eta) + \sqrt{(R+\eta)^2 - 4\eta|\rho|^2 R} \\ l(\eta) &= -\frac{\lambda_2(\eta)}{\lambda_1(\eta)}. \end{aligned}$$

Thus, η can be computed by finding the zero of the function

$$h_2(\eta) = F_{\hat{r}}(\eta; R, |\rho|, N) - P_{FA}. \tag{5.7}$$

5.2 Computing Thresholds for Two-Stage Detector

Given the probability of false alarm P_{FA}, the goal is to choose thresholds η_1 and η_2 such that the joint pdf of $|\hat{\rho}_a|$ and $\hat{r}$ under H_0 integrates to P_{FA} over the region Z_1. Considering the form of Z_1 for any of the two single-dimensional detectors $\hat{r}$ and $|\hat{\rho}_c|$, we can see that a given P_{FA} uniquely determines the region Z_1 (and hence η) for these two detectors. This is a consequence of the fact that the cdf of a continuous random variable is monotonically increasing. However, for the case of the two-dimensional detector, the decision space is partitioned in a way that there can exist a number of regions Z_1 such that when we integrate $p(|\hat{\rho}_a|, \hat{r}; H_0)$ over any such Z_1, we get the P_{FA}. In other words, a given P_{FA} does not uniquely determine η_1 and η_2. We require some additional information apart from the given P_{FA} in order to uniquely determine η_1 and η_2. In order to develop a method for threshold selection, we further partition Z_1 into two sets, Z_{11} ($0 \leq \hat{r} \leq \eta_1, 0 \leq |\hat{\rho}_a| \leq 1$) and Z_{12} ($\eta_1 \leq \hat{r} \leq 1, 0 \leq |\hat{\rho}_a| \leq \eta_2$), as shown in Figure 5.1. The region Z_1 must be such that:

$$P_{FA} = \iint\limits_{Z_1} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} = \iint\limits_{Z_{11}} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} + \iint\limits_{Z_{12}} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} \tag{5.8}$$

At this point, we introduce a parameter $\alpha \in [0, 1]$, which denotes the fraction of the total P_{FA} which is due to Z_{11}. In other words, if we integrate $p(|\hat{\rho}_a|, \hat{r}; H_0)$ over Z_{11}, we get αP_{FA}.

$$\begin{aligned} \alpha P_{FA} = \iint\limits_{Z_{11}} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} &= \int\limits_0^{\eta_1} \int\limits_0^{1} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} \\ &= \int\limits_0^{\eta_1} p(\hat{r}; H_0) d\hat{r} \\ &= F(\eta_1; H_0) \end{aligned} \tag{5.9}$$

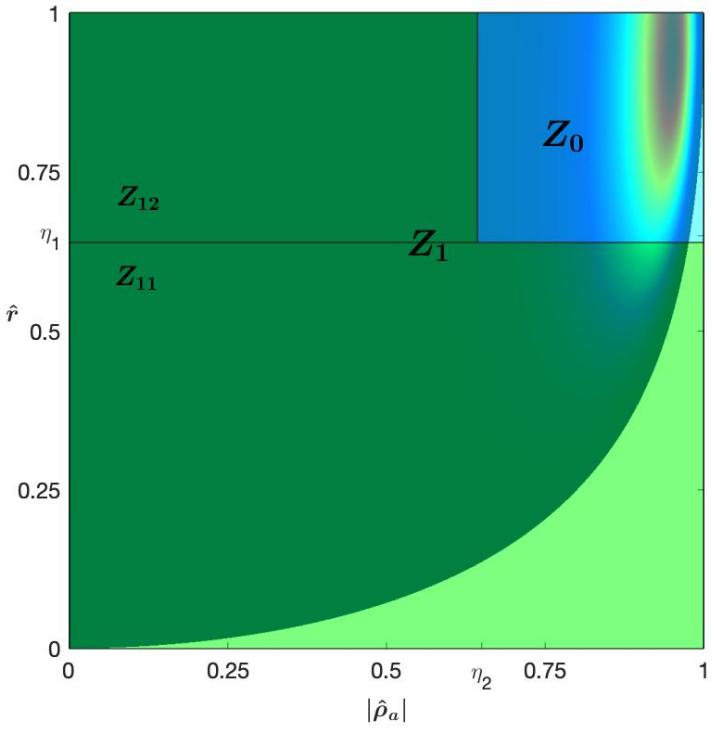


Figure 5.1: Partitions Z_{11} and Z_{12} of the Region Z_1.

where the subscript $\hat{r}$ has been dropped and now $F(\hat{r}; H_0)$ denotes the cdf of $\hat{r}$ as given in (5.4). Thus, since we have the closed form expression for $F(\hat{r}; H_0)$, we can find η_1, which is the value of $\hat{r}$ for which $F(\hat{r}; H_0) = \alpha P_{FA}$. This is equivalent to finding the threshold for the first stage detector $\hat{r}$ with the fixed probability of false alarm constraint equal to αP_{FA}, as discussed in the previous section. Note that when we fix α, η_1 is uniquely determined for the given P_{FA}. Since η_1 is now known, we go back to (5.8) in order to find η_2.

$$P_{FA} = \iint_{Z_1} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} = 1 - \iint_{Z_0} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r}$$
$$= 1 - \int_{\eta_1}^{1} \int_{\eta_2}^{1} p(|\hat{\rho}_a|, \hat{r}; H_0) d|\hat{\rho}_a| d\hat{r} \tag{5.10}$$

The only unknown quantity in the above equation is η_2, and it is found by computing the zero of the function

$$h(\eta_2) \triangleq \int_{\eta_1}^{1} \int_{\eta_2}^{1} p(|\hat{\rho}_a|, \hat{r}) d\,|\hat{\rho}_a|\, d\hat{r} - (1 - P_{FA})$$

by the Newton-Raphson method, where η_2 is the running variable, and the double integral is computed numerically in MATLAB for each η_2.

It is important to note that in this analysis, we initially fixed a value of α, and computed the thresholds for that α. In the beginning of this section, it was mentioned that a given P_{FA} does not uniquely determine η_1 and η_2. However, we can see that if α is specified in addition to the P_{FA}, the thresholds are uniquely determined. Thus, in order to implement this approach of threshold selection, we must fix the fraction of the total P_{FA} which is contributed by Z_{11}. The dependence of this parameter α on the detection performance is discussed in the next chapter.

In this chapter, the distribution of various detectors under H_0 was used to determine the thresholds. The pdfs under H_0 are obtained when we use the values of R and $|\rho|$ according to the H_0 scenario of the hypothesis test formulations given in (3.1), (3.4), and (3.9). In order to determine thresholds for $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$, it is obvious that some $|\rho| \approx 1$ should be assumed under H_0. However, in the case of $\hat{r}$, the pdf depends on R and $|\rho|$ both. Many studies including [3] have used the pdf of $\hat{r}$ with $R = 1$ and $|\rho| = 0$ to determine the thresholds. However, since $|\rho| = 0$ itself depicts an H_1 scenario, using this value of $|\rho|$ under H_0 might prove to degrade the detection performance for $\hat{r}$. This can also be seen in the hypothesis test for the two-stage detector given in (3.9), where under H_0, $R = 1$ and $|\rho| \approx 1$. Thus, throughout this study, $R = 1$ and some $|\rho| \approx 1$ are assumed under H_0 and are used for threshold selection for $\hat{r}$ and the two-stage detector.

Chapter 6

DETECTION PERFORMANCE FOR MODIFIED TWO-STAGE CHANGE DETECTOR

The detection performance of different detectors can be compared by generating their Receiver Operating Characteristic (ROC) curves. The ROC curve for a detector is generated by plotting it's P_D as a function of P_{FA}. The detector which can achieve a higher P_D for the same P_{FA} gives a better performance. In other words, the detector for which the ROC curve lies above is the one which performs better. In this chapter, the ROC curves for the three detectors discussed previously are compared for different values of the true parameters R and $|\rho|$. Also, the dependence of P_D on α for different values of R and $|\rho|$ is analyzed.

In order to generate the ROC curves for a test statistic T, the probability distribution of the test statistic under both the hypotheses, i.e. $p(T; H_0)$ and $p(T; H_1)$ should be known. In other words, if the probability distribution depends on a parameter θ, then the value of θ under H_0 and H_1 must be known. In the problem of change detection, the detectors which we deal with depend on the parameters R, $|\rho|$, and N. Out of these, N is just the window size, and it does not give any information on the underlying SAR returns. R and $|\rho|$ are the parameters which define change or no change, as seen in the hypothesis testing formulations for the non-coherent, CCD, and the two-stage change detection methods in (3.1), (3.4), and (3.9) respectively. In the case of R, the H_0 condition is fixed as $R = 1$, and the H_1 condition can be represented by any R significantly less than or greater than one. However, for $|\rho|$, the H_1 condition is assumed to be fixed as $|\rho| = 0$, and the H_0 condition can have any $|\rho| \approx 1$. As mentioned in [3], such an assumption is valid for fine resolution systems,

and is used in this section. Thus, throughout this section, the two main types of H_1 scenarios of interest are - 1) change due to $|\rho|$ only, i.e. $R = 1$ and $|\rho| = 0$, and 2) change due to R and $|\rho|$ both, i.e. $R \neq 1$ and $|\rho| = 0$.

The ROC curves for a test statistic T are generated as follows:

- Using the given P_{FA} constraint and the distribution $p(T; H_0)$, the region Z_1 (i.e. the thresholds) is determined, as shown in Chapter 5.
- The probability of detection (P_D) corresponding to the P_{FA} constraint is determined by integrating $p(T; H_1)$ over Z_1, as given in (2.6).
- The two steps given above are repeated for a range of P_{FA} constraints, and the ROC curves are generated by plotting the P_D as a function of the P_{FA}.

6.1 Generating ROC Curves for $|\hat{\rho}_c|$, $|\hat{\rho}_a|$ and $\hat{r}$

Once the threshold η is determined from the given P_{FA} constraint as shown in Section 5.1, the probability of detection (P_D) for the sample coherence statistic can be determined using:

$$P_D = \int_{Z_1} p(|\hat{\rho}_c|; H_1) d|\hat{\rho}_c| = \int_0^{\eta} p(|\hat{\rho}_c|; H_1) d|\hat{\rho}_c|. \tag{6.1}$$

The ROC curves for $|\hat{\rho}_c|$ (solid lines) are plotted in Figure 6.1 for $N = 6$, where the value of $|\rho|$ under H_1 (as per (3.4)) is fixed as $|\rho| = 0$, and the different curves correspond to different values of $|\rho|$ assumed under H_0. We can see that as the value of $|\rho|$ under H_0 increases, the performance of the detector becomes better. This is because as $|\rho|$ under H_0 moves further away from 0, the pdfs under H_0 and H_1 which are given by $p(|\hat{\rho}_c|; H_0)$ and $p(|\hat{\rho}_c|; H_1) = p(|\hat{\rho}_c|; |\rho| = 0)$ get separated further away from each other, which makes it easier to detect which of the two distributions generated a given data sample. Note that the value of R under both the hypotheses

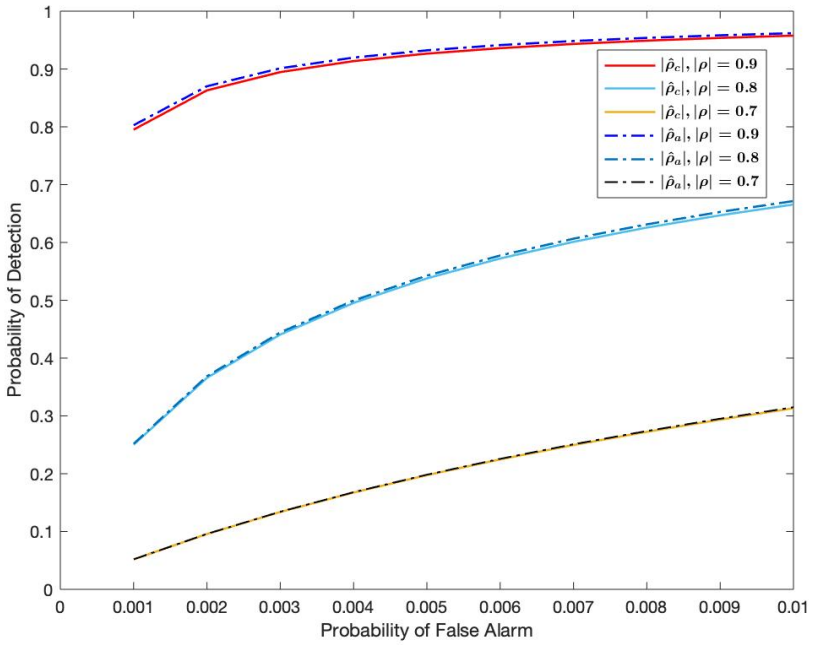


Figure 6.1: ROC Curves for $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ for Change Scenario H_1 Given by $|\rho| = 0$. The Legend Indicates Value of $|\rho|$ under H_0.

was not mentioned above, because the pdf of $|\hat{\rho}_c|$ is independent of the true underlying variances, and thus is independent of R, as seen in (3.6). Thus, $|\hat{\rho}_c|$ gives the same detection performance for any value of R under H_0 and H_1.

At this point, it is interesting to consider the detection performance of $|\hat{\rho}_a|$ as well. In Section 3.4, it was mentioned that $|\hat{\rho}_a|$ is an accurate estimator of $|\rho|$ only when the underlying variances are equal, i.e. $R = 1$. ROC curves for $|\hat{\rho}_a|$ are plotted as dashed-dot lines in Figure 6.1. The procedure for generating the ROC curves for $|\hat{\rho}_a|$ is the same as that of $|\hat{\rho}_c|$, the only difference is that the pdf in (3.8) is used instead of (3.6). We can clearly see that the performance of the two detectors is almost the same. It is important to note that (3.8) is valid only for $R = 1$, and thus the comparison between $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ in Figure 6.1 is valid only when $R = 1$, even though $|\hat{\rho}_c|$ is independent of R. In summary, when the change is due to $|\rho|$ only ($R = 1$ and $|\rho| = 0$ under H_1), $|\hat{\rho}_a|$ and $|\hat{\rho}_c|$ give the same detection performance.

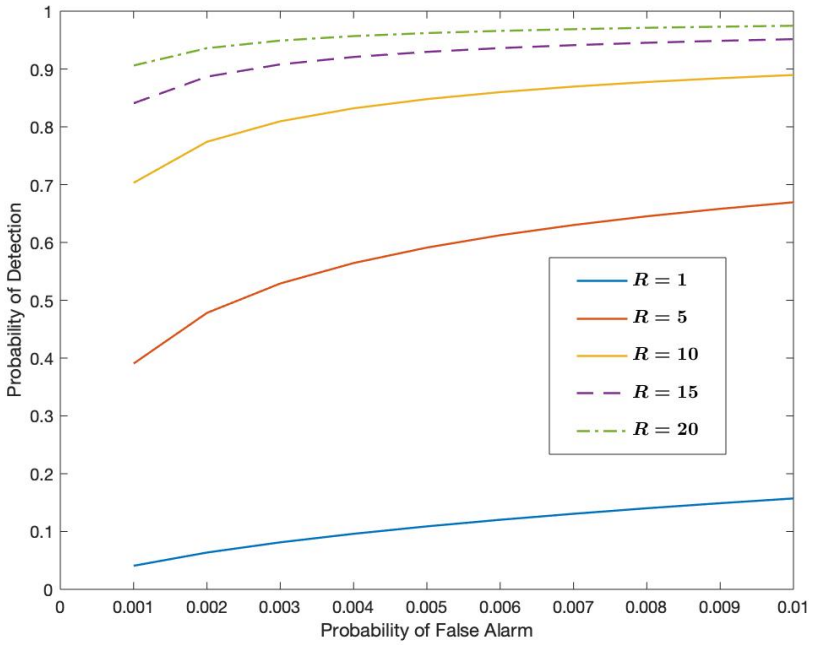


Figure 6.2: ROC Curves for $\hat{r}$ for the No Change Scenario H_0 Given by $R = 1$ and $|\rho| = 0.9$. Different ROC Curves are due to Different Values of R under H_1 where $|\rho|$ under H_1 is fixed to be zero.

Similarly, for the modified first stage $\hat{r}$, P_D can be computed using

$$P_D = \int_{Z_1} p(\hat{r}; H_1) d\hat{r} = \int_0^{\eta} p(\hat{r}; H_1) d\hat{r} \tag{6.2}$$

The parameter which characterized change for $|\hat{\rho}_c|$ was just $|\rho|$. However, in the case of $\hat{r}$, both R and $|\rho|$ characterize change. Figure 6.2 shows the comparison of ROC curves for $N = 3$ when the values of R and $|\rho|$ under H_0 are fixed to be 1 and 0.9 respectively. The ROC curves are plotted for fixed $|\rho| = 0$ and different values of R under H_1. As the value of R under H_1 deviates further from 1, the detection performance for $\hat{r}$ improves. We can see that for change due to $|\rho|$ only, $\hat{r}$ gives a very small value of P_D, and for values of R significantly larger (or smaller) than one, $\hat{r}$ achieves high values of P_D.

6.2 Generating ROC Curves for Two-Stage Detector

In case of the two-stage detector, for a given P_{FA} and α pair, the thresholds η_1 and η_2 can be determined as shown in the previous chapter. Thus, the probability of detection can be given as:

$$P_D = \iint_{Z_1} p(|\hat{\rho}_a|, \hat{r}; H_1) d|\hat{\rho}_a| d\hat{r} = 1 - \iint_{Z_0} p(|\hat{\rho}_a|, \hat{r}; H_1) d|\hat{\rho}_a| d\hat{r}$$
$$= 1 - \int_{\eta_1}^{1} \int_{\eta_2}^{1} p(|\hat{\rho}_a|, \hat{r}; H_1) d|\hat{\rho}_a| d\hat{r} \tag{6.3}$$

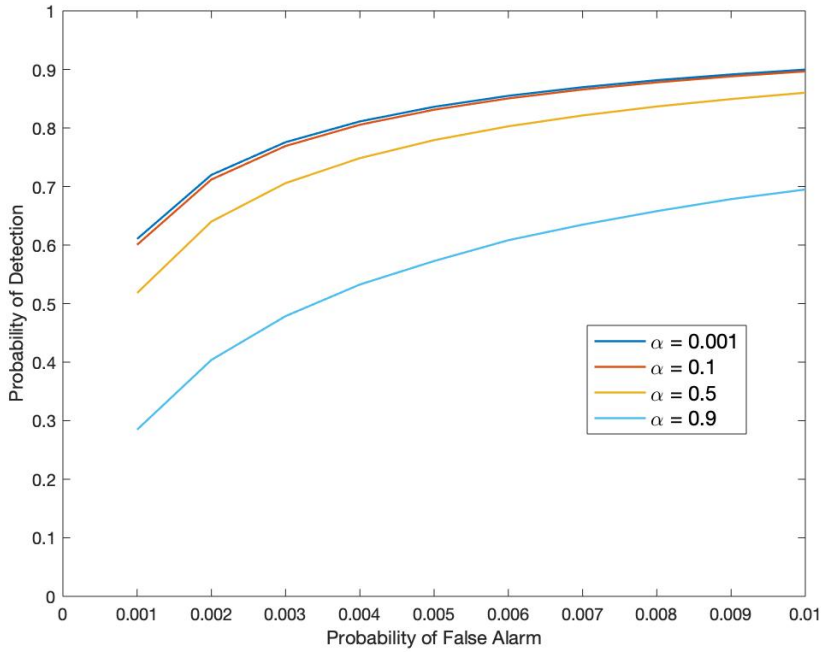


Figure 6.3: ROC Curves for Two-Stage Detector for Different Values of α when $R = 1$ and $|\rho| = 0$ under H_1.

Figure 6.3 shows ROC curves for the two-stage detector for $N = 5$, $R = 1$ and $|\rho| = 0.9$ under H_0, and $R = 1$ and $|\rho| = 0$ under H_1 for different values of α. A value of α chosen is constant for all P_{FA} in the ROC curve. We can see that as α

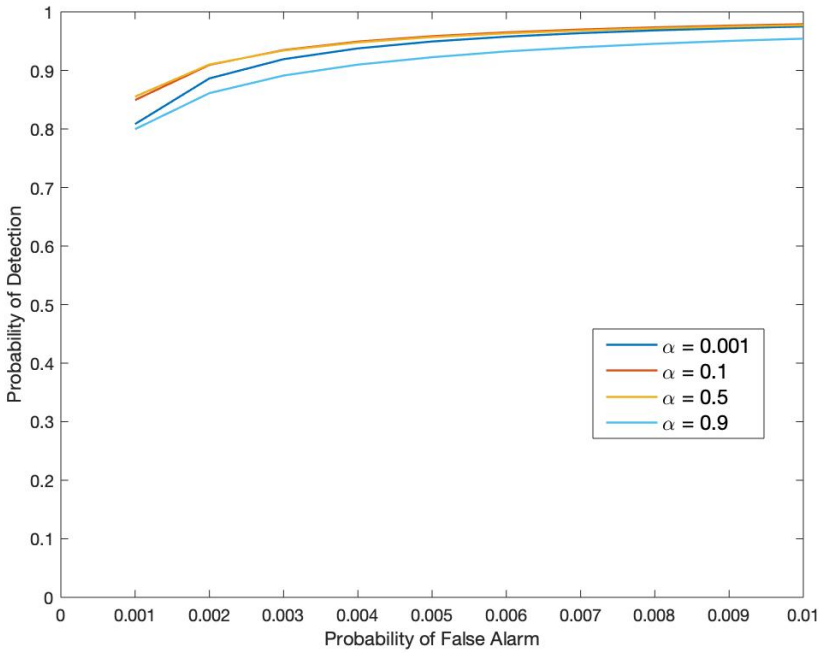


Figure 6.4: ROC Curves for Two-Stage Detector for Different Values of α when $R = 5$ and $|\rho| = 0$ under H_1.

increases, the performance of the two-stage detector degrades. However, when ROC curves are plotted for the H_1 case of $R = 5$ and $|\rho| = 0$, the detector behaves in a different manner, as we can see in Figure 6.4. The detection performance improves as we increase α from 0.001 to 0.1, remains the same as α goes from 0.1 to 0.5, and finally the performance degrades as α is further increased to 0.9. Thus, these two figures illustrate that the manner in which the probability of detection depends on α is different for different H_1 scenarios.

6.3 Behaviour of Two stage Detector for $\alpha = 0$ and $\alpha = 1$.

In order to further explore the dependence of P_D on α, it is important to understand the behaviour of the two-stage detector as α varies from 0 to 1. As mentioned earlier in Chapter 5, α is the fraction of the total P_{FA} which is allotted to Z_{11}. In other words, when the joint pdf of the two-stage detector under H_0 is integrated

over Z_{11}, we get αP_{FA}. It is evident from Figure 5.1 that as we decrease α, η_1 will also decrease. However, since the entire region Z_1 should integrate to P_{FA}, η_2 will increase as η_1 decreases. As α is made smaller, the rectangle representing the Z_0 region elongates in the vertical direction, as shown in Figure 6.5. As α approaches zero, η_2 stops changing and attains a constant value. Note that $\alpha = 0$ means that the contribution of Z_{11} to the total P_{FA} is zero. This effectively means that all the values of $\hat{r}$ are greater than η_1, and the two-stage detector is only comparing the value of $|\hat{\rho}_a|$ with η_2. Thus, α taking the value zero essentially reduces the two-stage detector to the single stage detector $|\hat{\rho}_a|$. Similarly, as α becomes equal to one, the two-stage detector reduces to $\hat{r}$. Therefore, as α goes from zero to one, the two-stage detector transitions from $|\hat{\rho}_a|$ to $\hat{r}$. This can be seen from Figure 6.6, where ROC curves are plotted for $\hat{r}$, $|\hat{\rho}_a|$, and the two-stage detector for $\alpha = 1$ and $\alpha = 0.001$ for the case of $N = 4$, $R = 1$ under both hypotheses, and the value of $|\rho|$ is 0.9 and 0 under H_0 and H_1 respectively.

An important point must be considered regarding the generation of theoretical ROC curves for $|\hat{\rho}_a|$. Equation (3.8) gives the pdf of $|\hat{\rho}_a|$ only when the underlying variances are equal, i.e. $R = 1$. Thus, the pdf in (3.8) can be used to generate the ROC curves for $|\hat{\rho}_a|$ only if $R = 1$ under both hypotheses. If $R \neq 1$ under H_1, we cannot use the same procedure to generate the ROC curves. In that case, the correct procedure to generate the ROC curves for $|\hat{\rho}_a|$ would be to generate the ROC curves for the two-stage detector for $\alpha \approx 0$ (say 0.001), because the pdf of $|\hat{\rho}_a|$ when $R \neq 1$ can be obtained by integrating the joint pdf of $|\hat{\rho}_a|$ and $\hat{r}$ over $\hat{r} = 0$ to $\hat{r} = 1$. In Section 6.1, it was shown that the detection performance of $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ is the same when $R = 1$ under both hypothesis. However, unlike $|\hat{\rho}_c|$, the detection performance of $|\hat{\rho}_a|$ is not invariant to R, and thus for $R \neq 1$, the detection performance of $|\hat{\rho}_a|$ and $|\hat{\rho}_c|$ will not be the same.

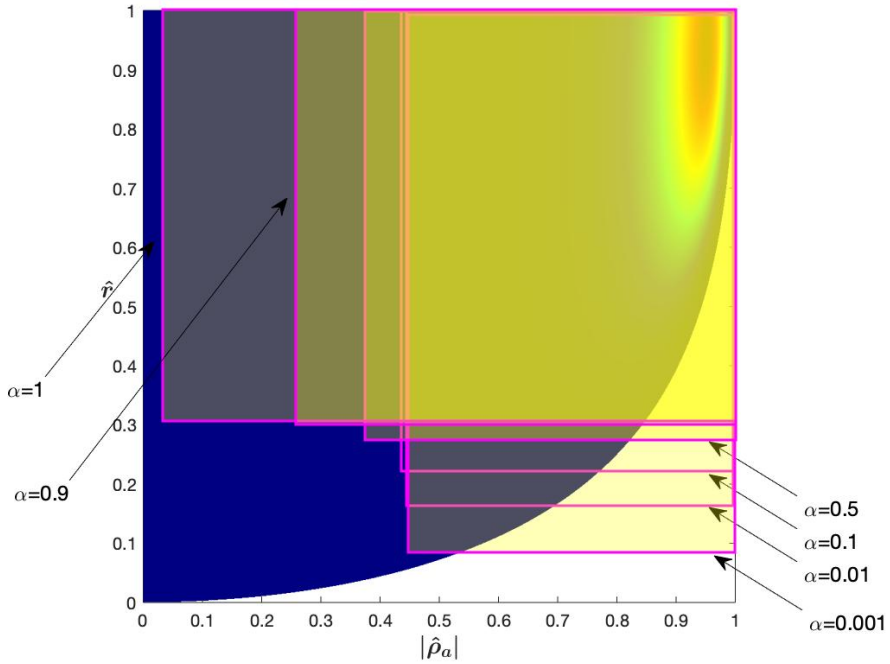


Figure 6.5: Transition of the Region Z_0 as α Varies from Zero to One.

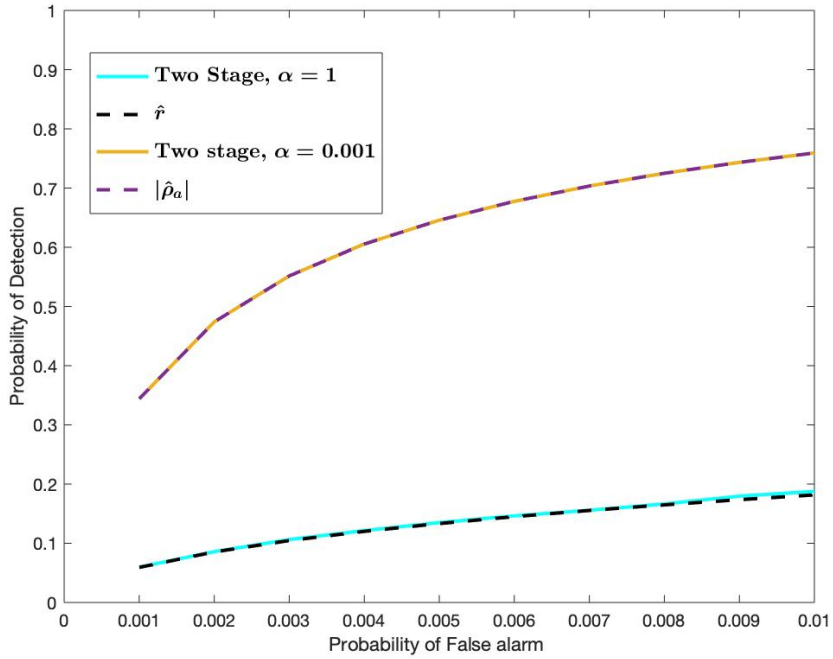


Figure 6.6: ROC Curves Showing that the Two-Stage Detector for $\alpha = 1$ and $\alpha \approx 0$ give the Same P_D as $\hat{r}$ and $|\hat{\rho}_a|$ respectively.

6.4 Dependence of P_D on α

Plots of P_D as a function of α for a fixed P_{FA} have been shown in this section for the two types of change scenarios (change due to only $|\rho|$ and change due to both R and $|\rho|$). Throughout this section, the thresholds are computed using the H_0 values $R = 1$ and $|\rho| = 0.9$, and the value of the spatial window size is taken as $N = 5$. Figure 6.7 shows α vs P_D for $P_{FA} = 0.001$, where the H_1 scenario is given by $|\rho| = 0$ and $R = 1$. The curve suggests that as α takes values closer to 1, the probability of detection is very low. This P_D starts increasing as α is brought closer to 0, and the maximum P_D is attained when $\alpha = 0$. It was shown in the previous section that when $\alpha = 0$, the two-stage detector gives the same performance as $|\hat{\rho}_a|$. Furthermore, since $|\hat{\rho}_a|$ gives a performance almost equal to $|\hat{\rho}_c|$ (as mentioned in Section 6.1) when $R = 1$ under both hypotheses, we can conclude that when the change is due to $|\rho|$ only, the two-stage detector for $\alpha \approx 0$ gives the highest probability of detection, and this P_D is equal to that achieved by $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$. Also, we can observe that when the two-stage detector takes the form of $\hat{r}$ (i.e. when $\alpha = 1$), the probability of detection is very low. This fact that $\hat{r}$ performs very poorly when the change is due to $|\rho|$ only was also mentioned in the Section 6.1. Summarizing, when the change is given by $|\rho|$ only, that is, when $|\rho| = 0$ and $R = 1$ under H_1, $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ give the same P_D, which is significantly higher than $\hat{r}$, and this high value of P_D can be achieved by the two-stage detector when $\alpha \approx 0$.

Similarly, the α vs P_D plot for $|\rho| = 0$ and $R = 5$ under H_1 is shown in Figure 6.8 for $P_{FA} = 0.001$. In this case, the value of α which gives maximum P_D is somewhere around 0.3. Similarly for $|\rho| = 0$ and $R = 10$ under H_1, the optimal value of α was found to be 0.47, as shown in Figure 6.9. Thus, when the change is given by both $|\rho|$ and R, the optimal value of α is a number between 0 and 1, and not 0 as in the

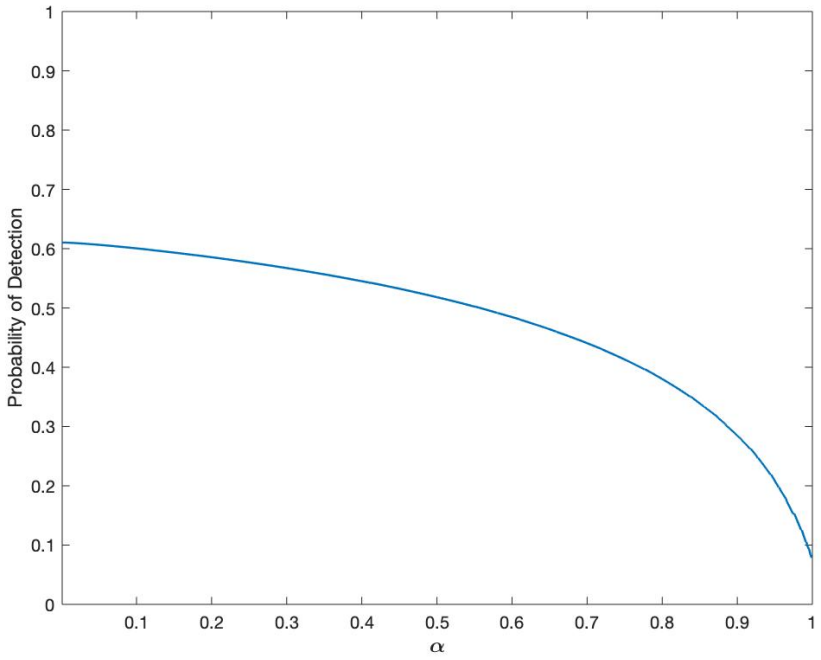


Figure 6.7: Plot of α vs P_D for $P_{FA} = 0.001$ when $R = 1$ and $|\rho| = 0$ under H_1 (i.e. Change Due to $|\rho|$ only).

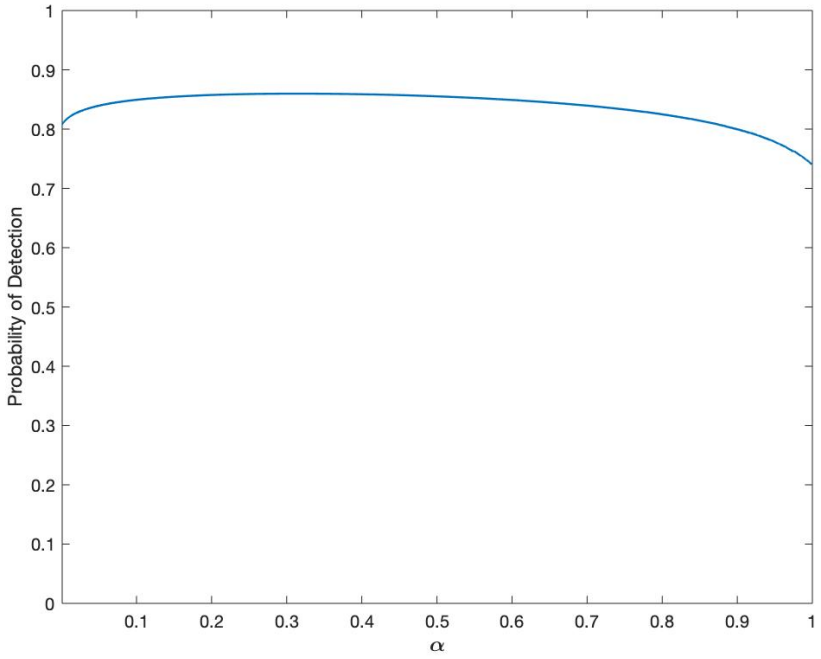


Figure 6.8: Plot of α vs P_D for $P_{FA} = 0.001$ when $R = 5$ and $|\rho| = 0$ under H_1 (i.e. Change Due to $|\rho|$ and R both).

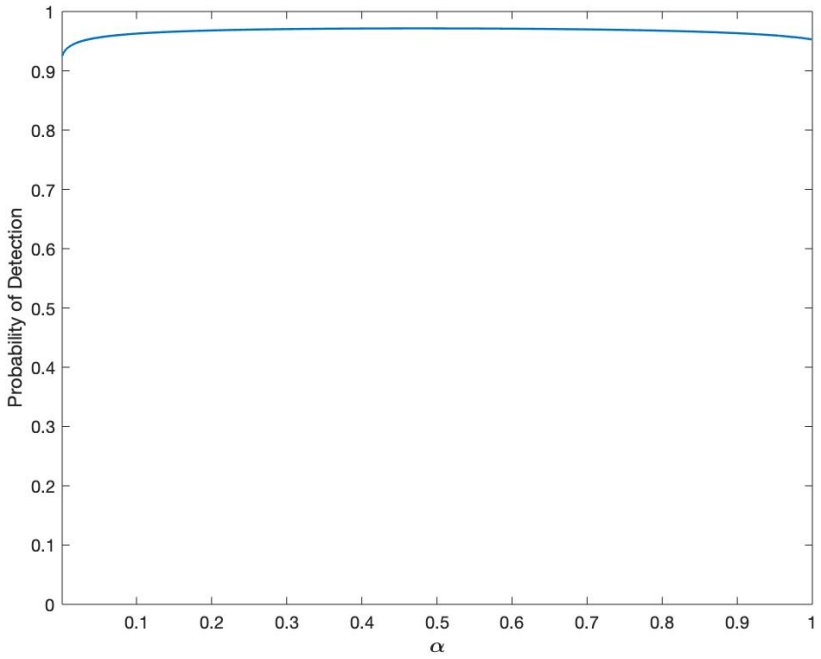


Figure 6.9: Plot of α vs P_D for $P_{FA} = 0.001$ when $R = 10$ and $|\rho| = 0$ under H_1 (i.e. Change Due to $|\rho|$ and R both).

case when change is due to $|\rho|$ only. Figures 6.10 and 6.11 show the optimal value of α for different values of $|\rho|$ and R under H_1 for $P_{FA} = 0.001$ and $P_{FA} = 0.0001$ respectively, when $N = 5$. The two types of change which we are interested in are represented by the top-left ($R \approx 1$ and $|\rho| \approx 0$) and bottom-left regions ($R \neq 1$ and $|\rho| \approx 0$). We can see that when change is due to $|\rho|$ only (top-left), the optimal value of α is approximately zero, and this value increases as the value of R under H_1 starts increasing. When R under H_1 deviates significantly from 1, the optimal value of α approaches 1. Thus, when $R = 1$ under H_1, the optimal performance is achieved by $|\hat{\rho}_a|$, and when R under H_1 is significantly greater or lesser than 1, the optimal performance is achieved by $\hat{r}$. For intermediate values of R, the optimal performance is achieved by the two-stage detector where the optimal value of α lies between zero and one.

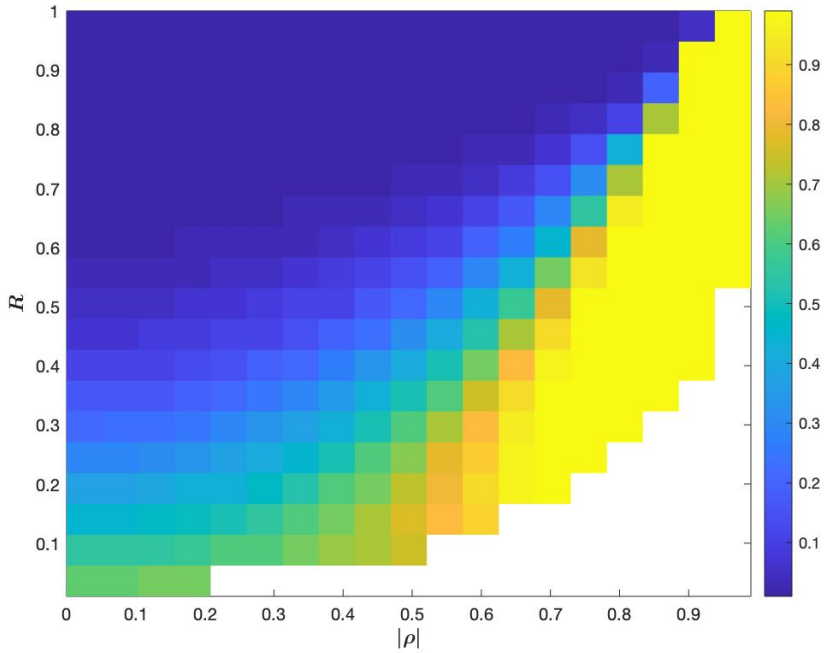


Figure 6.10: Optimal Value of α for Different Values of R and $|\rho|$, for $P_{FA} = 0.001$ and $N = 5$.

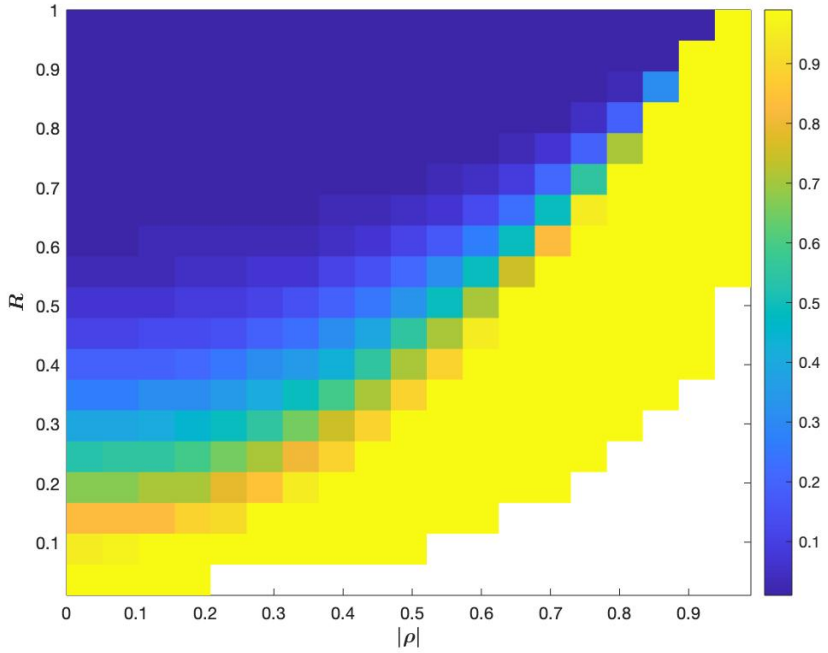


Figure 6.11: Optimal Value of α for Different Values of R and $|\rho|$, for $P_{FA} = 0.0001$ and $N = 5$.

It is important to determine whether this optimal α (α_{opt}) gives an improved performance as compared to $|\hat{\rho}_c|$ or $\hat{r}$. Figures 6.12 and 6.13 illustrate the improvement in P_D achieved by the two-stage detector using the optimal value of α, as compared to $|\hat{\rho}_c|$ and $\hat{r}$ respectively. Figure 6.12 shows the difference in P_D obtained by subtracting the P_D achieved by $|\hat{\rho}_c|$ from the P_D achieved by the two-stage detector using the optimal value of α for every true (R,$|\rho|$) pair under H_1, for $P_{FA} = 0.001$. A similar plot showing the comparison of $\hat{r}$ and the two-stage detector is shown in Figure 6.13. The first observation which we can make is that the difference is always non-negative, which suggests that using this α_{opt} for the two-stage change detector will always give a better (or at least the same) performance as compared to $|\hat{\rho}_c|$ or $\hat{r}$. As discussed earlier, when the change is due to $|\rho|$ only (i.e. when $R \approx 1$ and $|\rho| = 0$), the two-stage detector can only perform as good as $|\hat{\rho}_c|$ (or $|\hat{\rho}_a|$). This can be seen in the top-left region of Figure 6.12, as the improvement is almost close to zero. Significant improvement can be seen in the bottom-left region, where change is due to R and $|\rho|$ both. Similarly, as discussed earlier, $\hat{r}$ gives a poor performance when change is due to $|\rho|$ only, and thus we can see the significant improvement achieved by the two-stage detector in the top-left region of Figure 6.13. Also, the bottom-left region shows only a slight improvement, which suggests that when change is due to $|\rho|$ and R both, the two-stage detector with the optimal α shows only a slight improvement over $\hat{r}$.

Figure 6.10 shows the optimal value of α for given R and $|\rho|$ under H_1. In real applications, these values R and $|\rho|$ are not known beforehand, and thus we cannot use this information to select α_{opt} based on Figure 6.10. However, we can observe Figures 6.8 and 6.9 (i.e. the case when change is due to R and $|\rho|$ both) and conclude that, even though the optimal value of α is an arbitrary number between 0 and 1, the P_D achieved by setting $\alpha = 0$ is not significantly less than the highest achievable P_D. Additionally, it was shown earlier that when change is due to $|\rho|$ only, the

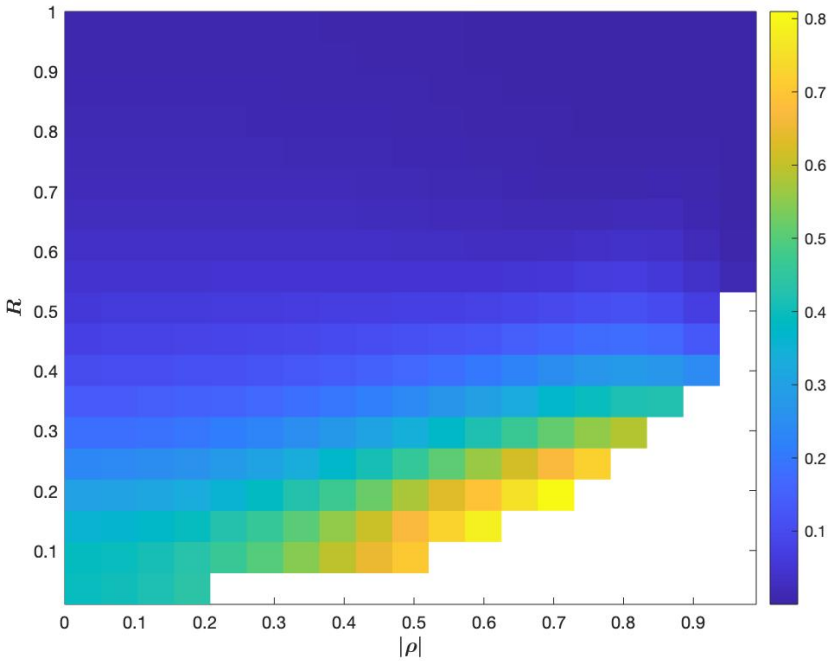


Figure 6.12: Two-Stage Detector with Optimal α vs $|\hat{\rho}_c|$ for $P_{FA} = 0.001$

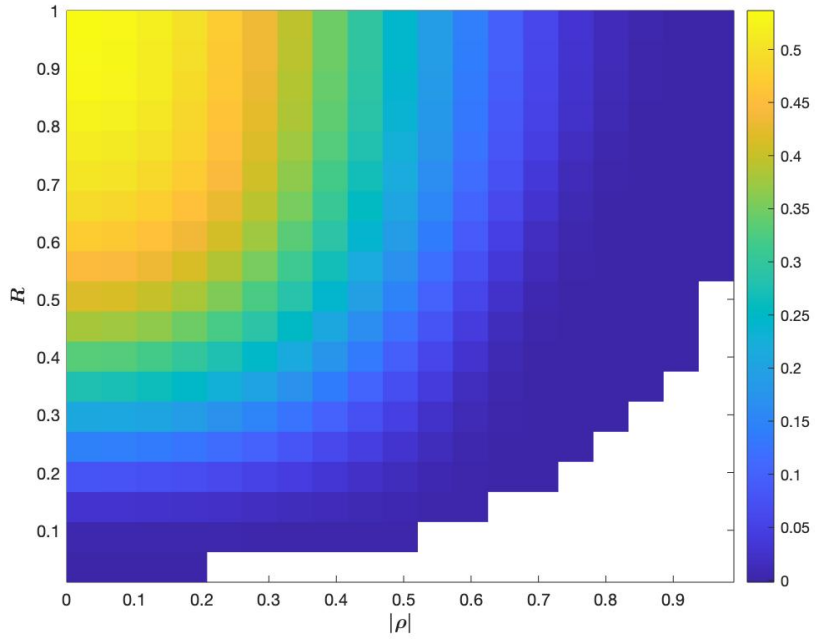


Figure 6.13: Two-Stage Detector with Optimal α vs $\hat{r}$ for $P_{FA} = 0.001$

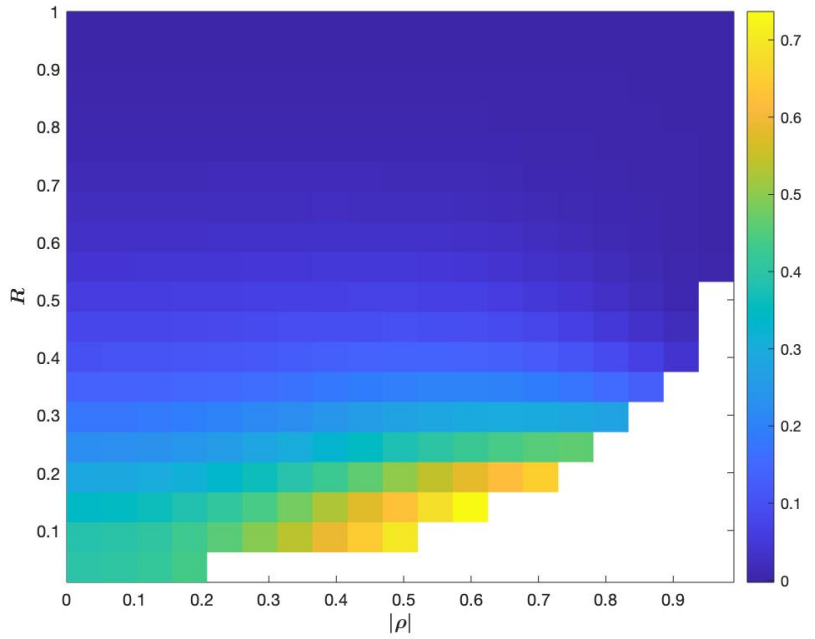


Figure 6.14: Two-Stage Detector with $\alpha = 0.1$ vs $|\hat{\rho}_c|$ for $P_{FA} = 0.001$.

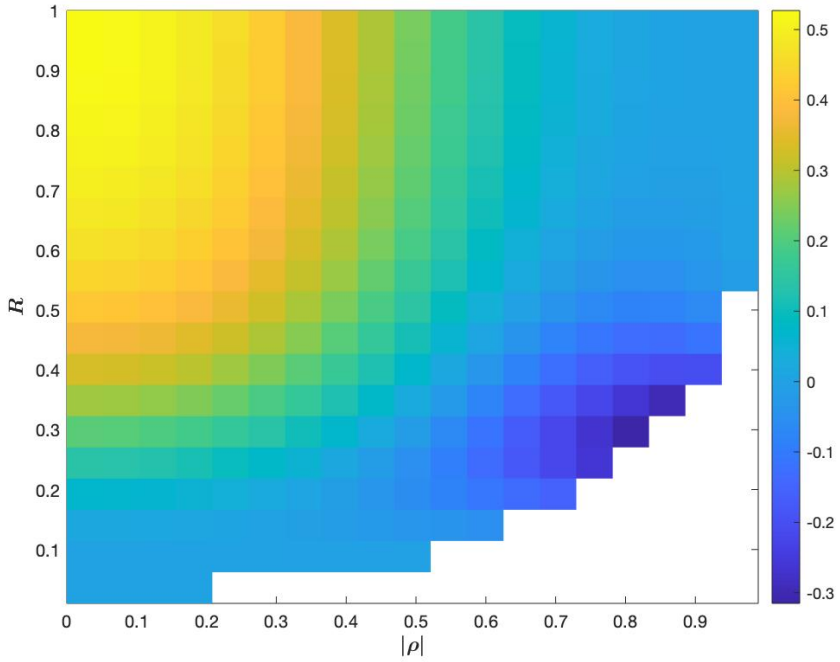


Figure 6.15: Two-Stage Detector with $\alpha = 0.1$ vs $\hat{r}$ for $P_{FA} = 0.001$.

optimal value of α is zero. Thus, we can conclude that using $\alpha \approx 0$ for the two-stage detector (which effectively is $|\hat{\rho}_a|$) for any value of R and $|\rho|$ under H_1 will give a significantly improved overall performance as compared to using $|\hat{\rho}_c|$ or $\hat{r}$ individually. Furthermore, we can see that in Figures 6.8 and 6.9, the P_D slightly drops as we approach zero. In order to avoid this, we can use a value of $\alpha = 0.1$, considering the fact that the P_D for $\alpha = 0.1$ is not significantly less as compared to $\alpha = 0$ in Figure 6.7. As a result, we can expect that the two-stage detector with $\alpha \approx 0.1$ can provide an higher overall P_D as compared to $|\hat{\rho}_c|$ and $\hat{r}$. However, if a computationally simpler detector is preferred, the single dimensional detector $|\hat{\rho}_a|$ can be used, which also gives an improved performance over $|\hat{\rho}_c|$ and $\hat{r}$.

Similar to Figures 6.12, the difference between the P_D achieved by the two-stage detector and $|\hat{\rho}_c|$ is shown in Figure 6.14. The only point of difference is that for all values of R and $|\rho|$ under H_1, the value of α used is 0.1, and not the optimal value of α. Similar to Figure 6.12, we can see that the difference in P_D is always positive, suggesting that the two-stage detector with $\alpha = 0.1$ always gives a performance which is at least as good as $|\hat{\rho}_c|$. Additionally, in the bottom-left part of Figure 6.14 where the change is due to $|\rho|$ and R both, the improvement shown by the two-stage detector as compared to $|\hat{\rho}_c|$ is significantly high. Similarly, Figure 6.15 shows the difference between the P_D achieved by the two-stage detector with $\alpha = 0.1$ and that achieved by $\hat{r}$. When change is given by $|\rho|$ only, we can see a significant improvement, whereas when change is due to $|\rho|$ and R both, the P_D is almost the same in both the cases. The region where the difference is negative can be assumed to be very less likely to occur, because higher values of $|\rho|$ are unlikely when R starts to deviate from 1, as explained in [2].

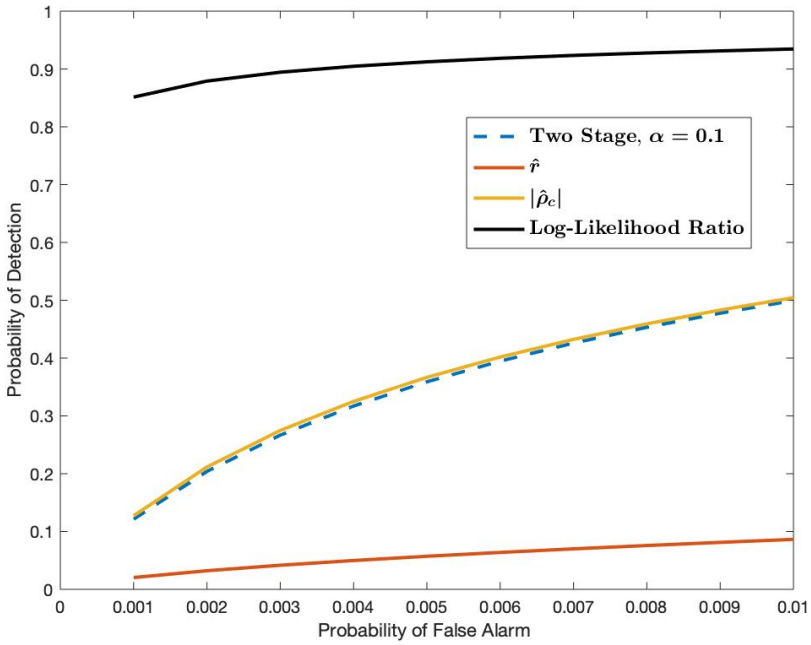


Figure 6.16: Comparison of ROC Curves for the Case $\sigma_{f0} = \sigma_{f1} = \sigma_{g0} = \sigma_{g1} = 1$ ($R_0 = 1$ and $R_1 = 1$), $\rho_0 = 0.8$, and $\rho_1 = 0$

6.5 Comparison of ROC Performance for Different Change Detectors

As mentioned in Section 2.2, the optimal test statistic is given by the likelihood ratio, which maximizes the probability of detection for a given probability of false alarm. The log-likelihood test statistic for this problem of change detection has been derived in [3]. The P_D achieved by the log-likelihood ratio is an upper bound for the P_D achievable by any other test statistic. Thus, comparing the detectors discussed in this study with the log-likelihood ratio tells us how well do these detectors perform in comparison with the highest achievable performance.

Figures 6.16 - 6.19 show the comparison of ROC curves between the sample coherence $|\hat{\rho}_c|$, the sample variance ratio $\hat{r}$, the log-likelihood test statistic, and the two-stage change detector for $\alpha = 0.1$ for different scenarios assumed under H_0 and H_1, for the case of $N = 5$. Note that unlike the detectors discussed in this study, the

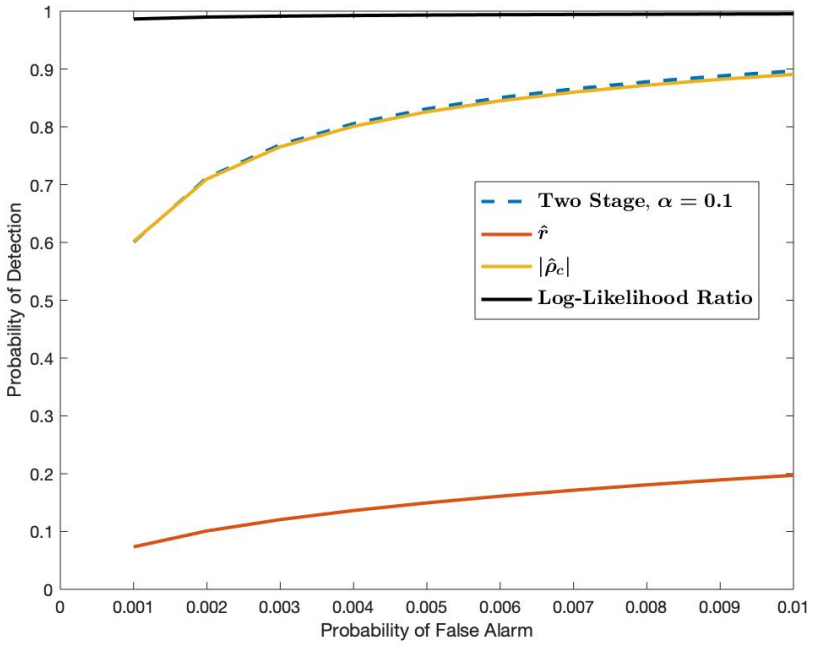


Figure 6.17: Comparison of ROC Curves for the Case $\sigma_{f0} = \sigma_{f1} = \sigma_{g0} = \sigma_{g1} = 1$ ($R_0 = 1$ and $R_1 = 1$), $\rho_0 = 0.9$, and $\rho_1 = 0$

log-likelihood test statistic depends directly on the values σ_{f0}, σ_{g0}, σ_{f1}, σ_{g1}, ρ_0, and ρ_1 (the subscript i indicates the value of the parameter under H_i), i.e. the dependence on these terms is not only through R and $|\rho|$.

Figures 6.16 and 6.17 show the comparison of ROC curves for the four detectors for the case when change is given by ρ only. Since we know that under these conditions, $|\hat{\rho}_c|$ and $|\hat{\rho}_a|$ give the same performance, the ROC curves for $|\hat{\rho}_a|$ have not been plotted in these two figures, in order to make these plots easily readable. In both the figures, ρ under H_1 is zero, whereas $|\rho|$ under H_0 is 0.8 and 0.9 in Figures 6.16 and 6.17 respectively. By comparing the two figures, we can see that when the values of ρ under H_0 and H_1 are further away from each other, the detection performance of each of these detectors is better. Furthermore, it was shown in the previous section, that under this type of change, the P_D achieved by the two-stage detector by setting $\alpha = 0.1$ is almost equal to that of $|\hat{\rho}_c|$ (or $|\hat{\rho}_a|$), and this can also be seen in Figures

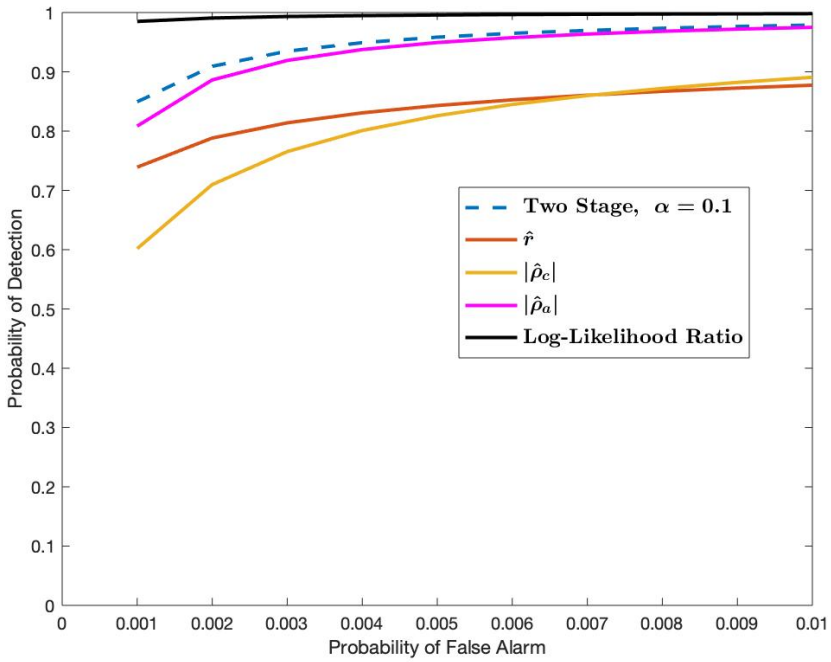


Figure 6.18: Comparison of ROC Curves for the Case $\sigma_{f0} = \sigma_{f1} = \sigma_{g0} = 1$, $\sigma_{g1} = \sqrt{0.2}$ ($R_0 = 1$ and $R_1 = 5$), $\rho_0 = 0.9$, and $\rho_1 = 0$

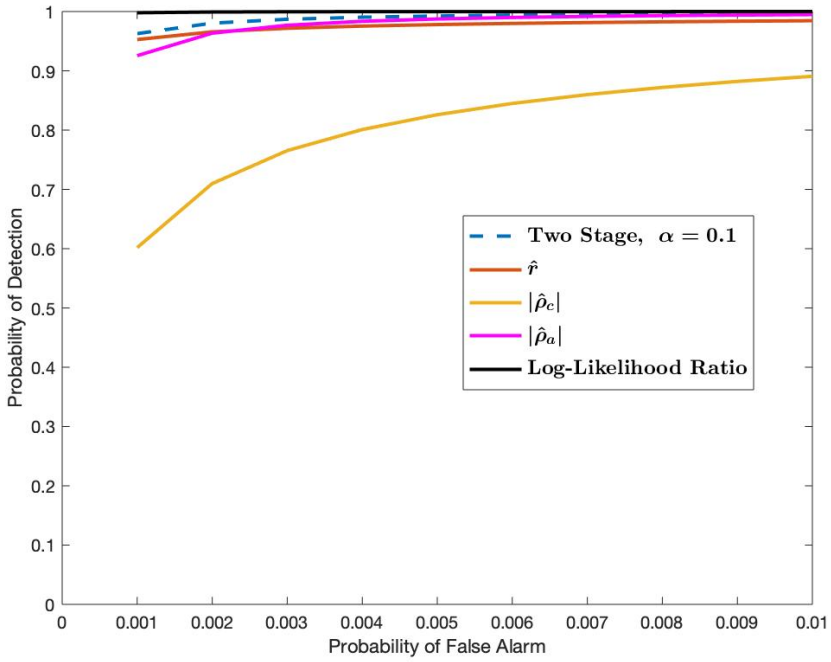


Figure 6.19: Comparison of ROC Curves for the Case $\sigma_{f0} = \sigma_{f1} = \sigma_{g0} = 1$, $\sigma_{g1} = \sqrt{0.1}$ ($R_0 = 1$ and $R_1 = 10$), $\rho_0 = 0.9$, and $\rho_1 = 0$

6.16 and 6.17. Thus, the two-stage detector, $|\hat{\rho}_c|$, and $|\hat{\rho}_a|$ display significantly better performance as compared to $\hat{r}$, which performs poorly. Figures 6.18 and 6.19 show the comparison of ROC curves when change is given by both R and ρ. By comparing Figures 6.16, 6.18, and 6.19, we can see that for fixed values of ρ under H_0 and H_1, as the value of R under H_1 deviates further from one, the performance of $\hat{r}$ starts improving, and eventually it outperforms $|\hat{\rho}_c|$ uniformly. The two-stage detector and $|\hat{\rho}_a|$ achieve a higher P_D as compared to both $|\hat{\rho}_c|$ and $\hat{r}$. The increase in P_D achieved by the two-stage detector with respect to $|\hat{\rho}_c|$ increases as R_1 deviates further from one, whereas the increase in P_D of the two-stage detector with respect to $\hat{r}$ decreases as R_1 deviates further from one. In all of the cases, it is evident that the log-likelihood statistic achieves higher P_D as compared to any other detector.

Thus, the comparison of the ROC curves between $|\hat{\rho}_c|$, $|\hat{\rho}_a|$, $\hat{r}$, and the two-stage detector for $\alpha = 0.1$ have been shown in this section. Summarizing, for the case when change is due to $|\rho|$ only, $|\hat{\rho}_c|$ (and $|\hat{\rho}_a|$) outperforms $\hat{r}$ significantly, and the two-stage detector achieves performance almost equal to that of $|\hat{\rho}_c|$. When the change is given by R and $|\rho|$ both, $\hat{r}$ starts outperforming $|\hat{\rho}_c|$ as R under H_1 starts deviating further from one. In this case, the two-stage detector and $|\hat{\rho}_a|$ perform better than $\hat{r}$. Thus, the two-stage detector for $\alpha = 0.1$ and $|\hat{\rho}_a|$ give a better overall detection performance as compared to $\hat{r}$ and $|\hat{\rho}_c|$, where the two-stage detector performs slightly better than $|\hat{\rho}_a|$.

The overall improvement in detection performance shown by $|\hat{\rho}_a|$ as compared to $|\hat{\rho}_c|$ and $\hat{r}$ is both unexpected and interesting. In Section 3.3, it was mentioned that $|\hat{\rho}_a|$ is an accurate estimator of $|\rho|$ only under the condition $R = 1$. However, we have seen in this section that $|\hat{\rho}_a|$ gives a significantly higher probability of detection as compared to $|\hat{\rho}_c|$ when $R \neq 1$. This unexpected improvement in performance is explained in [2], and is because as R deviates from one, the arithmetic mean present in

the denominator of (3.7) becomes greater than the geometric mean in the denominator of (3.5), and thus $|\hat{\rho}_a|$ becomes lesser than $|\hat{\rho}_c|$. Since $|\hat{\rho}_a| \leq |\hat{\rho}_c|$, it becomes easier to detect change using $|\hat{\rho}_a|$ as compared to $|\hat{\rho}_c|$, considering the form of Z_1 for both these detectors. Thus, even though $|\hat{\rho}_a|$ does not provide accurate estimates of the coherence when $R \neq 1$, it gives better change detection performance as compared to $|\hat{\rho}_c|$.

Chapter 7

EXPERIMENTAL RESULTS

In this chapter, the four change detectors are implemented on a SAR data set, made available publicly by the Air Force Research Laboratory, as described in [17]. Figures 7.1 and 7.2 are the two SAR images, and the goal is to generate an image which shows whether a pixel has undergone change. It is important to note that throughout this entire thesis, binary thresholding algorithms have been discussed, which classify a pixel as change or no change. Thus, applying the detectors to the SAR data set must result in a binary (black and white) image, where a black pixel denotes change and a white pixel denotes no change. However, Figure 7.3 is a gray scale image, where every pixel contains the actual value of $|\hat{\rho}_c|$ for that pixel, and this image can be used as a reference to determine areas of change or no change. Some areas of change have been described in [17], which are shown in Figure 7.3. Small clusters of dark pixels shown in the elliptical region correspond to car displacements, and the dark region showed near the top left corresponds to changes caused by moving foliage. Apart from these, other distinctive features in Figure 7.3 can also be considered as references, like the vertical elongated rectangle-like path situated above the elliptical region.

Figures 7.4 - 7.7 show the change images for $\hat{r}$, $|\hat{\rho}_a|$, $|\hat{\rho}_c|$, and the two-stage detector respectively, for $N = 4$. As discussed in the previous chapter, the two-stage detector with $\alpha = 0.1$ has been considered in this chapter. We can see that even though none of the detectors produce a good change image, $\hat{r}$, $|\hat{\rho}_a|$, and the two-stage detector do detect some areas of change, whereas $|\hat{\rho}_c|$ does not perform as good as the other three. As we increase the value of N to 9, all the detectors show a significant improvement in change detection, as seen in Figures 7.8 - 7.12. The car displacements can be thought

of as significant changes, and thus $\hat{r}$ does a better job in detecting them as compared to $|\hat{\rho}_c|$. The two-stage detector and $|\hat{\rho}_a|$ show considerably good change images in this case.

As we further increase the value of N to 12 and 16, all four detectors generate significantly good change images. Even for these higher values of N, $|\hat{\rho}_c|$ does not do as great a job as the other three detectors in detecting the car displacements. However, it detects the subtle changes due to moving foliage better than $\hat{r}$. This can be seen from the more densely-packed black pixels in the foliage region of the image generated by $|\hat{\rho}_c|$. We can see that for lower values of N, $\hat{r}$ performs better, whereas $|\hat{\rho}_c|$ requires higher values of N to produce better performance. Overall, $|\hat{\rho}_a|$ and the two-stage detector perform better than $\hat{r}$ and $|\hat{\rho}_c|$. Also, the images generated by $|\hat{\rho}_a|$ and the two-stage detector are almost similar, thus only $|\hat{\rho}_a|$ can be used to detect changes, as it provides a simpler alternative as compared to the two-stage detector.

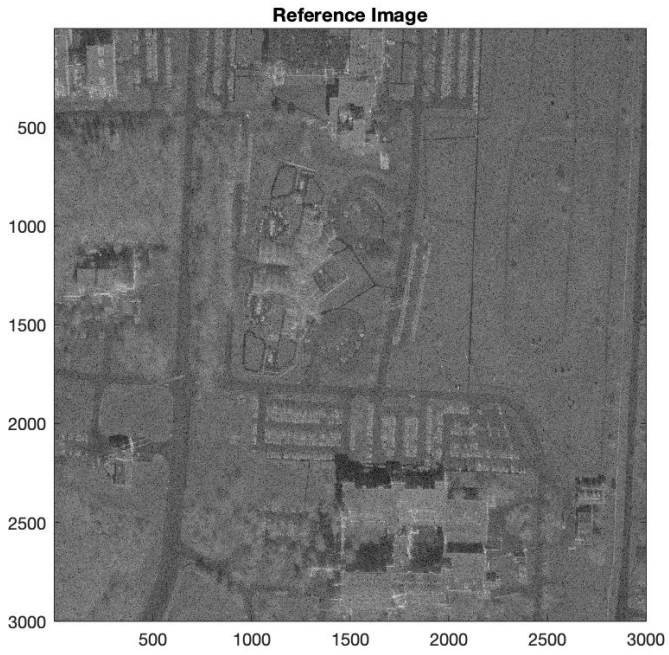


Figure 7.1: Reference SAR Image

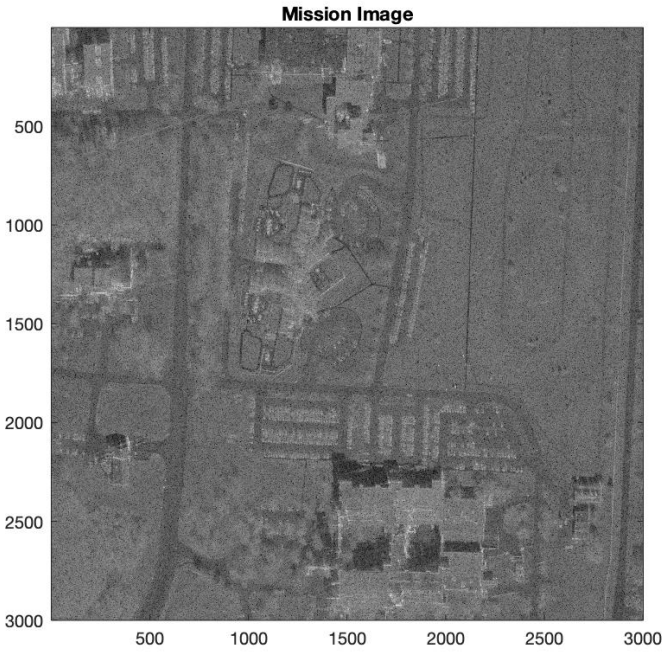


Figure 7.2: Mission SAR Image

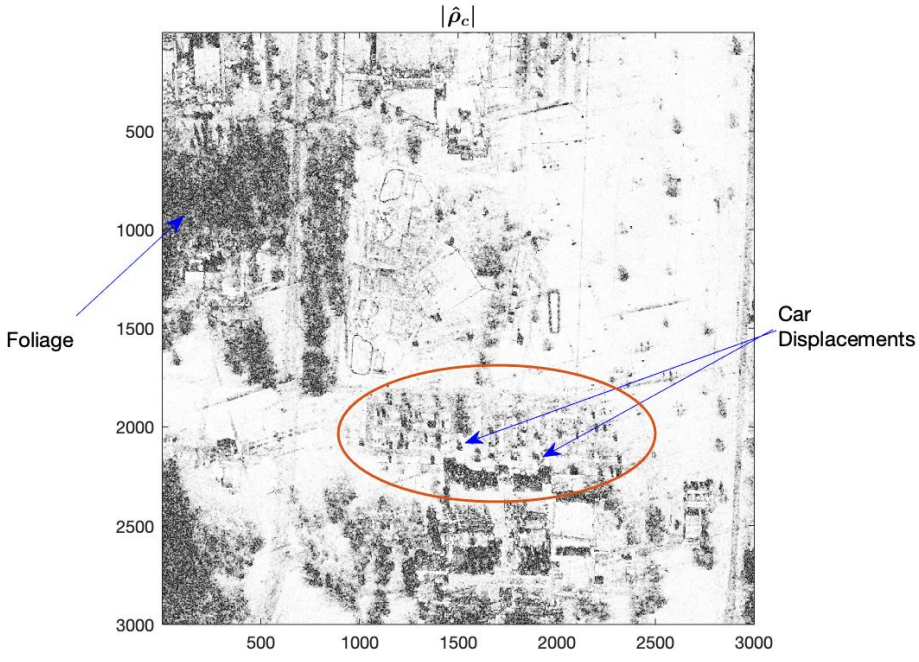


Figure 7.3: Gray Scale Change Image showing Raw Statistic $|\hat{\rho}_c|$

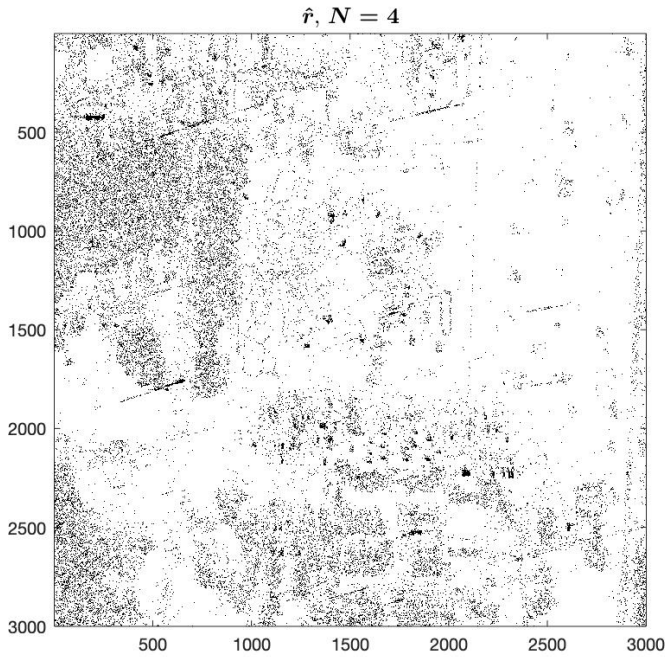


Figure 7.4: Change Image for $\hat{r}$, $N = 4$.

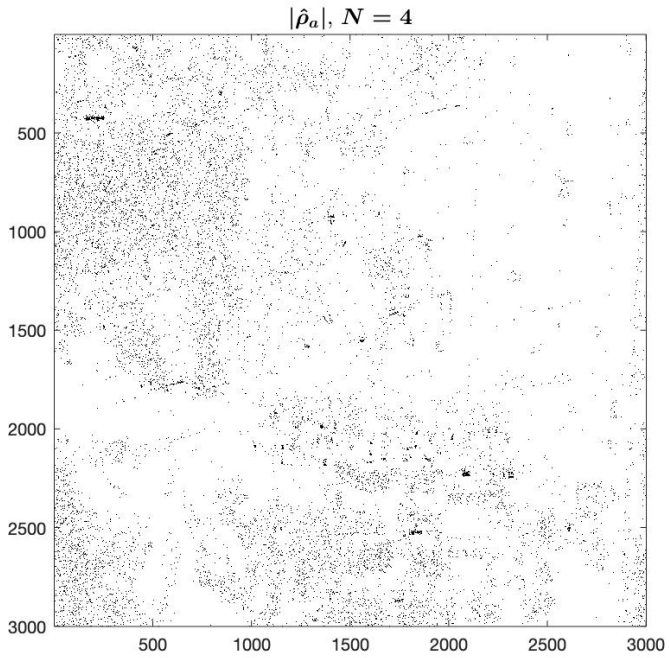


Figure 7.5: Change Image for $|\hat{\rho}_a|$, $N = 4$

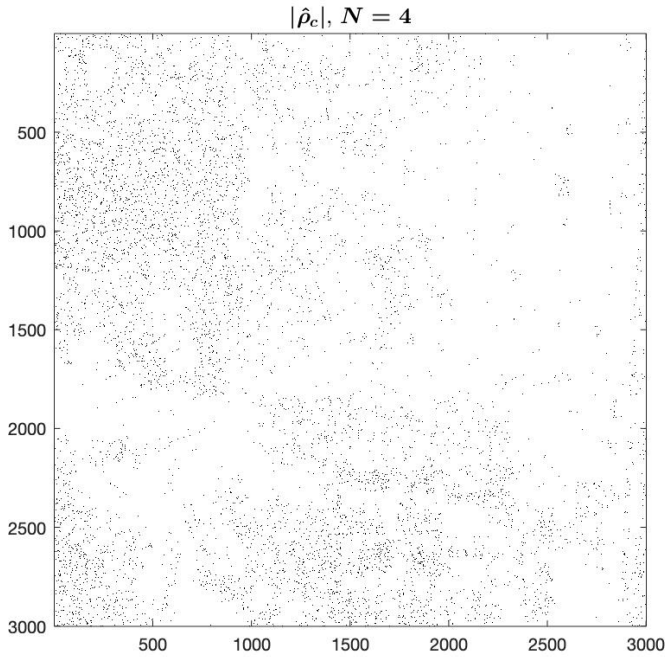


Figure 7.6: Change Image for $|\hat{\rho}_c|$, $N = 4$

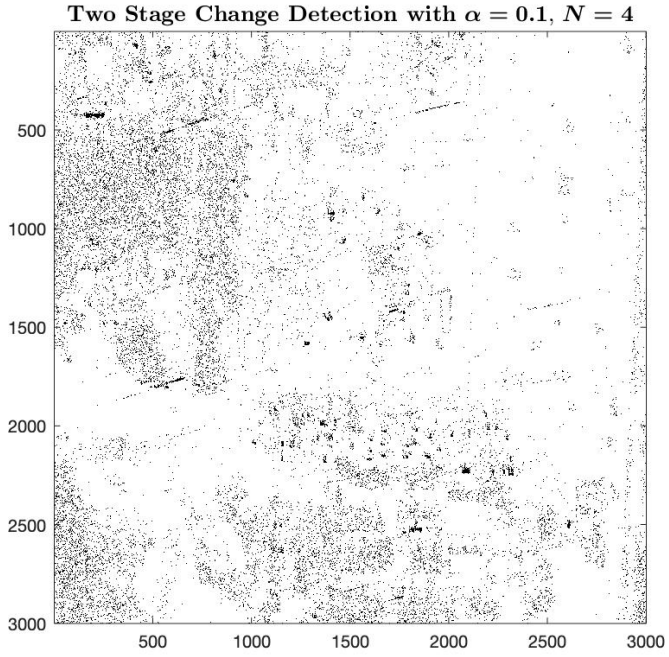


Figure 7.7: Change Image for Two-Stage Detector with $\alpha = 0.1$, $N = 4$.

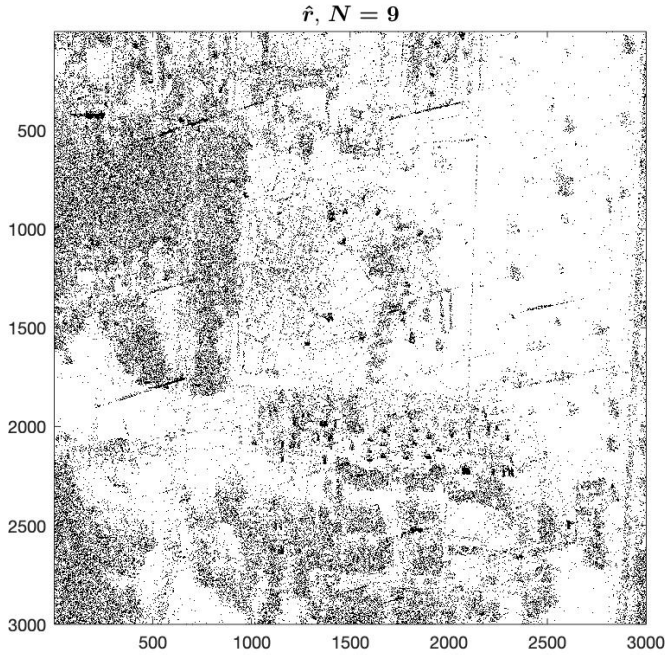


Figure 7.8: Change Image for $\hat{r}$, $N = 9$.

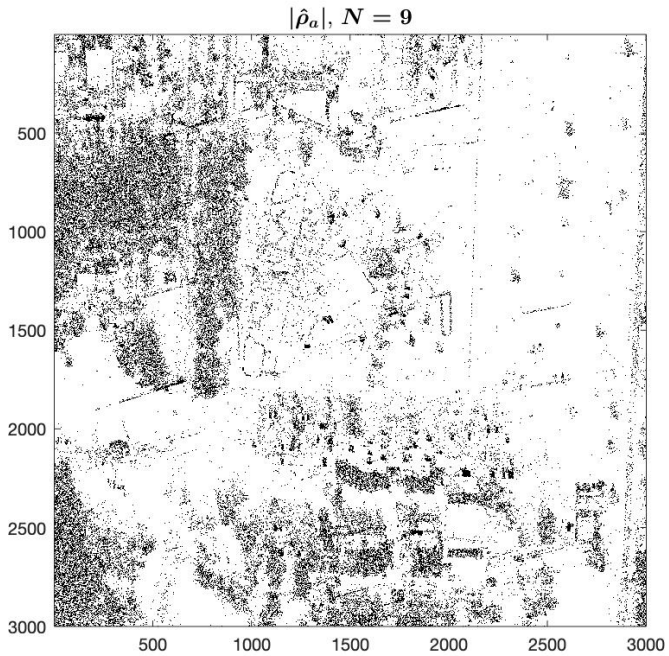


Figure 7.9: Change Image for $|\hat{\rho}_a|$, $N = 9$

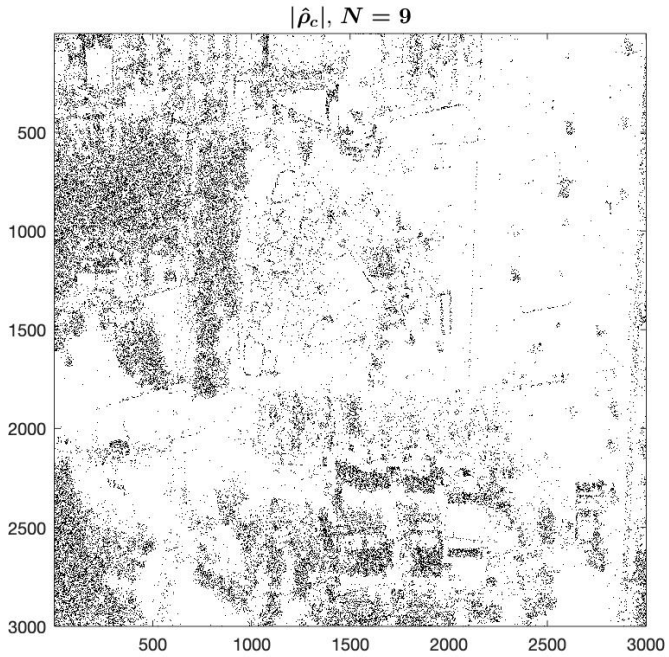


Figure 7.10: Change Image for $|\hat{\rho}_c|$, $N = 9$

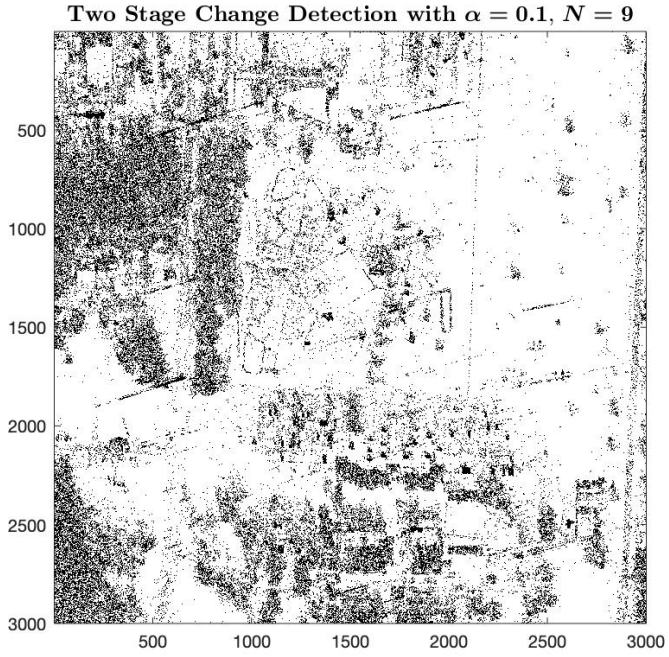


Figure 7.11: Change Image for Two-Stage Detector with $\alpha = 0.1$, $N = 9$

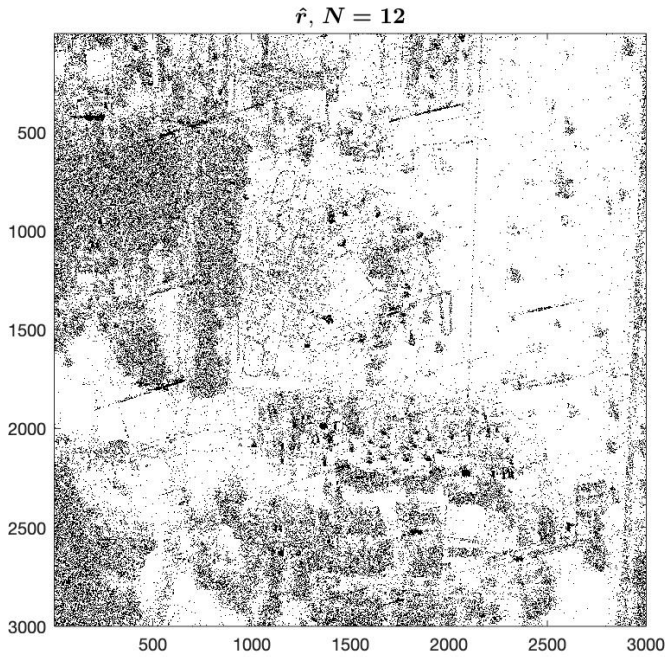


Figure 7.12: Change Image for $\hat{r}$, $N = 12$.

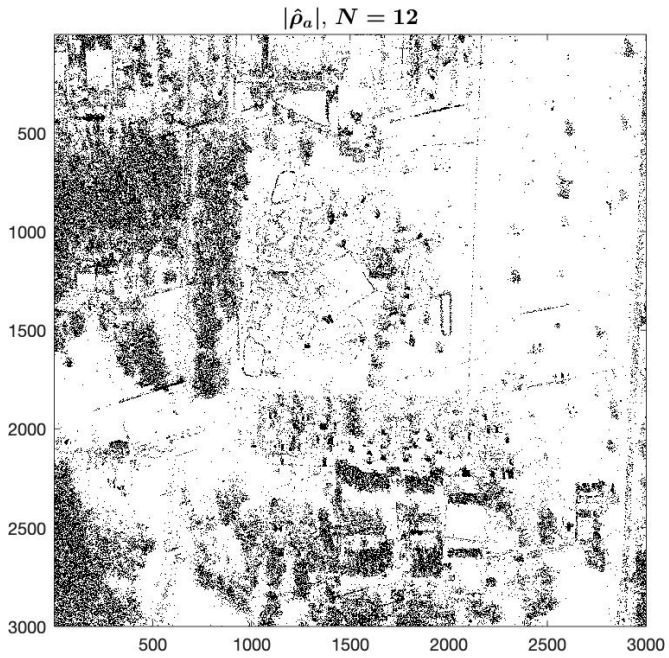


Figure 7.13: Change Image for $|\hat{\rho}_a|$, $N = 12$

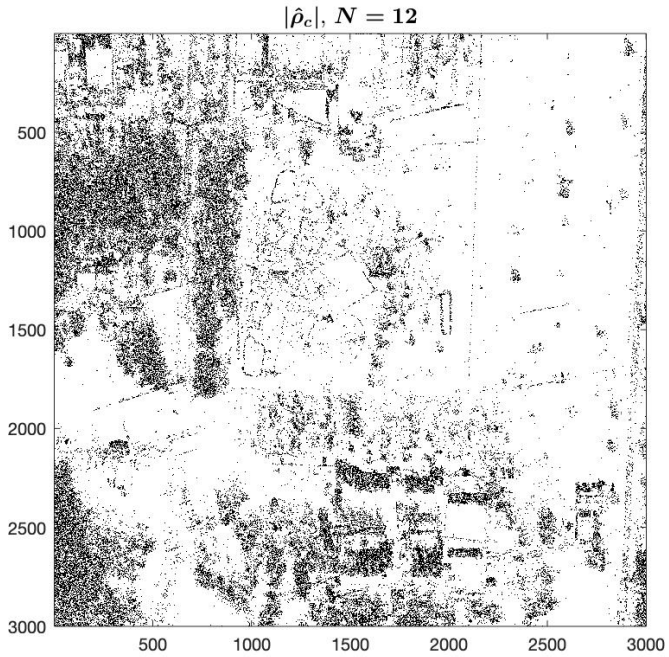


Figure 7.14: Change Image for $|\hat{\rho}_c|$, $N = 12$

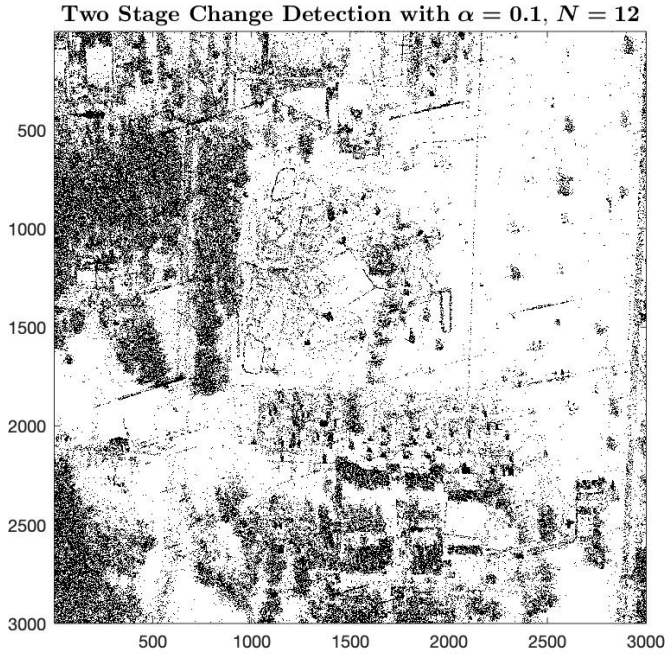


Figure 7.15: Change Image for Two-Stage Detector with $\alpha = 0.1$, $N = 12$

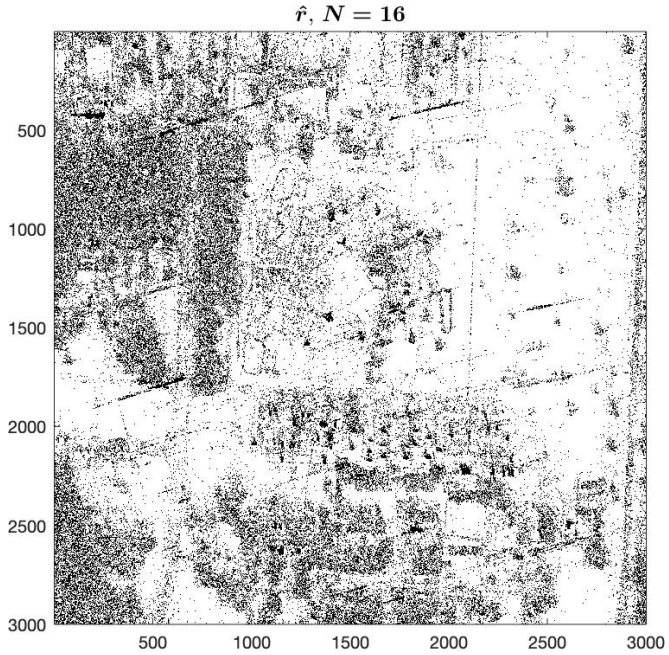


Figure 7.16: Change Image for $\hat{r}$, $N = 16$.

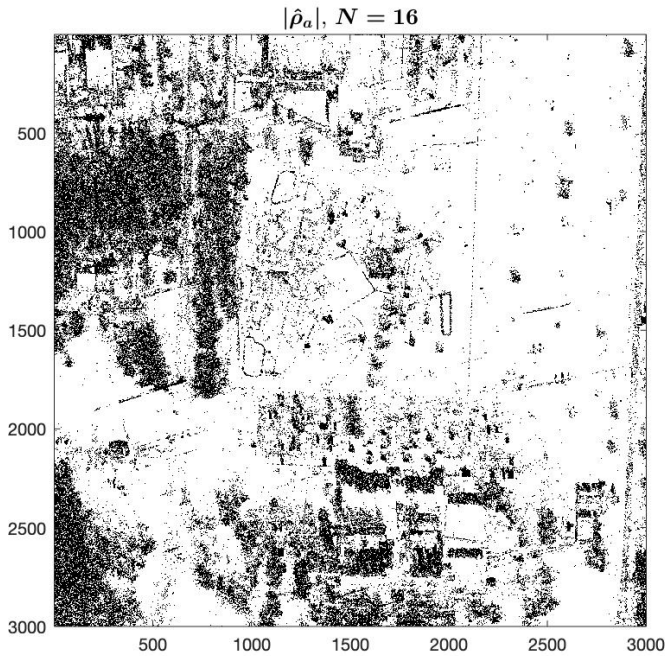


Figure 7.17: Change Image for $|\hat{\rho}_a|$, $N = 16$

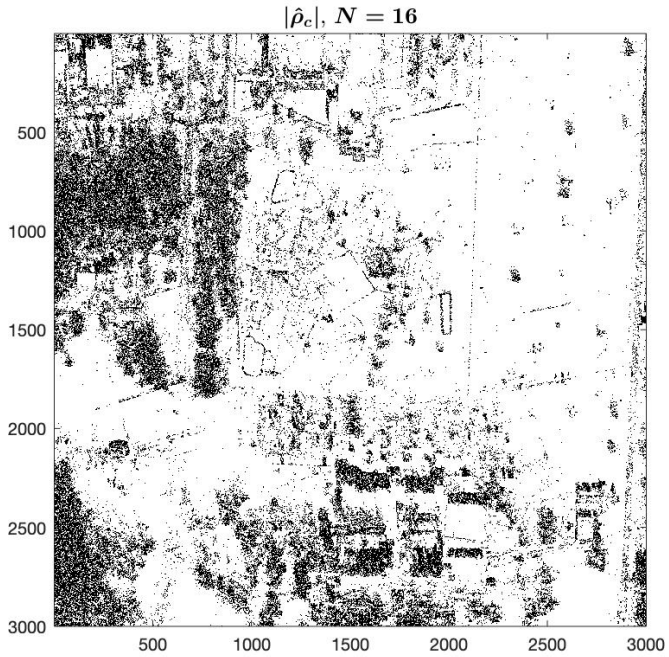


Figure 7.18: Change Image for $|\hat{\rho}_c|$, $N = 16$

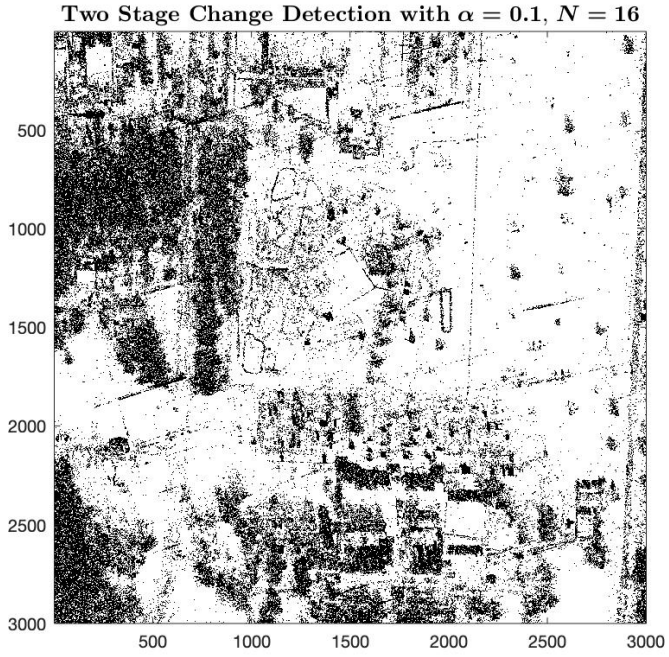


Figure 7.19: Change Image for Two-Stage Detector with $\alpha = 0.1$, $N = 16$

Chapter 8

CONCLUSION

The problem of detecting areas of change between SAR images of the same scene, captured at two different time instants is explored in this study. Particularly, the two-stage change detector proposed by [2] has been studied in detail. This detector is a two-dimensional detector, where the dimensions correspond to the sample variance ratio $\hat{R}$, and the Berger's coherence estimator $|\hat{\rho}_a|$. It has been shown that replacing the sample variance ratio test statistic by the symmetric detector $\hat{r}$ yields a simpler form for the two-stage detector. The probability density function and the cumulative distribution function for this symmetric ratio detector have been derived. The probability density function for the modified two-stage detector has also been derived. An approach for computing the thresholds with the help of this density function for a fixed probability of false alarm constraint is proposed. It is shown that in order to uniquely determine the thresholds, the introduction of another parameter α is required. The two-stage detector effectively reduces to the symmetric sample variance ratio detector when $\alpha = 1$, whereas when $\alpha \approx 0$, it reduces to the Berger's coherence estimator. α taking a value between 0 and 1 results in the two-stage detector behaving as a combination the two statistics $|\hat{\rho}_a|$ and $\hat{r}$.

The dependence of the detection performance on α is explored in detail, and it has been shown that the optimal value of α depends on the true value of R and $|\rho|$. However, considering a fixed value of $\alpha = 0.1$ gives a better detection performance as compared to the classical coherence estimator and the $\hat{r}$. Theoretical ROC curves of all the four detectors were compared with the optimal log-likelihood ratio detector, which demonstrated that the two-stage detector with $\alpha = 0.1$ gives a better

performance compared to the other three detectors. However, since the probability of detection achieved by Berger's estimate is only slightly less than that of the two-stage detector, Berger's estimate may be a better option since it only involves a single threshold. All of the four detectors were implemented on a SAR data set, and qualitatively better change images were generated by $|\hat{\rho}_a|$ and the two-stage detector as compared to $\hat{r}$ and $|\hat{\rho}_c|$.

REFERENCES

[1] Chris Oliver and Shaun Quegan. *Understanding synthetic aperture radar images*. SciTech Publishing, 2004.

[2] Miriam Cha, Rhonda D Phillips, Patrick J Wolfe, and Christ D Richmond. Two-stage change detection for synthetic aperture radar. *IEEE Transactions on Geoscience and Remote sensing*, 53(12):6547–6560, 2015.

[3] Mark Preiss, Douglas A Gray, and Nick JS Stacy. Detecting scene changes using synthetic aperture radar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8):2041–2054, 2006.

[4] Douglas G Corr and Alex Rodrigues. Coherent change detection of vehicle movements. In *IGARSS'98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings.(Cat. No. 98CH36174)*, volume 5, pages 2451–2453. IEEE, 1998.

[5] Toby Berger. On the correlation coefficient of a bivariate, equal variance, complex gaussian sample. *The Annals of Mathematical Statistics*, pages 2000–2003, 1972.

[6] Rex K Andrew and R Lynn Kirlin. An equality test for variances of two complex correlated gaussian processes. *IEEE transactions on signal processing*, 48(8):2452–2454, 2000.

[7] Ridha Touzi, Armand Lopes, and Pierre Bousquet. A statistical and geometrical edge detector for sar images. *IEEE Transactions on geoscience and remote sensing*, 26(6):764–773, 1988.

[8] Mark A Richards, Jim Scheer, William A Holm, and William L Melvin. *Principles of modern radar*. Citeseer, 2010.

[9] Giorgio Franceschetti and Riccardo Lanari. *Synthetic aperture radar processing*. CRC press, 1999.

[10] Nathaniel R Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of mathematical statistics*, 34(1):152–177, 1963.

[11] Steven M Kay. Fundamentals of statistical signal processing. detection theory, volume ii. *Printice Hall PTR*, pages 1545–5971, 1998.

[12] Tarek Medkour and Andrew T Walden. Attenuation estimation from correlated sequences. *IEEE transactions on signal processing*, 55(1):378–383, 2006.

[13] G Carter, C Knapp, and A Nuttall. Statistics of the estimate of the magnitute-coherence function. *IEEE transactions on audio and electroacoustics*, 21(4):388–389, 1973.

[14] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of integrals, series, and products*. Academic press, 2014.

[15] Christ D Richmond. Capon algorithm mean-squared error threshold snr prediction and probability of resolution. *IEEE Transactions on Signal Processing*, 53(8):2748–2764, 2005.

[16] Christ D Richmond. Performance of the adaptive sidelobe blanker detection algorithm in homogeneous environments. *IEEE Transactions on Signal Processing*, 48(5):1235–1247, 2000.

[17] Steven M Scarborough, LeRoy Gorham, Michael J Minardi, Uttam K Majumder, Matthew G Judge, Linda Moore, Leslie Novak, Steven Jaroszewksi, Laura Spoldi, and Alan Pieramico. A challenge problem for sar change detection and data compression. In *Algorithms for Synthetic Aperture Radar Imagery XVII*, volume 7699, page 76990U. International Society for Optics and Photonics, 2010.

APPENDIX A

DERIVATION OF PDF OF $\hat{r}$

This test statistic $\hat{R}$, which is an estimator of the ratio of the true variances , has probability density function(pdf) given by :

$$p_{\hat{R}}\left(x; R, |\rho|, N\right) = \frac{\Gamma(2N)}{\Gamma(N)^2} \frac{(1-|\rho|^2)^N (x+R) R^N x^{N-1}}{[(x+R)^2 - 4xR|\rho|^2]^{N+\frac{1}{2}}} \qquad \text{for } 0 < x < \infty. \tag{A.1}$$

where $R = \sigma_f^2/\sigma_g^2$ is the ratio of the true variances, ρ is the complex correlation coefficient between the two complex random radar returns and $\Gamma(.)$ denotes the gamma function. The symmetric detector $\hat{r}$ is given by :

$$\hat{r} = \begin{cases} \hat{R} & \text{when } \hat{R} \leq 1 \\ \hat{R}^{-1} & \text{when } \hat{R} > 1 \end{cases} \tag{A.2}$$

The approach to derive the pdf of $\hat{r}$ is similar to the one used in [7]. It can be seen from (A.2) that the support of the pdf of $\hat{r}$, $p_{\hat{r}}\left(x; R, |\rho|, N\right)$ is $0 \leq x \leq 1$. Thus, for $0 \leq a < b \leq 1$, we can write

$$\begin{aligned} \Pr(a < \hat{r} < b\,;\, R, |\rho|, N) = \Pr(a < \hat{R} < b\,,\, \hat{R} \leq 1\,;\, R, |\rho|, N) \\ + \Pr(a < \hat{R}^{-1} < b\,,\, \hat{R} > 1\,;\, R, |\rho|, N) \end{aligned} \tag{A.3}$$

Since a and b lie between 0 and 1, we can see that the event $\hat{R} \in [a, b]$ is a subset of the event $\hat{R} \leq 1$, and thus

$$\Pr(a < \hat{R} < b\,,\, \hat{R} \leq 1\,;\, R, |\rho|, N) = \Pr(a < \hat{R} < b\,;\, R, |\rho|, N) \tag{A.4}$$

Similarly, $\hat{R}^{-1} \in [a, b]$, which can also be written as $\hat{R} \in [1/b, 1/a]$, is a subset of the event $\hat{R} > 1$. Hence,

$$\Pr\left(a < \hat{R}^{-1} < b\,,\, \hat{R} > 1\,;\, R, |\rho|, N\right) = \Pr\left(\frac{1}{b} < \hat{R} < \frac{1}{a}\,;\, R, |\rho|, N\right). \tag{A.5}$$

Thus, using (A.4) and (A.5) in (A.3), we get

$$\begin{aligned} \Pr(a < \hat{r} < b\,;\, R, |\rho|, N) = \Pr(a < \hat{R} < b\,;\, R, |\rho|, N) \\ + \Pr\left(\frac{1}{b} < \hat{R} < \frac{1}{a}\,;\, R, |\rho|, N\right) \end{aligned} \tag{A.6}$$

We now consider $p_{\hat{R}}\left(x^{-1};R^{-1},|\rho|,N\right)$, which can be obtained from the expression for $p_{\hat{R}}\left(x;R,|\rho|,N\right)$ which is given in (A.1).

$$\begin{aligned}
p_{\hat{R}}\left(x^{-1};R^{-1},|\rho|,N\right) &= \frac{\Gamma(2N)}{\Gamma(N)^2}\frac{(1-|\rho|^2)^N(x^{-1}+R^{-1})R^{-N}x^{1-N}}{[(x^{-1}+R^{-1})^2-4x^{-1}R^{-1}|\rho|^2]^{N+\frac{1}{2}}} \\
&= \frac{\Gamma(2N)}{\Gamma(N)^2}\frac{(1-|\rho|^2)^N\left(\frac{x+R}{xR}\right)R^{-N}x^{1-N}}{\left[\left(\frac{x+R}{xR}\right)^2-4\frac{xR}{(xR)^2}|\rho|^2\right]^{N+\frac{1}{2}}} \\
&= \frac{\Gamma(2N)}{\Gamma(N)^2}\frac{(1-|\rho|^2)^N(x+R)R^{-N}x^{1-N}}{[(x+R)^2-4xR|\rho|^2]^{N+\frac{1}{2}}}.(xR)^{2N} \\
&= \frac{\Gamma(2N)}{\Gamma(N)^2}\frac{(1-|\rho|^2)^N(x+R)R^{N}x^{N-1}}{[(x+R)^2-4xR|\rho|^2]^{N+\frac{1}{2}}}.x^2 \\
p_{\hat{R}}\left(x^{-1};R^{-1},|\rho|,N\right) &= p_{\hat{R}}\left(x;R,|\rho|,N\right).x^2. \qquad \text{(A.7)}
\end{aligned}$$

We now go back to (A.6), and consider the second term in the right hand side of (A.6),

$$\Pr\left(\frac{1}{b}<\hat{R}<\frac{1}{a}\,;\,R,|\rho|,N\right)=\int_{1/b}^{1/a}p_{\hat{R}}\left(x;R,|\rho|,N\right)dx. \qquad \text{(A.8)}$$

Using the substitution $x=1/y$, we get $dx=-(1/y^2)dy$, and

$$\begin{aligned}
\Pr\left(\frac{1}{b}<\hat{R}<\frac{1}{a}\,;\,R,|\rho|,N\right) &= -\int_{b}^{a}p_{\hat{R}}\left(\frac{1}{y};R,|\rho|,N\right)\frac{dy}{y^2} \\
&= \int_{a}^{b}p_{\hat{R}}\left(\frac{1}{y};R,|\rho|,N\right)\frac{dy}{y^2}. \qquad \text{(A.9)}
\end{aligned}$$

Now, using the property given by (A.7), we get

$$\Pr\left(\frac{1}{b}<\hat{R}<\frac{1}{a}\,;\,R,|\rho|,N\right)=\int_{a}^{b}p_{\hat{R}}\left(y;\frac{1}{R},|\rho|,N\right)dy. \qquad \text{(A.10)}$$

Thus, substituting this result in (A.6), we get

$$\int_{a}^{b}p_{\hat{r}}\left(y;R,|\rho|,N\right)dy=\int_{a}^{b}\left[p_{\hat{R}}\left(y;R,|\rho|,N\right)+p_{\hat{R}}\left(y;\frac{1}{R},|\rho|,N\right)\right]dy \qquad \text{(A.11)}$$

which gives us the final result

$$p_{\hat{r}}\left(y;R,|\rho|,N\right)=p_{\hat{R}}\left(y;R,|\rho|,N\right)+p_{\hat{R}}\left(y;\frac{1}{R},|\rho|,N\right) \qquad \text{(A.12)}$$

which can be written as

$$p(\hat{r}; R, |\rho|, N) = \frac{\Gamma(2N)}{\Gamma(N)^2} \hat{r}^{N-1}(1-|\rho|^2)^N \cdot \left(\frac{R^N(\hat{r}+R)}{[(\hat{r}+R)^2 - 4\hat{r}R|\rho|^2]^{N+\frac{1}{2}}} + \frac{R^{-N}(\hat{r}+R^{-1})}{[(\hat{r}+R^{-1})^2 - 4\hat{r}R^{-1}|\rho|^2]^{N+\frac{1}{2}}} \right) \tag{A.13}$$

for $0 \leq \hat{r} \leq 1$ and zero elsewhere. Furthermore, using $a = 0$ and $b = \eta$ in (A.11), we get

$$\int_0^\eta p_{\hat{r}}\,(y; R, |\rho|, N)\, dy = \int_0^\eta p_{\hat{R}}\,(y; R, |\rho|, N)\, dy + \int_0^\eta p_{\hat{R}}\left(y; \frac{1}{R}, |\rho|, N\right) dy. \tag{A.14}$$

The integral on the left hand side gives the cdf of $\hat{r}$, and the two integrals on the right hand side are in terms of the cdf of $\hat{R}$, under true parameter values R and R^{-1}. Thus, the cdf of $\hat{r}$ is given by :

$$F_{\hat{r}}(\eta; R, |\rho|, N) = F_{\hat{R}}(\eta; R, |\rho|, N) + F_{\hat{R}}(\eta; R^{-1}, |\rho|, N) \tag{A.15}$$

where $F_{\hat{R}}(\eta; R, |\rho|, N)$ is the cdf of $\hat{R}$, given by (5.5).

APPENDIX B

DERIVATION OF JOINT PDF OF $|\hat{\rho}_a|$ AND $\hat{r}$

The joint pdf of $\hat{R}$ and $|\hat{\rho}_a|$ is given by :

$$\begin{aligned} p_{|\hat{\rho}_a|,\hat{R}}(x,y;R,|\rho|,N) =& \frac{(1-|\rho|^2)^N\Gamma(2N)}{\Gamma(N)\Gamma(N-1)}\frac{x}{2(y+1)^2} \\ & \cdot\left[\frac{y}{(y+1)^2}-\frac{x^2}{4}\right]^{N-2} \\ & \cdot\left[x|\rho|+\frac{y+R}{(y+1)\sqrt{R}}\right]^{-2N} \\ & \cdot{}_2F_1\left(\frac{1}{2},2N;1;\frac{2x|\rho|}{x|\rho|+\frac{y+R}{(y+1)\sqrt{R}}}\right) \end{aligned} \tag{B.1}$$

for $0 \le y < \infty$, $x \le [4y/(y+1)^2]^{1/2}$ and zero elsewhere. We begin by considering

$$\begin{aligned} P(a<\hat{r}<b\,,\ c<|\hat{\rho}_a|<d\,;\ R,|\rho|,N) = & \\ P(a<\hat{R}<b\,,\ c<|\hat{\rho}_a|<d\,,\ \hat{R}\le 1\,;\ R,|\rho|,N) & \\ +P(a<\hat{R}^{-1}<b\,,\ c<|\hat{\rho}_a|<d\,,\ \hat{R}>1\,;\ R,|\rho|,N) & \end{aligned} \tag{B.2}$$

where $0 \le a < b \le 1$ and $0 \le c < d \le 1$. Similar to (A.4) and (A.5), we can write the above equation as

$$\begin{aligned} P(a<\hat{r}<b\,,\ c<|\hat{\rho}_a|<d\,;\ R,|\rho|,N) = P(a<\hat{R}<b\,,\ c<|\hat{\rho}_a|<d\,;\ R,|\rho|,N) & \\ + P(\frac{1}{b}<\hat{R}<\frac{1}{a}\,,\ c<|\hat{\rho}_a|<d\,;\ R,|\rho|,N) & \end{aligned} \tag{B.3}$$

Considering $p_{|\hat{\rho}_a|,\hat{R}}(x,y^{-1};R^{-1},|\rho|,N)$,

$$\begin{aligned} p_{|\hat{\rho}_a|,\hat{R}}(x,y^{-1};R^{-1},|\rho|,N) =& \frac{(1-|\rho|^2)^N\Gamma(2N)}{\Gamma(N)\Gamma(N-1)}\frac{x}{2(y^{-1}+1)^2}\cdot\left[\frac{y^{-1}}{(y^{-1}+1)^2}-\frac{x^2}{4}\right]^{N-2} \\ & \cdot\left[x|\rho|+\frac{y^{-1}+R^{-1}}{(y^{-1}+1)\sqrt{R^{-1}}}\right]^{-2N} \\ & \cdot{}_2F_1\left(\frac{1}{2},2N;1;\frac{2x|\rho|}{x|\rho|+\frac{y^{-1}+R^{-1}}{(y^{-1}+1)\sqrt{R^{-1}}}}\right) \\ =& \frac{(1-|\rho|^2)^N\Gamma(2N)}{\Gamma(N)\Gamma(N-1)}\frac{xy^2}{2(y+1)^2}\cdot\left[\frac{y}{(y+1)^2}-\frac{x^2}{4}\right]^{N-2} \\ & \cdot\left[x|\rho|+\frac{y+R}{(y+1)\sqrt{R}}\right]^{-2N} \\ & \cdot{}_2F_1\left(\frac{1}{2},2N;1;\frac{2x|\rho|}{x|\rho|+\frac{y+R}{(y+1)\sqrt{R}}}\right) \end{aligned}$$

$$p_{|\hat{\rho}_a|,\hat{R}}(x, y^{-1}; R^{-1}, |\rho|, N) = p_{|\hat{\rho}_a|,\hat{R}}(x, y; R, |\rho|, N).y^2 \tag{B.4}$$

which is similar to (A.7). Considering the second term in the right hand side of (B.3),

$$\begin{aligned} P(\frac{1}{b} < \hat{R} < \frac{1}{a}\ ,\ c < |\hat{\rho}_a| < d\ ;\ R, |\rho|, N) &= \int_c^d \int_{1/b}^{1/a} p_{|\hat{\rho}_a|,\hat{R}}(x, y; R, |\rho|, N) dy dx \\ &= \int_c^d \int_a^b p_{|\hat{\rho}_a|,\hat{R}}(x, \frac{1}{z}; R, |\rho|, N) \frac{dz dx}{z^2} \\ &= \int_c^d \int_a^b p_{|\hat{\rho}_a|,\hat{R}}(x, z; \frac{1}{R}, |\rho|, N) dz dx \end{aligned} \tag{B.5}$$

where the substitution $y = 1/z$ and (B.4) have been used. We can now write (B.3) as :

$$\begin{aligned} \int_c^d \int_a^b p_{|\hat{\rho}_a|,\hat{r}}(x, z; R, |\rho|, N) dz dx &= \int_c^d \int_a^b p_{|\hat{\rho}_a|,\hat{R}}(x, z; R, |\rho|, N) dz dx \\ &+ \int_c^d \int_a^b p_{|\hat{\rho}_a|,\hat{R}}(x, z; \frac{1}{R}, |\rho|, N) dz dx \end{aligned} \tag{B.6}$$

and thus we get the final result for the expression for the joint pdf of $\hat{r}$ and $|\hat{\rho}_a|$ as

$$p_{|\hat{\rho}_a|,\hat{r}}(x, z; R, |\rho|, N) = p_{|\hat{\rho}_a|,\hat{R}}(x, z; R, |\rho|, N) + p_{|\hat{\rho}_a|,\hat{R}}(x, z; \frac{1}{R}, |\rho|, N) \tag{B.7}$$