5-1-2015

Testing for Two Signals with Unknown Locations in Functional Magnetic Resonance Images Using Gaussian Random Field: A Monte Carlo Simulation Study

Pei-Chin Lu

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

TESTING FOR TWO SIGNALS WITH UNKNOWN LOCATIONS IN FUNCTIONAL MAGNETIC RESONANCE IMAGES USING GAUSSIAN RANDOM FIELD: A MONTE CARLO SIMULATION STUDY

A Dissertation Submitted in Partial Fulfillment
of the Requirement for the Degree of
Doctoral of Philosophy

Pei-Chin Lu

College of Educational and Behavioral Sciences
Department of Applied Statistics and Research Methods

May 2015

This dissertation by: Pei-Chin Lu
Entitled: *Testing for Two Signals with Unknown Locations in Functional Magnetic Resonance Images Using Gaussian Random Field: A Monte Carlo Simulation Study*

has been approved as meeting the requirement for the Degree of Doctoral of Philosophy in College of Education and Bahavioral Sciences in Department of Applied Statistics and Research Methods

Accepted by the Doctoral Committee

Khalil Shafie, Ph.D., Research Advisor

Jay Schaffer, Ph.D., Committee Member

Trent Lalonde, Ph.D., Committee Member

Robert Heiny, Ph.D., Faculty Representative

Date of Dissertation Defense _____

Accepted by the Graduate School

Linda L. Black, Ed.D., LPC
Dean of the Graduate School and International Admissions

ABSTRACT

Lu, Pei-Chin. *Testing for Two Signals with Unknown Locations in Functional Magnetic Resonance Images Using Gaussian Random Field: A Monte Carlo Simulation Study.* Published Doctor of Philosophy dissertation, University of Northern Colorado, 2015.

Gaussian random field theory has been used extensively for correcting the multiple comparisons problem in neuroimaging over the past few decades. Traditionally the global maximum X_{max} of the field is used as the test statistic for thresholding, and it was proved to be the likelihood ratio test statistic when testing one signal in a Gaussian scale space random field. Nonetheless, it is not uncommon to test for more than one signal in practice. Hence, the primary purpose of the current study was to propose a new likelihood ratio test statistic Y_{max} for testing two signals simultaneously in fMRI images based on Gaussian random field theory. Monte Carlo simulation was used to approximate the probability of Y_{max} through the empirical distribution in two-dimensional images and its power was also assessed under different conditions, varying the levels of distance, amplitude, and scale of the signals. This study also sought to explore the result in scale space where the width of smoothing kernel was added as an extra dimension. Critical values were successfully obtained for Y_{max} using simulation. In scale space, the thresholds were more stringent than the ones with fixed kernel width, but it also revealed that the power of scale space Y_{max} was higher. In both scale space and fixed smoothing width, distance, amplitude and scale of the signals all had effect on the power of

Y_{max} to some extent. Nonetheless, Y_{max} did not seem to surpass the other conventional test statistics in terms of power. The reason could be the limited conditions being simulated in the present study. Further investigation is required to provide more information about the behavior of Y_{max}.

ACKNOWLEDGEMENTS

This journey of dissertation began from a casual conversation back in summer 2013. At that time, Dr. Khalil Shafie was trying to talk me into doing some fancy statistical theory of which I have never heard, and he was making jokes that this is like an "arranged marriage" that I will have to spend quite some time in order to get familiar with this new topic. Indeed it took me almost two years to finish the work (not to mention how many times I was pulling my hair out), and to date I still cannot say I fully grasp the idea of Gaussian random field theory. But somehow in a hindsight this "arranged marriage" was not really a terrible idea as it seemed after all. I am sincerely thankful to Dr. Shafie for being my research advisor. Without his guidance and support, this dissertation would never have been made into reality. And I must thank him for his amiable insistence on making everything difficult for me, and his endless patience for my periodic tears of frustration throughout these years. It is my honor to be your student, Dr. Shafie.

To my dissertation committee, Dr. Jay Schaffer, Dr. Trent Lalonde, and Dr. Robert Heiny, I would like to thank them for their helpful feedback. Dr. Susan Hutchinson has been a wonderful teacher and mentor, always inspiring me to become a better researcher. I also wish to thank Dr. Steven Pulos whose classes I enjoy all the time, for his warm-hearted encouragement and witty sense of humor. Special thanks goes to Dr. Bob Pearson, who was always there to listen and gave me

valuable advice. I am especially indebted to him for teaching me so many programming tricks.

There are many friends at ASRM and UNC that I would like to express my gratitude: Jamil, Samy, JD, David Reavill, Wipanee (my lovely roommate), Lin, Katie, Kirk, Steve Hoff, Aman, and many other people whom I did not have the space to mention here. Thank you for being there with me and always putting up with my crankiness. I am lucky to have your friendship. I also wish to thank Keyleigh Gurney for her help with almost everything. You are the best, Keyleigh.

Lastly, I would like to thank my dearest family. Even though they were not with me all the time during this journey, it was their love and support that keep me going through every single difficult situation. Dad and Mom, thank you for kicking me out of the house seven years ago so I was able to keep the family tradition and became the 3rd Ph.D. in the Lu family. Bekki and Scott, thank you always for your warm hospitality. To my two fur-nephews, Meichan and Heimdall, I sincerely thank you both for being fluffy.

TABLE OF CONTENTS

CHAPTER		Page
I	INTRODUCTION	1
	Statistical Analysis of Functional Magnetic Resonance Imaging Data	2
	Problem Statement	5
	Purpose of the Study	6
	Research Questions	6
	Delimitations of the Study	6
	Definition of Terms	7
II	REVIEW OF LITERATURE	9
	Functional Magnetic Resonance Imaging	9
	Multiple Comparisons Problem in Functional Magnetic Resonance Imaging	15
	Random Field Theory	17
	From One Signal to Multiple Signals	26
	Summary	27
III	METHODOLOGY	29
	Likelihood Ratio Test Statistic for Testing Two Signals When the Kernel Width Is Fixed	31

Likelihood Ratio Test Statistic for Testing Two Signals in Scale Space 41

IV RESULTS 44

Results for Fixed Smoothing Kernel 44

Results for the Scale Space 85

V CONCLUSIONS 96

Findings 96

Limitations and Suggestions for Future Research 100

Closing Remarks 102

REFERENCES 104

APPENDIX A R CODE FOR MONTE CARLO SIMULATION 108

Code for Simulating Critical Values with Fixed Kernel 109

Code for Generating Power with Fixed Kernel 111

Code for Simulating Critical Values in Scale Space 114

Code for Generating Power in Scale Space 117

LIST OF TABLES

1 Schemes of the Parameters 35

2 Critical Values When σ Is Fixed 45

3 Power Table When $\sigma = 0.02$ at $\alpha = .05$ 47

4 Power Table When $\sigma = 0.02$ at $\alpha = .01$ 48

5 Power Table When $\sigma = 0.0258$ at $\alpha = .05$ 49

6 Power Table When $\sigma = 0.0258$ at $\alpha = .01$ 50

7 Power Table When $\sigma = 0.0334$ at $\alpha = .05$ 51

8 Power Table When $\sigma = 0.0334$ at $\alpha = .01$ 52

9 Power Table When $\sigma = 0.0431$ at $\alpha = .05$ 53

10 Power Table When $\sigma = 0.0431$ at $\alpha = .01$ 54

11 Power Table When $\sigma = 0.0557$ at $\alpha = .05$ 55

12 Power Table When $\sigma = 0.0557$ at $\alpha = .01$ 56

13 Power Table When $\sigma = 0.0719$ at $\alpha = .05$ 57

14 Power Table When $\sigma = 0.0719$ at $\alpha = .01$ 58

15 Power Table When $\sigma = 0.0928$ at $\alpha = .05$ 59

16 Power Table When $\sigma = 0.0928$ at $\alpha = .01$ 60

17 Power Table When $\sigma = 0.1199$ at $\alpha = .05$ 61

18 Power Table When $\sigma = 0.1199$ at $\alpha = .01$ 62

19 Power Table When $\sigma = 0.1549$ at $\alpha = .05$ 63

20 Power Table When $\sigma = 0.1549$ at $\alpha = .01$ 64

21 Power Table When $\sigma = 0.2$ at $\alpha = .05$ 65

22 Power Table When $\sigma = 0.2$ at $\alpha = .01$ 66

23 Critical Values in the Scale Space 85

24 Power Table in Scale Space at $\alpha = .05$ 87

25 Power Table in Scale Space at $\alpha = .01$ 88

26 Comparison of the Power of Y_{max}'s at $\alpha = .05$ 89

27 Comparison of the Power of Y_{max}'s at $\alpha = .01$ 90

LIST OF FIGURES

1 A Statistical Parametric Map 3

2 Relation between standard deviation and FWHM 13

3 Effect of the width of smoothing kernel 14

4 An example of height thresholding in two-dimensional images 21

5 Three levels of distance between signals when $\xi_1 = \xi_2 = 2$ 37

6 The empirical distribution of $Y_{max}95$ for different σ values at $\alpha = .05$ 46

7 Power vs. distance 68

7 Simulated power vs. distance between the two signals 69

8 Simulated power vs. amplitude of the two signals 71

8 Simulated power vs. amplitude of the two signals 72

9 Simulated power vs. σ 73

9 Simulated power vs. σ 74

10 Simulated power of Y_{max} vs. norm for selected σ values at $\alpha = .05$ . . 80

10 Simulated power of Y_{max} vs. norm for selected σ values at $\alpha = .05$ . . 81

11 Power vs. norm grouped by scale 82

12 Comparison between Y_{max} and X^2_{max} on power against norm at $\alpha = .05$ 83

12 Comparison between Y_{max} and X^2_{max} on power against norm at $\alpha = .05$ 84

13 The empirical distribution of scale space $Y_{max}95$ 85

14 Simulated power vs. distance in scale space 92

15 Simulated power vs. amplitude in scale space 93

16 Simulated power vs. norm in scale space at $\alpha = .05$ 94

CHAPTER I

INTRODUCTION

Neuroimaging, or "brain mapping", has flourished for the last two decades thanks to modern technology. Functional neuroimaging is especially gaining increasing interest among neuroscientists given that it is a profound technique to map out the location of neural function within a living brain, and enable us to "see into our minds" (Schweitzer, 2010). Functional neuroimaging, compared to structural neuroimaging which only concentrates on the structure of the brain, facilitates understanding of how neural activity relates to human thinking process, emotional function, and reaction in response to stimulus (Ashby, 2011; Moran & Zaki, 2013; Schweitzer, 2010). There are several techniques used in functional neuroimaging—among them, functional magnetic resonance imaging (fMRI), a neuroimaging procedure using MRI technology, has become the dominant tool since the beginning of 1990s given that it carries fewer health risks as well as the wide accessibility of the equipment. It also provides much better spatial resolution than other techniques (Banich, 2010; Huettel, Song, & McCarthy, 2008).

The ultimate goal of fMRI data analysis is to detect the correlation between the stimulus and brain activation. To this end, fMRI measures a signal called the blood-oxygen level dependent (BOLD) response, and generates brain images showing the local changes in neuronal activity. Typical fMRI data refer to an image

of a brain, which are comprised of a set of voxels (the three-dimensional counterpart of a two-dimensional pixel). These voxels are arranged in a matrix, corresponding to spatial locations with a living brain (Poldrack, Mumford, & Nichols, 2011; Schweitzer, 2010).

Statistical Analysis of Functional Magnetic Resonance Imaging Data

Because fMRI study relies on detecting changes in the BOLD response, the experiment is usually designed to compare two conditions: a baseline condition and a treatment condition. During the experiment, the participants are instructed to perform certain cognitive tasks or receive a series of stimuli (Ashby, 2011; Maydeu-Olivares & Brown, 2013). After acquiring the data, statistical analyses, mostly based on the general linear model (GLM), are applied to the fMRI data in order to detect brain activation. As mentioned earlier, typical fMRI data consist of a bunch of voxels where each voxel represents a data point. As a result, a t-test, for instance, is applied to each voxel. Finally, the regions which show significant differences in brain activation are visualized by adding colors on a brain image, with the color gradient representing the value of the test statistic (e.g., z-scores or t-values). Such images are generally referred as statistical parametric maps (Figure 1)

It is worth noting that a standard MRI image generally has more than 130,000 voxels. Consequently, multiple comparisons problem arises as a result of numerous statistical tests being done in each image when making statistical inferences (Ashby, 2011; Bennett, Wolford, & Miller, 2009; Brett, Penny, & Kiebel,

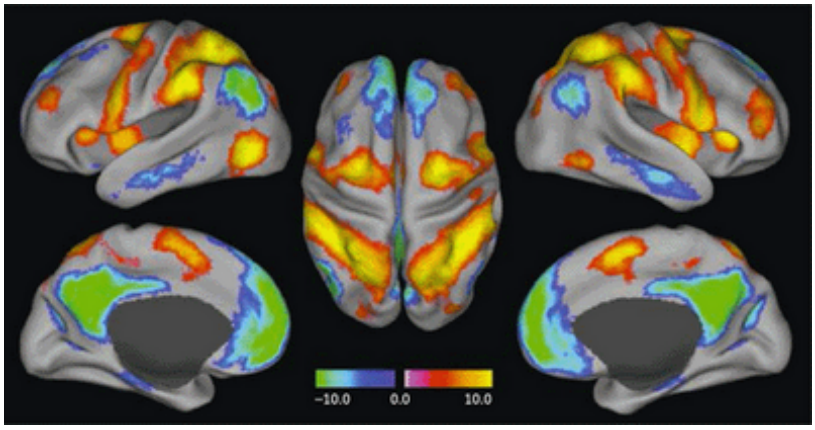


Figure 1. A Statistical Parametric Map

Source: `http://neurocritic.blogspot.com/2007_01_01_archive.html`

2003). With a Type I error rate of .05, we would expect 6,500 false positives if none of these voxels were stimulus-activated. A famous paper published by Bennett, Baird, Miller, and Wolford (2010) brought up the argument that the multiple comparisons problem is often ignored by researchers and a vast majority of fMRI studies did not utilize proper correction. They conducted a regular fMRI experiment in which they put the subject on an fMRI scanner and showed the subject a series of photographs of people engaging in social situations. In the end, several voxels in the subject's brain were found to be significant with an uncorrected $p < .001$ (Bennett, Baird, Miller, & Wolford, 2010, p.4). The catch in this experiment, despite everything, was that the subject is an Atlantic salmon—and it is dead. The experiment, which won the 2012 Ig Nobel Prize, may be absurd but the message behind it is clear: Correction for the multiple comparisons in fMRI study is necessary.

The multiple comparisons problem is not an unfamiliar issue in statistics. The most commonly-seen method to tackle multiple comparisons is the Bonferroni correction. However, fMRI data are large. An image of 130,000 voxels means there are 130,000 tests. Thus, the adjusted threshold based on the Bonferroni correction would be $.05/130000 = .000000385$, which corresponds to a z-score of 4.94. Obviously the correction is too conservative. Further, the Bonferroni correction assumes independent tests, but most fMRI data have some degree of spatial autocorrelation, meaning the voxels are not independent of each other. The autocorrelation usually comes from the intrinsic dependency among neighboring voxels in the image as well as the spatial smoothing applied to the image during preprocessing stage (Brett et al., 2003; Poldrack et al., 2011). Several methods other than the Bonferroni correction were proposed to solve the multiple comparison issues in fMRI study, and Gaussian random field theory is one of them (Bennett et al., 2009; Marchini & Presanis, 2004).

Random field theory is a sophisticated mathematical work where its fundamental work was developed by Adler in 1981, and later the theory was applied extensively by Keith Worsley to functional neuroimaging to remedy the multiple comparisons problem (Worsley, 1994; Worsley, Evans, Marrett, & Neelin, 1992). In fMRI, random field is used to find the global maximum as a test statistic controlling for the family-wise error rate. It takes account of the spatial correlation through the use of a topological measure so-called the Euler characteristic (Worsley, 1996). At high thresholds, the Euler characteristic is either zero or one, and the corresponding

expected value of the Euler characteristic is approximately equal to the probability of the global maximum, that is, the p-value (Worsley, 1996). Thus, this p-value could serve as a new threshold for statistical inferences.

Most of Worsley's work is based on the Gaussian random field, which is a random field whose finite dimensional distributions are all multivariate Gaussian (normal) distributions. Worsley and his colleagues also extended the work to other fields, such as χ^2, t, and F, to name a few (Cao & Worsley, 2001; Worsley, 1994); Gaussian scale- as well as rotation-space random fields were also proposed (Shafie, Sigal, Siegmund, & Worsley, 2003; Siegmund & Worsley, 1995). For all the fields mentioned above, explicit calculations of the probability of the global maximum using the expected value of the Euler characteristic were obtained.

Problem Statement

Siegmund and Worsley (1995) proposed Gaussian scale-space random field for testing one signal with unknown location and scale. When testing for a positive-valued signal ($H_a : \xi > 0$), they showed the global maximum X_{max} of a Gaussian random field is equivalent to the likelihood ratio test statistic in a $N + 1$ dimensional "scale space", with N dimensions for location and one dimension for the width of a smoothing kernel. Further, they were able to approximate the probability of X_{max} through the use of the expected value of the Euler characteristic and the volume of tubes method. Shafie (2014) continued Siegmund and Worsley's (1995) work and proved that when testing one signal with a two tailed hypothesis ($H_a : \xi \neq 0$), the likelihood ratio test statistics is X^2_{max} instead.

In applied fMRI study, however, it is not uncommon to test more than one signal simultaneously (Huettel et al., 2008), and in this case the likelihood ratio test statistic when testing two signals simultaneously in an image is neither X_{max} nor X^2_{max} anymore. Shafie (2014) proved a new test statistic, denoted as Y_{max}, is the likelihood ratio test statistic when testing two signals both in scale space and with fixed kernel width. However, not only the underlying distribution of Y_{max} is unknown, but also there is no closed form solution to approximate the probability of Y_{max}. Thus, there is little we know about Y_{max} regarding its sensitivity for signal detection.

Purpose of the Study

So far there is no published research on how to obtain p-values of the likelihood ratio test statistic using Gaussian random field theory when testing two signals. Therefore, the current study was intended to extend Siegmund and Worsley's (1995) and Shafie's (2014) work specifically for the two-signal case. Monte Carlo method was used to simulate the empirical distribution of Y_{max} in two-dimensional images to obtain the associated p-values. The power of Y_{max} was also examined under different conditions, such as distance, amplitude, and scale of the signals. To be more meaningful, the author was interested in knowing if Y_{max} outperforms X^2_{max} in power when testing two signals.

Research Questions

Q1 How to obtain p-values of the likelihood ratio test statistic Y_{max} for testing two signals simultaneously in two-dimensional images using Monte Carlo simulation?

Q2 To what extent does the distance between the locations of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q3 To what extent does the amplitude of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q4 To what extent does the scale of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q5 Does Y_{max} outperform X^2_{max} in power when testing two signals simultaneously under the above conditions?

Delimitations of the Study

A few limitations of the current study should be addressed. First, we assumed all the simulated error fields are Gaussian random fields (Adler, 1981). Second, we only investigated the situation with two signals; thus, there is no guarantee the results of the current study will be valid for other conditions. Third, the simulation results were only limited to two-dimensional images. Last but not least, the results of this study were based on Monte Carlo simulation method. As a result the data might not necessarily represent the diversity that can be found in real data.

Definition of Terms

Blood oxygen level dependent (BOLD). A ratio of oxygenated hemoglobin and deoxygenated hemoglobin.

Euler characteristic. A topological measure used to described a set.

Excursion set. A set of points where a field exceeds a fixed threshold.

Full width at half maximum (FWHM). The width of the smoothing kernel at the point where it is at half of its maximum.

Functional magnetic resonance imaging (fMRI). A functional neuroimaging procedure using MRI technology to detect brain activity.

Gaussian random field. A special class of random fields whose finite dimensional distributions are all multivariate Gaissian (normal) distributions.

Pixel. The basic element used in two-dimensional image.

Random fields. One type of random functions where a set of random variables defined over an N-dimensional Euclidean space.

Scale space. A range of smoothing widths is used to create an extra scale dimension to the data.

Signal coactivation. Two or more distinct brain regions are simultaneously activated during an experimental task.

Smoothing kernel. A function used to filter images.

Spatial Smoothing. To filter the image by replacing the values on each voxel with a weighted average of neighboring voxels. It is also called filtering or blurring.

Statistical parametric map (SPM). A graphical representation to display the statistical result of the brain activation where color gradient representing the values of the test statistic.

Voxel. The smallest unit used to form a three-dimensional image.

Width of the smoothing kernel. The parameter used to determine the amount of spatial smoothing applied to the image. In practice it is measured in FWHM but in the current study it was controlled by σ.

CHAPTER II

REVIEW OF LITERATURE

This chapter begins with an introduction of functional magnetic resonance imaging (fMRI), and follows with a discussion of the statistical challenges encountered in fMRI study to provide readers with background information. The next section describes random field theory as well as Gaussian random fields. The author then focuses on the Gaussian scale space random field developed by Siegmund and Worsley (1995) to lay down the framework for the methods proposed in Chapter III. Following the section of the Gaussian scale space random field, functional coactivation is briefly discussed. The chapter concludes with a summary of the key points illustrated by the review.

Functional Magnetic Resonance Imaging

How the human brain works has always been a mystery for scientists. The history of human brain study could date back to the time of Ancient Egypt. The word "brain" first appeared (in any language) in the famous Edwin Smith Papyrus nearly 3600 years ago. However, by the 20^{th} century neuroscience did not progress much through all the years, and what scientists had found is a mere glimpse of the human brain. To date, neuroscience is still barely-explored compared to other disciplines.

The modern technology, however, has accelerated the development of neuroscience, or more specifically, neuroimaging. Researchers, nowadays, can easily examine the anatomy of the brain (e.g., shape, size, volume) by using X-ray or computer assisted tomography (CAT), which is considered the domain of structural neuroimaging. Although structural neuroimaging is useful in the diagnosis of brain injury and corresponding neurological disorders, it cannot reveal "how the brain thinks or feels".

Functional neuroimaging, on the other hand, can overcome the limitations of structural neuroimaging by identifying and characterizing the brain regions required for given mental processes, which enables neuroscientists to discover more about how brain works related to emotions, thinking, or performing tasks in response to certain stimuli (Huettel et al., 2008; Schweitzer, 2010). There are multiple techniques for functional neuroimaging, including computed tomography (CT), positron emission tomography (PET), magnetoencephalography (MEG), electroenphalography (EEG), and fMRI, to name a few. Until mid-1990, PET was the most commonly seen neuroimaging technique and had shown to be useful to localize mental functions in the brain. For PET, it requires the injection of radioactive tracers into the bloodstream and measures the changes of oxygen consumption or glucose directly in the brain. Nonetheless, the use of PET is limited due to multiple reasons, including the safety concerns caused by the radioactive injections, the high costs of generating radioactive isotopes, and the slow data acquisition (Huettel et al., 2008; Poldrack et al., 2011).

Functional MRI has many advantages over PET, and nowadays it is the most popular neuroimaging technique. First, because fMRI is non-invasive, requiring no radioactive tracers, multiple scans can be run on subjects in a short time. Second, fMRI provides excellent spatial resolution compared to the other neuroimaging techniques such as EEG and MEG. Third, almost all the hospitals and medical centers have multiple MRI machines, making fMRI readily available (Banich, 2010; Huettel et al., 2008).

Functional MRI does not directly map out the neural activity of the human brain but the physiological changes that are related to neural activity (Ashby, 2011). The majority of the fMRI studies use the blood-oxygen level dependent (BOLD) signal, proposed by Ogawa, Lee, Kay and Tank (1990). The BOLD signal is a ratio of oxygenated hemoglobin and deoxygenated hemoglobin where the two have very different magnetic properties—oxygenated hemoglobin is diamagnetic and deoxygenated blood is paramagnetic (Ashby, 2011; Huettel et al., 2008; Pauling & Coryell, 1936). When neurons become active, the excess oxygen-rich blood flows through active regions. As a result, the relative proportion of deoxygenated blood in that particular region decreases, and the BOLD signal increases. There are other contrast mechanisms being used, such as vascular space occupancy (VASO), signal enhancement by extravascular water protons (SEEP), perfusion-weighted imaging (PWI); nevertheless, these mechanisms are not as sensitive as the BOLD signal (Stroman, 2011). Thus, the BOLD signal is still considered as the mainstream of the fMRI study (Logothetis, 2002).

Signal Detection and Spatial Smoothing

In fMRI, data are presented and analyzed in image format; in a three-dimensional image the basic unit is called a voxel (volume element), and in a two-dimensional image, the unit is referred as a pixel (picture element). A successful fMRI study is to build an image , or a "map", to localize psychological functions to brain regions. This is generally achieved by conducting an experiment on the subjects to evoke the cognitive (or perceptual, motor, mnemonic, etc.) processes with a stimulus, such as seeing faces, listening to music, or performing tasks, and at the same time measuring the response in the brain. Thus, the goal of any fMRI study comes down to detecting signals embedded in the functional imaging data, which is called "signal detection" (Petersson, Nichols, Poline, & Holmes, 1999).

Typically, the quality of the raw data coming right out of the MRI scanner is not good enough for any analysis. To remove uninteresting artifacts and to allow statistical analyses on the fMRI data, several preprocessing operations are required, such as motion correction, spatial normalization, and spatial smoothing (Ashby, 2011; Poldrack et al., 2011). Spatial smoothing, above all, plays an important role in signal detection for fMRI data.

Smoothing, also called filtering or blurring, is to filter the image by replacing the BOLD response on each voxel with a weighted average of neighboring voxels. The voxel being smoothed always has strongest weight, while the weight decreases with distance at a rate in accordance with the amount of smoothing. As a result,

spatial smoothing removes the variation in the image by blurring the sharp edges of the peaks and valleys.

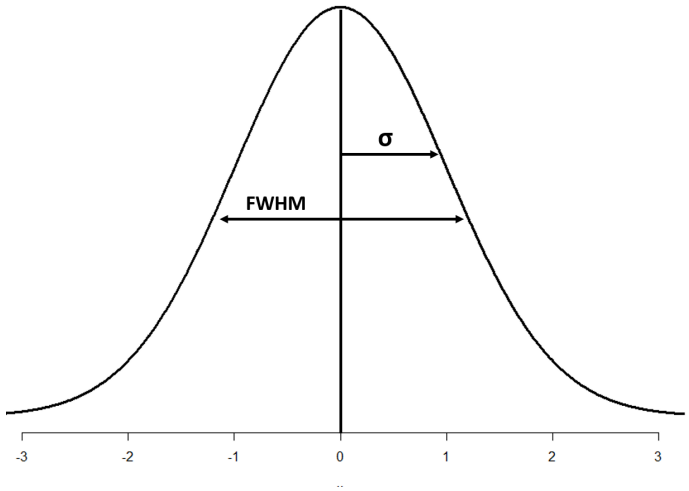
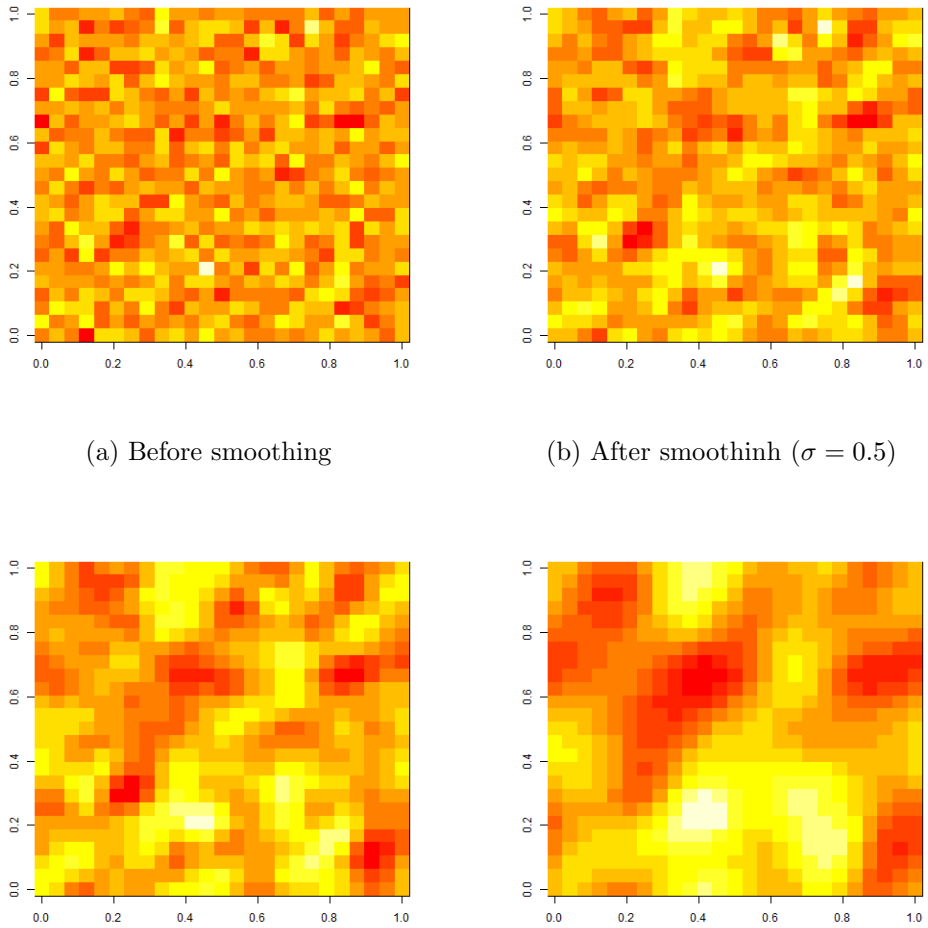


Figure 2. Relation between standard deviation and FWHM

The foremost advantage of spatial smoothing is to reduce noise, and thus increase the signal-to-noise ratio. In most cases, the BOLD signal (if there is any) in the fMRI data is weak compared to the background noise, making it hard to detect (Huettel et al., 2008; Worsley, 1996). Therefore, increasing signal-to-noise ratio would greatly improve the chance of detecting activation. In addition, according to the matched filter theorem in signal processing, the signal-to-noise ratio will be its maximum if the signal is smoothed with a kernel whose shape and size match the signal (Rodenfeld & Kak, 1982). Hence a Gaussian-shaped 10 mm signal is best detected by a Gaussian kernel with a width of 10 mm. Nevertheless, it is important to remember that smoothing also reduces the resolution of the image and decreases the number of independent voxels (Brett et al., 2003). Thus to some degree the autocorrelation embedded in the fMRI data comes from spatial smoothing.



(a) Before smoothing

(b) After smoothinh ($\sigma = 0.5$)

(c) After smoothing ($\sigma = 1$)

(d) After smoothing ($\sigma = 2$)

Figure 3. Effect of the width of smoothing kernel

Spatial smoothing is implemented by convolving the fMRI image with a function, often called a kernel. The most commonly used one is the Gaussian kernel, with the shape of a normal curve (Ashby, 2011; Lazar, 2008; Poldrack et al., 2011; Worsley, 2002). The amount of smoothing is determined by the width of the normal curve. In practice the width is measured in full width at half-maximum (FWHM), which is the width of the kernel at the point where it is at half of its maximum, and is related to standard deviation (σ) as follows: FWHM=$\sigma\sqrt{8\ln 2}$ (see Figure 2).

Note that the wider the FWHM the greater the smoothing. Figure 3 shows the effect of the smoothing parameter. (a) is a 25×25 image with Gaussian white noise, whereas (b), (c), and (d) are the images after smoothing with a Gaussian kernel at $\sigma = 0.5, 1$, and 2, respectively. In practice, a general recommendation is to smooth the image using a kernel with a FWHM twice the voxel width (Ashby, 2011; Poldrack et al., 2011; Stroman, 2011). Nevertheless, there is no universal golden rule regarding how much smoothing should be applied.

Multiple Comparisons Problem in Functional Magnetic Resonance Imaging

Since in the fMRI study signal detection is through changes in the BOLD response, the easiest way to test the hypothesis is to compare an experimental condition with a control condition through the use of a Student's t-test. Note that depending on the nature of the experiment the statistical analysis could be more complicated; nevertheless, without loss of generality, a t-test is assumed at this moment.

To test the statistical hypothesis in fMRI, generally it is done by applying the statistical test on each voxel (or pixel if it is a two-dimensional image) and then each voxel would be assigned a test statistic, such as a t-value. With the test-statistics from all the voxels, the end product is called a statistical parametric map. The statistical parametric map is color-coded based on the value of the test statistic and is a useful graphical representation to display the result of the brain activation (Huettel et al., 2008; Poldrack et al., 2011). Since a t-test is applied to each voxel of

the image, the map is sometimes referred as a t-map; likewise, if an F-test or a χ^2-test is used, then the map is called an F-map or a χ^2-map, respectively.

Standard functional imaging data usually contain around 130,000 voxels (or pixels), that is to say there are enormous amount of tests being conducted when producing a statistical parametric map (Ashby, 2011; Huettel et al., 2008), resulting in the multiple comparisons problem in the fMRI study. Consequently, the Type I error rate would be inflated if no correction is done. Suppose we have a t-map with 130,000 voxels. At an α value of 0.05, it means approximately 6,500 tests are false positive that none of these significant results are related to real signals (Brett et al., 2003; Petersson et al., 1999).

One way to solve the multiple comparisons problem is through controlling for the family-wise error rate (FWER), which is often called "height thresholding" in neuroimaging. The idea is to choose a threshold so that any test statistics above the threshold are unlikely to be false-positive while maintaining the FWER (i.e., Type I errors among all the tests) at an acceptable level, say, .05. The Bonferroni correction is often used for controlling FWER in traditional statistical analysis, and it is easy to implement (Bender & Lange, 2001; Oehlert, 2000). To use the Bonferroni correction, simply divide the significance threshold (α) by the number of tests being conducted simultaneously.

However, given the nature of the fMRI data, the Bonferroni correction is not appropriate for the following reasons. First of all, the Bonferroni correction is notoriously conservative when the number of tests becomes too large (Huettel et al.,

2008; Poldrack et al., 2011). This is especially the case in fMRI given that fMRI data are abundant. If the image includes 10,000 voxels, applying the Bonferroni correction will reduce the α value from .05 to .000005, corresponding to a z-score of 4.42. While the Bonferroni correction minimizes the probability of Type I error across all voxels in a brain image, it also decreases the power of detecting any real activation with such a stringent threshold. Second, the Bonferroni correction assumes the tests are independent of each other (Gelman, Hill, & Yajima, 2012). However, not only smoothing introduces spatial correlation but also some degree of spatial autocorrelation is almost always expected in fMRI data. In general the value from any voxel in an image tends to be similar to adjacent voxels (Brett et al., 2003; Petersson et al., 1999). Thus, it is not possible to know exactly how many independent voxels exist in an image. Because of the above limitations of the Bonferroni correction, alternatives were proposed. The most commonly seen approaches include false discovery rate (Benjamini & Hochberg, 1995), permutation method, and Gaussian random field theory. The next section presents a detailed discussion of random field and Gaussian random field theory.

Random Field Theory

Random fields, like stochastic processes, are one kind of random functions. As a stochastic process is a collection of random variables defined over time (e.g, time-series), a random field is a generalization of a stochastic process and is defined as a set of random variables over an N-dimensional Euclidean space (Adler, 1981; Worsley, 2002). The history of random fields could date back to the early 20$^{\text{th}}$

century but no serious advances were made until the middle of the century. Longuet-Higgins' classical work (1952) on ocean surface is considered to be the first one studying the properties of Gaussian random fields. Meanwhile Yaglom, a Soviet statistician, also published an important paper in 1957, establishing the spectral theory of homogeneous random fields. The theory of random fields was consolidated by Adler (1981) which is based on his doctoral thesis. In his work, Adler explored the sample path properties of Gaussian and Gaussian related random fields. At the same time random fields were also applied in a variety of disciplines such as astrophysics, forestry, geology, meteorology, seismology, turbulence, just to name a few (Adler, Bartz, & Kou, 2011).

There are different types of random fields depending on how the parameters are defined. In the current study, the random field is defined as a function $X(\boldsymbol{t}), \boldsymbol{t} \in T$ that the domain T is a subset of N-dimensional Euclidean space, and $X(\boldsymbol{t})$ is smooth or at least continuous in $\boldsymbol{t}$. When $N = 1$, we have a continuous-time stochastic process; when $N = 2$, it is a random surface; and when $N = 3$, a spatial process (Worsley, 2002).

Gaussian Random Field Theory

An important special class of random fields is Gaussian random field, which is one that is multivariate Gaussian at all finite sets of points, that is, $[X(\boldsymbol{t_1}), \ldots, X(\boldsymbol{t_n})]$ must be multivariate Gaussian for all $0 < n < \infty$ and all $\boldsymbol{t_j} \in T$. Because the multivariate Gaussian is completely defined by its mean vector and covariance matrix, the Gaussian random field is also defined uniquely by its mean

function

$$\mu(\boldsymbol{t}) = E[X(\boldsymbol{t})], \tag{1}$$

and its covariance function

$$R(\boldsymbol{s}, \boldsymbol{t}) = Cov[X(\boldsymbol{s}), X(\boldsymbol{t})]. \tag{2}$$

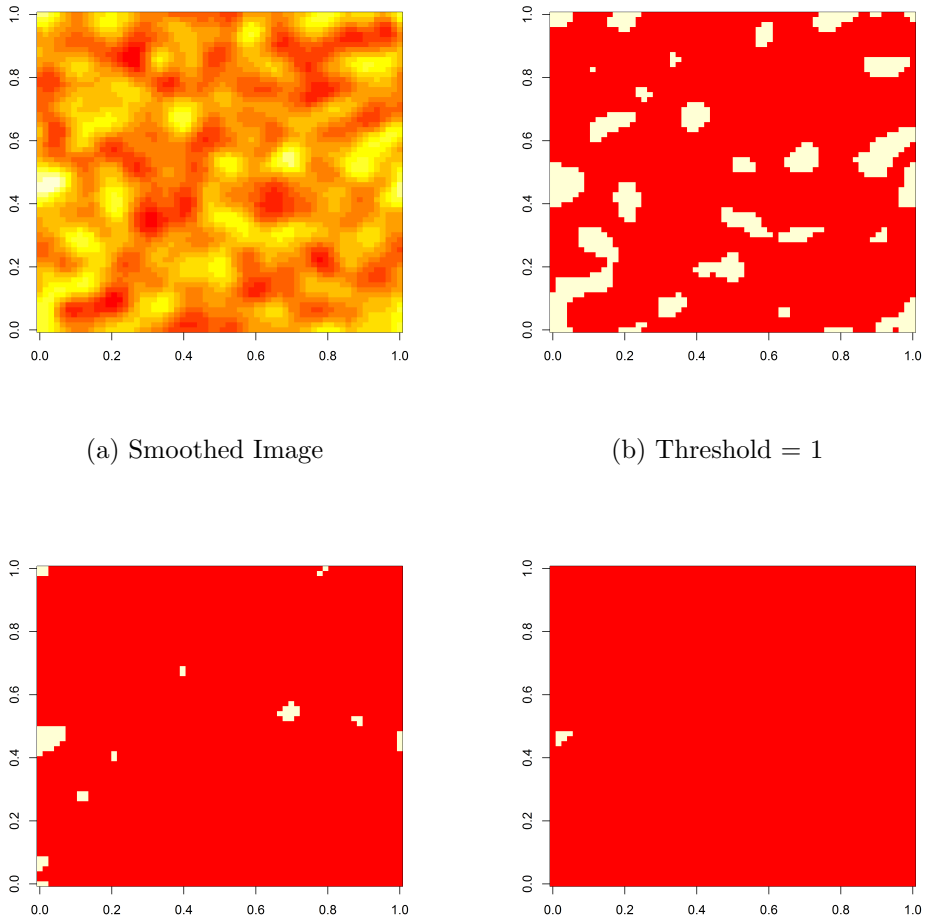
One of the most fruitful applications of Gaussian random fields is in neuroimaging, where the theory is largely expanded by Worsley and his colleagues. When fMRI was introduced in the early 1990s, Worsley noticed the statistical parametric map could be treated as a lattice representation of a continuous field and thus, the topological features of the map could be modeled by using Adler's theory of random fields (1981). Worsley's idea to height thresholding in neuroimaging is to approximate the probability of the global maximum (i.e., p-value) of a smooth image through the use of the expected Euler characteristic of an excursion set (which is simply the region above a fixed threshold). The theory is explained as follows.

The Euler characteristic is a topological measure which is used to describe the surface of a polyhedron. In one-dimensional case the Euler characteristic counts the number of connected components; for two-dimensions, the Euler characteristic counts the number of connected components minus the number of holes; in three-dimension the Euler characteristic counts the number of connected components of the excursion set minus the number of holes plus the number of hollows (Worsley, 1996). As the threshold increases the holes in the excursion set

begin disappearing until only local maxima are left. At this point the Euler characteristic only counts the number of local maxima. For a much higher threshold, only the global maximum remains—the Euler characteristic counts one if the global maximum exceeds the threshold and zero otherwise. At this point, the expected Euler characteristic approximates the probability of the global maximum of the statistical image, that is, the p-value.

Figure 4 provides a visual example illustrating the relationship between the height threshold and the Euler characteristic. (a) represents a two-dimensional smoothed random field. (b), (c) and (d) show the excursion set after thresholding at cutoff value of 1, 2, and 3, respectively. Any value less than the threshold have been set to zero (displayed in red); otherwise the values have been set to one (displayed in white). As we can see, when the threshold increases to 2 and 3, only the local maxima remain and we could just count the number of blobs. Take (d) for an instance, after thresholding at cutoff value of 3, there is only one blob left, indicating only one area with values larger than 3. Thus the observed Euler characteristic is one in this case.

Gaussian random field is considered a better approach than the Bonferroni correction because it takes the spatial correlation into account, which results in a less conservative threshold. To use random field theory in fMRI, it requires three assumptions. The first assumption is that the error field is a Gaussian random field, which is essential to the results derived by Adler (1981) and Worsley, Evans, Marrett, and Neelin (1992), although this is not easy to check in practice (Brett et



(a) Smoothed Image

(b) Threshold = 1

(c) Threshold = 2

(d) Threshold = 3

Figure 4. An example of height thresholding in two-dimensional images

al., 2003; Petersson et al., 1999). The second assumption is that the error fields should be smooth (Friston, Holmes, Poline, Price, & Frith, 1996; Worsley, 2002). It has been shown that if the image is not smooth enough the threshold obtained from random field theory may actually be more conservative than the Bonferroni correction (Marchini & Presanis, 2004; Nichols & Hayasaka, 2003). Last, the autocorrelation function of the field should be twice-differentiable (Adler, 1981).

Global maximum X_{max}

Define X_{max} as the global maximum of a statistical field. Explicit calculation of the probability of X_{max} using the expected value of the Euler characteristic were already obtained for the z-, t-, χ^2-, F-, Hotelling's T^2, Gaussian scale space, χ^2 scale space, and Gaussian rotation space field (Adler, 1981; Cao & Worsley, 1999; Shafie et al., 2003; Siegmund & Worsley, 1995; Worsley, 1994). A detailed summary is provided by Cao and Worsley (2001).

The idea of using random field theory for thresholding in fMRI is to obtain p-value through the use of X_{max}. However, there was no justification for X_{max} as a test statistic until Siegmund and Worsley's paper in 1995. They showed X_{max} is the likelihood ratio test statistic in the Gaussian scale space random field when testing one signal with unknown location and scale. The scale space random filed refers to a random field with $N + 1$ dimensions that N dimensions for the location plus one dimension for the width of the smoothing kernel. In practice, the scale of the signal is usually unknown; hence different sizes of smoothing kernel should be applied to improve detection sensitivity (Brett et al., 2003; Rohani, Shafie, & Noorbaloochi, 2006; Shafie et al., 2003). The idea of the multifilter approach was first suggested by Poline and Mazoyer (1994), which led to Gaussian scale space random field by Siegmund and Worsley (1995). Worsley (2001) also extended the scale space to χ^2 field, which considers the case of several independent Gaussian images, each with the same width signal at the same location but may be with different amplitudes.

The current study is based on Siegmund and Worsley's paper (1995). Thus a brief discussion of their work about the Gaussian scale space random field is presented below. Suppose we model the observed data with a random field $Z(\boldsymbol{t})$, which comprises of two parts: the signal and the noise. With a fixed value $t_0 \in C$, $\xi \geq 0$, and $\sigma_0 > 0$, $Z(\boldsymbol{t})$ could be expressed using the stochastic differential equation as follows, where $\boldsymbol{t} \in \mathbb{R}^N$

$$dZ(\boldsymbol{t}) = \xi \sigma_0^{-N/2} f\left\{\sigma_0^{-1}(\boldsymbol{t} - \boldsymbol{t}_0)\right\} d\boldsymbol{t} + dW(\boldsymbol{t}). \tag{3}$$

Here C is the search region and is a subset of N-dimensional Euclidean space. The function dW is a white noise. The unknown vector parameter $(\xi, \boldsymbol{t_0}, \sigma_0)$ represents the amplitude, location, and scale of the signal. f is an N-dimensional function for the shape of the signal, which is usually positive, symmetric and unimodal. Let k be an N-dimensional kernel and normalized such that $\int k(\boldsymbol{t})^2 = 1$. A commonly-used kernel is the Gaussian kernel where

$$k(\boldsymbol{t}) = \pi^{(-N/4)} exp(-\frac{1}{2}\|\boldsymbol{t}\|^2). \tag{4}$$

Given a kernel function k, Siegmund and Worsley (1995) defined the Gaussian scale space random field

$$X(\boldsymbol{t}, \sigma) = \sigma^{-N/2} \int_{\mathbb{R}^n} k[\sigma^{-1}(\boldsymbol{h} - \boldsymbol{t})] dZ(\boldsymbol{h}), \tag{5}$$

where the term $\sigma^{-N/2}$ is to make the variance =1. We can expand Equation (5) and obtain a general form of a Gaussian scale space random field,

$$\begin{aligned} X(\boldsymbol{t},\sigma) = (\sigma_0\sigma)^{-N/2}\xi \int f\left\{\sigma_0^{-1}(\boldsymbol{h}-\boldsymbol{t_0})\right\} k\left\{\sigma^{-1}(\boldsymbol{h}-\boldsymbol{t})\right\} d\boldsymbol{h} \\ + \sigma^{-N/2} \int k\left\{\sigma^{-1}(\boldsymbol{h}-\boldsymbol{t})\right\} dW(\boldsymbol{h}). \end{aligned} \tag{6}$$

If we denote the first and the second terms in Equation (6) by μ and W^*, then

$$X(\boldsymbol{t},\sigma) = \mu(\boldsymbol{t},\sigma;\xi,\boldsymbol{t_0},\sigma_0) + W^*(\boldsymbol{t},\sigma), \tag{7}$$

where ξ, $\boldsymbol{t_0} = (\boldsymbol{t_{01}},\ldots,\boldsymbol{t_{0N}})'$ are unknown parameters, σ_0 is also an unknown parameter but lie in the interval $[\sigma_1,\sigma_2]$, and σ is the scale parameter of the kernel function. $W^*(\boldsymbol{t},\sigma)$ is an N-dimensional Gaussian random field with zero mean, unit variance, and covariance function of the form

$$(\sigma_1\sigma_2)^{-N/2} \int k\left\{\sigma_1^{-1}(\boldsymbol{h}-\boldsymbol{t_1})\right\} k\left\{\sigma_2^{-1}(\boldsymbol{h}-\boldsymbol{t_2})\right\} d\boldsymbol{h}. \tag{8}$$

Let $X(\boldsymbol{t})$ be the scale space Gaussian random field defined in Equation (5), where $\boldsymbol{t} \in \mathbb{R}^N$, and $k = f$ (that is, the condition of matched filter theorem mentioned in Chapter II is met.). Thus, for testing the null hypothesis $H_0 : \xi = 0$ against the alternative $H_a : \xi > 0$, the log likelihood function (Siegmund & Worsley, 1995) is

$$\xi X(\boldsymbol{t_0},\sigma_0) - \xi^2/2. \tag{9}$$

Thus, the test statistic is

$$X_{max} = \max_{\boldsymbol{t},\sigma} X(\boldsymbol{t}, \sigma), \tag{10}$$

where the maximum extends over $\boldsymbol{t} \in C$ and $\sigma_1 < \sigma < \sigma_2$; the proof can be found in Rohani (2003). As a result, Siegmund and Worsley (1995) showed the global maximum X_{max} is the likelihood ratio test statistic. Furthermore, though the distribution of X_{max} is unknown, Siegmund and Worsley (1995) were able to approximate the probability of X_{max} using the Euler characteristic approach and the volume of tubes.

It is important to note that X_{max} is the likelihood ratio test statistic when testing $H_0 : \xi = 0$ against $H_a : \xi > 0$. If one would like to have a two-tailed alternative ($H_a : \xi \neq 0$), then Shafie (2014) proved in scale space the likelihood ratio test statistic is X^2_{max}, defined as

$$X^2_{max} = \max_{\boldsymbol{t},\sigma} X^2(\boldsymbol{t}, \sigma). \tag{11}$$

With $X^2(\boldsymbol{t}, \sigma)$ being a scale space χ^2 field, the probability of X^2_{max} could be approximated using the Euler characteristic formula mentioned in Worsley (2001). Shafie (2014) also showed that when σ is fixed, then X^2_{max} is still the likelihood ratio statistic and Equation (11) is reduced to

$$X^2_{max} = \max_{\boldsymbol{t}} X^2(\boldsymbol{t}). \tag{12}$$

Likewise, its probability could be approximated through the use of the Euler characteristic formula provided in Worsley (1994).

Nonetheless, X_{max} and X^2_{max} are the test statistics for testing one signal in the search region. If one would like to test multiple signals simultaneously, in this case the likelihood ratio test statistic is neither X_{max} nor X^2_{max} anymore. Consequently, none of the existing expected Euler characteristic formulas can be used to obtain the p-values for thresholding when testing for more than one signal.

From One Signal to Multiple Signals

A lot of early brain research focuses on the localization of brain function that different brain regions are specialized for different functions (McIntosh, 2010). Recently, more and more evidence is gathered and supports that no brain region exists in isolation. Connectivity analysis, for example, is one way for neuroscientists to identify functional relationship among brain regions. Functional MRI is widely used for functional connectivity analysis because it has the capability to detect the activation throughout the entire brain (Huettel et al., 2008).

The simplest case in connectivity analysis is coactivation, where two or more distinct brain areas are simultaneously activated during an experimental task. Huettel, Song, and McCarthy (2008) provided an example of coactivation in motor task: "squeezing your left hand will result in activity in the precentral gyrus in the right hemisphere and cerebellum in the left hemisphere" (p.389). Of course, coactivation does not imply the brain regions are truly connected as a neural network.

To do the connectivity analysis to understand deeper about how neural functions are integrated in the human brain, it is necessary to test multiple signals simultaneously. Unfortunately, not much work has been done. Hartvig (1999) proposed a spatio-temporal model for estimating the activation patterns formed by multiple signals, using a Bayesian approach. However, the focus of Hartvig's model is estimation, rather than hypothesis testing. There is no result regarding thresholding using Gaussian random field for multiple signals. Therefore, the current study attempts to fill the gap.

Summary

For the last 20 years, fMRI has flourished and now becomes one of the dominant techniques used in neuroimaging. One important challenge of the statistical analysis in fMRI studies is the multiple comparisons problem. Given the nature of fMRI data, correction must be done in order to control for Type I error rate. Traditionally the Bonferroni correction is used—the adjustment, however, is said to be too conservative with the massive number of statistical tests being conducted in fMRI. Random field theory, on the other hand, is an alternative to height thresholding. The key of using random field theory for thresholding is to obtain the p-value by approximating the probability of the global maximum X_{max}, often through the use of the expected value of the Euler characteristic. The justification for X_{max} was provided by Siegmund and Worsley (1995) that they showed it is the likelihood ratio test statistic when testing one signal in the search region. However, the explicit calculations of using the Euler characteristic are only

available for certain fields with known distributions, such as Gaussian, t, χ^2, etc. If the neuroscientists are interested in observing more than one signal, then it is inappropriate to use X_{max} as the test statistic. Further, there is closed form solution in this case to approximate the probability of the likelihood ratio test statistic. Therefore, the current study attempts to use Monte Carlo simulation to approximate the probability of the proposed new test statistic for testing two signals.

CHAPTER III

METHODOLOGY

The purpose of the current study was to propose a new likelihood ratio test statistic Y_{max} for testing two signals and to obtain the associated p-values for thresholding. The power of Y_{max} was also examined under various conditions. To be more meaningful, the author was interested in comparing the power of Y_{max} to another test statistic in order to see how well Y_{max} performed in terms of signal detection. There were two choices: X_{max} and X^2_{max}. Note that the statistical hypotheses for Y_{max} is two-tailed as follows,

$$H_0 : \xi_1 = \xi_2 = 0$$

$$H_a : \text{At least one of the } \xi_i \neq 0,\ i = 1, 2.$$

Recall for X_{max} the hypothesis is one-tailed ($H_a : \xi > 0$) whereas for X^2_{max} it is two-tailed ($H_a : \xi \neq 0$). Hence, even though X_{max} is more often used in practice than X^2_{max}, to be consistent the comparison of the power in this study was focused on Y_{max} against X^2_{max}.

This chapter illustrates the methodology to investigate the following research questions:

Q1 How to obtain p-values of the likelihood ratio test statistic Y_{max} for testing two signals simultaneously in two-dimensional images using Monte Carlo simulation?

Q2 To what extent does the distance between the locations of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q3 To what extent does the amplitude of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q4 To what extent does the scale of two signals affect the power of the likelihood ratio test statistic Y_{max}?

Q5 Does Y_{max} outperform X^2_{max} in power when testing two signals simultaneously under the above conditions?

There were two scenarios being considered in this study: the width of the smoothing kernel (σ) being fixed, and the scale space method where the kernel width σ was not fixed. Thus, this chapter comprises of two major sections, starting with the case of fixed σ; then the section of σ not fixed comes later. In each section, first the forms of the likelihood ratio test statistic Y_{max} are introduced. Second, the Monte Carlo simulation procedures used to obtain the empirical distribution of Y_{max} are described, following the illustration of the simulation procedures along with the parameter schemes to address the power of Y_{max}.

As mentioned in Chapter II, Siegmund and Worsley (1995) showed the likelihood ratio test statistic is equivalent to

$$X_{max} = \max_{\boldsymbol{t},\sigma} X(\boldsymbol{t}, \sigma) \tag{13}$$

when testing one positive-valued signal in a Gaussian scale space random field. And if one would like to test one signal with a non-directional hypothesis (that the amplitude could be either positive or negative), Shafie (2014) showed the likelihood ratio test statistic should be

$$X^2_{max} = \max_{\boldsymbol{t},\sigma} X^2(\boldsymbol{t}, \sigma). \tag{14}$$

Nevertheless, when the number of signals increases from one to n, the random field $Z(\boldsymbol{t})$ is now defined as

$$dZ(\boldsymbol{t}) = \sum_{i=1}^{n} \xi_i \sigma_{0_i}^{-N/2} f\left\{\sigma_{0i}^{-1}(\boldsymbol{t} - \boldsymbol{t}_{0i})\right\} d\boldsymbol{t} + dW(\boldsymbol{t}), \tag{15}$$

where ξ_i, $\boldsymbol{t}_{0i}$, σ_{0i}, $i = 1, \cdots, n$ represent the amplitude, location, and scale of the i^{th} signal, respectively. Thus, both X_{max} and X^2_{max} are no longer the likelihood ratio test statistics when testing multiple signals (Shafie, 2014).

Likelihood Ratio Test Statistic for Testing Two Signals When the Kernel Width Is Fixed

When testing two signals simultaneously, Shafie (2014) showed the log likelihood ratio can be written as

$$Y(\boldsymbol{t_1}, \boldsymbol{t_2}) = \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X} \tag{16}$$

where $\boldsymbol{X}' = (X(\boldsymbol{t_1}), X(\boldsymbol{t_2}))'$ in which $X(\boldsymbol{t})$ is a Gaussian random field and $(\boldsymbol{t_1}, \boldsymbol{t_2}) \in C \times C$. $\boldsymbol{R}$ is the correlation function. Hence, with fixed σ, the likelihood ratio test statistic for testing two signals is

$$Y_{max} = \max_{\boldsymbol{t_1}, \boldsymbol{t_2}} \left\{ \frac{1}{1 - R^2(\boldsymbol{t_1}, \boldsymbol{t_2})} \left[X^2(\boldsymbol{t_1}) + X^2(\boldsymbol{t_2}) - 2X(\boldsymbol{t_1})X(\boldsymbol{t_2})R(\boldsymbol{t_1}, \boldsymbol{t_2}) \right] \right\}. \quad (17)$$

The correlation function $R(\boldsymbol{t_1}, \boldsymbol{t_2})$ is

$$(\sigma)^{-N} \int k\left\{\sigma^{-1}(\boldsymbol{h} - \boldsymbol{t_1})\right\} k\left\{\sigma^{-1}(\boldsymbol{h} - \boldsymbol{t_2})\right\} d\boldsymbol{h}, \quad (18)$$

and if k is a Gaussian smoothing kernel, in two-dimension the correlation function simplifies to

$$R(\boldsymbol{t_1}, \boldsymbol{t_2}) = exp(-(\boldsymbol{t_1} - \boldsymbol{t_2})'(\boldsymbol{t_1} - \boldsymbol{t_2})/4\sigma^2). \quad (19)$$

It should be noted that the log likelihood ratio $Y(\boldsymbol{t_1}, \boldsymbol{t_2})$, like $X^2(\boldsymbol{t})$, is a χ^2 field. However, the Euler characteristic formula developed in Worsley (1994) cannot be used to approximate the probability of Y_{max} because $Y(\boldsymbol{t_1}, \boldsymbol{t_2})$ is not stationary which is one of the key assumptions to use the Euler characteristic for approximation.

Empirical Distribution of Y_{max}

Since the distribution of Y_{max} is unknown and there is no closed form solution at hand to approximate the probability of Y_{max}, a Monte Carlo simulation was used to obtain the empirical distribution of Y_{max} in two dimensional images. R Version 3.1.1 was employed to simulate images: the built-in function rnorm() was

used to generate Gaussian white noise, and the fft() function was used for Fourier transformation. Each image was a unit square with 64×64 grids. The resolution the author initially planned to have was 128×128, which is commonly used in practice and has also been employed in previous simulation studies (Shafie et al., 2003; Siegmund & Worsley, 1995). However, when calculating the correlation matrices the author encountered memory issues with 128×128 images. Thus, the image resolution was compromised and a size of 64×64 was chosen instead in the current study. A Gaussian-shaped kernel was used for spatial smoothing and it was centered to ensure the entire image was smoothed evenly.

Under the null hypothesis, there is no signal. The following explains the simulation procedure: (1) generate an independent Gaussian white noise image and smooth the image with a Gaussian kernel, (2) calculate the correlation matrix using Equation (19), and (3) for each image use Equation (17) to obtain Y_{max}. It is important to note that when $\boldsymbol{t_1} = \boldsymbol{t_2}$, Equation (17) would become undefined and these cases were tossed out. Given that the width of the kernel was fixed, ten σ values were used as the smoothing kernel: 0.02, 0.0258, 0.0334, 0.0431, 0.0557, 0.0719, 0.0928, 0.1199, 0.1549, 0.2. The smallest and biggest values, 0.02 and 0.2, were chosen empirically based on the image size being used in this study, and the other values were equally spaced over the interval $[log(0.02), log(0.2)]$ and rounded to four decimal places (Siegmund & Worsley, 1995). The log scale was to ensure the resolution in scale space is constant, and note the same values were also used later for the scale space method. For each fixed kernel σ, 5000 replications were run.

Because it took nearly three hours for only 5000 replications; thus, more number of replications was not attempted. An empirical null distribution of Y_{max} was formed based on these 5000 replications for each σ, and both the 95^{th} and 99^{th} percentile were calculated to serve as the critical values at α of .05 and .01, respectively.

Power of the Likelihood Ratio Test Statistic Y_{max}

In this study the power was calculated as the number of Y_{max}'s obtained under the alternative hypothesis exceeding the critical value divided by the number of replications. Under the alternative hypothesis, the simulated image contained not just the noise but also the signals; hence, images with two signals under various schemes were generated. As Equation (15) implies, when there are multiple signals, the random field $Z(\boldsymbol{t})$ depends on the amplitude, location, and scale of all the signals. Therefore, the signals were manipulated through the following parameters: (a) amplitude (ξ), where there are three levels: $\xi = 0.5, 2, 4$. With two signals simulated in the present study, the amplitude of two signals, ξ_1 and ξ_2, could be different, resulting in six combinations of ξ_1 and ξ_2, as displayed in Table 1; (b) distance, where there are three levels: the signals far from each other (Far), the signals close to each other (Near), and the signals neither far from nor close to each other (Middle). The operational definition of these three levels is illustrated in the following paragraph; and (c) scale of the signals (σ_0), which has two levels: $\sigma_0 = 0.02, 0.04$. Note that the scale was held constant for both signals. In summary, a total of 36 schemes were simulated as described in Table 1. During simulation,

each σ value was combined with all 36 schemes; therefore, a total of 360 conditions were implemented, each with 5000 replications.

Table 1

Schemes of the Parameters

Scheme No.	Amplitude (ξ_1, ξ_2)	Distance	Scale (σ_0)
1	(0.5, 0.5)	Near	0.02
2			0.04
3		Middle	0.02
4			0.04
5		Far	0.02
6			0.04
7	(0.5, 2)	Near	0.02
8			0.04
9		Middle	0.02
10			0.04
11		Far	0.02
12			0.04
13	(0.5, 4)	Near	0.02
14			0.04
15		Middle	0.02
16			0.04
17		Far	0.02
18			0.04
19	(2, 2)	Near	0.02
20			0.04
21		Middle	0.02
22			0.04
23		Far	0.02
24			0.04
25	(2, 4)	Near	0.02
26			0.04
27		Middle	0.02
28			0.04
29		Far	0.02
30			0.04
31	(4, 4)	Near	0.02
32			0.04
33		Middle	0.02
34			0.04
35		Far	0.02
36			0.04

Choices of the levels of the parameters. Given that there are few simulation studies out there in this field, as the first investigation of this topic, the levels of the above parameters (amplitude, distance, and scale) were mostly chosen empirically and arbitrarily — the rationale is explained as below. For the locations of the two signals, the three levels were operationally defined according to the values of the scale. When $\sigma_0 = 0.02$, the coordinates (based on a unit square with 64×64 grids) of the 1^{st} signal (t_{11}, t_{12}) and the 2^{nd} signal (t_{21}, t_{22}) of the three levels were:

- Near: (0.04, 0) and (-0.05, 0)
- Middle: (0.17, -0.17) and (-0.17, 0.17)
- Far: (0.45, -0.45) and (-0.45, 0.45)

When $\sigma_0 = 0.04$, the coordinates of (t_{11}, t_{12}) and (t_{21}, t_{22}) were:

- Near: (0.09, 0) and (-0.1, 0)
- Middle: (0.2, -0.2) and (-0.2, 0.2)
- Far: (0.41, -0.41) and (-0.41, 0.41)

Figure 5 shows the visualization of the above conditions. Note that for the case of the signals being close, the coordinates were chosen so the two signals did not touch each other. Likewise, for the case of the signals being far away from each other, the coordinates were chosen so the signals did not touch the boundary of the image.

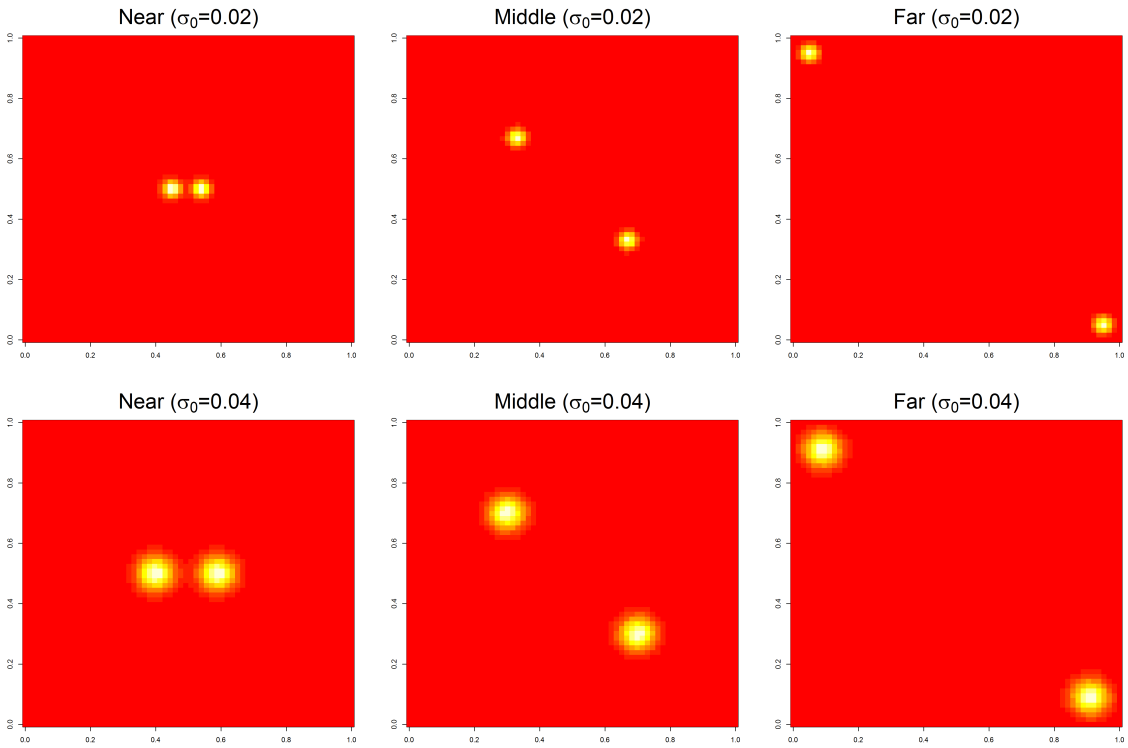


Figure 5. Three levels of distance between signals when $\xi_1 = \xi_2 = 2$

As for the values of amplitude and scale, previous research used $\xi = 6$ and $\sigma_0 = 1$ for simulation (Shafie et al., 2003; Siegmund & Worsley, 1995). However, these values were not applicable because with all the simulated images being unit squares in the current study, a signal with $\xi = 6$ and $\sigma_0 = 1$ was too big to be contained nicely within a unit square. Further, such a signal made the power go up to one regardless what the other condition is. Thus, more reasonable values for amplitude and scale were chosen instead.

To estimate the power, similar simulation procedure was used as the one for the thresholds. The only difference is that under the alternative hypothesis, images were generated with Gaussian white noise as well as two Gaussian-shaped signals. It should also be noted that when adding the signals to the noise image, the signals

were not smoothed along with the noise using the Gaussian kernel, but rather were created using the following equation prior to being added to the noise image,

$$\frac{2\xi_i\sigma\sigma_0}{\sigma^2+\sigma_o^2}exp\left(\frac{-\|\boldsymbol{t}-\boldsymbol{t_{0i}}\|^2}{2(\sigma^2+\sigma_0^2)}\right). \tag{20}$$

This is because the form of the kernel (k) and signal (f) are known (as defined in Equation (4)), when $N=2$ it is

$$k(\boldsymbol{t})=f(\boldsymbol{t})=\pi^{(-1/2)}exp\left(-\frac{1}{2}\|\boldsymbol{t}\|^2\right). \tag{21}$$

Recall when testing n signals, the random field $Z(\boldsymbol{t})$ is

$$dZ(\boldsymbol{t})=\sum_{i=1}^{n}\xi_i\sigma_{0_i}^{-N/2}f\left\{\sigma_{0i}^{-1}(\boldsymbol{t}-\boldsymbol{t}_{0i})\right\}d\boldsymbol{t}+dW(\boldsymbol{t}), \tag{22}$$

and with a kernel function k, the Gaussian scale space random field is

$$X(\boldsymbol{t},\sigma)=\sigma^{-N/2}\int_{\mathbb{R}^n}k[\sigma^{-1}(\boldsymbol{h}-\boldsymbol{t})]dZ(\boldsymbol{h}). \tag{23}$$

Assuming all n signals have the same scale σ_0, Equation (23) can be expanded and the mean of the i^{th} signal is expressed as,

$$(\sigma_0\sigma)^{-N/2}\xi_i\int f\left(\frac{\boldsymbol{h}-\boldsymbol{t}_{0i}}{\sigma_0}\right)k\left(\frac{\boldsymbol{h}-\boldsymbol{t}}{\sigma}\right)d\boldsymbol{h}. \tag{24}$$

In two-dimension, the above equation becomes

$$(\sigma_0\sigma)^{-1}\xi_i \int f\left(\frac{\boldsymbol{h}-\boldsymbol{t}_{0i}}{\sigma_0}\right) k\left(\frac{\boldsymbol{h}-\boldsymbol{t}}{\sigma}\right) d\boldsymbol{h}. \tag{25}$$

If we assume $\boldsymbol{t}|\boldsymbol{h} \sim N(\boldsymbol{h}, \sigma^2\boldsymbol{I})$, which has the probability density function

$$f(\boldsymbol{t}|\boldsymbol{h}) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{\|\boldsymbol{t}-\boldsymbol{h}\|^2}{2\sigma^2}\right),$$

and $\boldsymbol{h} \sim N(\boldsymbol{t}_{oi}, \sigma_0^2\boldsymbol{I})$ with the probability density function

$$f(\boldsymbol{h}) = \frac{1}{2\pi\sigma_0^2} exp\left(-\frac{\|\boldsymbol{h}-\boldsymbol{t}_{0i}\|^2}{2\sigma_0^2}\right).$$

And we further assume $\boldsymbol{t} = \boldsymbol{h} + \epsilon$ and $\epsilon \sim N(0, \sigma^2\boldsymbol{I}) \Rightarrow \boldsymbol{t} \sim N(\boldsymbol{t}_{0i}, (\sigma_0^2+\sigma^2)\boldsymbol{I})$. Thus Equation (25) becomes

$$\begin{aligned}
&(\sigma_0\sigma)^{-1}\xi_i \int f\left(\frac{\boldsymbol{h}-\boldsymbol{t}_{0i}}{\sigma_0}\right) k\left(\frac{\boldsymbol{h}-\boldsymbol{t}}{\sigma}\right) d\boldsymbol{h} \\
&= (\sigma_0\sigma)^{-1}\xi_i \times 2\pi^{\frac{1}{2}} \times 2\pi^{\frac{1}{2}} \times \sigma^2 \times \sigma_0^2 \underbrace{\int \frac{1}{2\pi\sigma_0^2} exp\left(\frac{-\|\boldsymbol{h}-\boldsymbol{t}_{0i}\|^2}{2\sigma_0^2}\right) \times \frac{1}{2\pi\sigma^2} exp\left(\frac{-\|\boldsymbol{t}-\boldsymbol{h}\|^2}{2\sigma^2}\right) d\boldsymbol{h}}_{\boldsymbol{t}\sim N(\boldsymbol{t}_{0i},(\sigma_0^2+\sigma^2)\boldsymbol{I})} \\
&= (\sigma_0\sigma)^{-1}\xi_i 4\pi\sigma^2\sigma_0^2 \times \frac{1}{2\pi(\sigma^2+\sigma_0^2)} exp\left(\frac{-\|\boldsymbol{t}-\boldsymbol{t}_{0i}\|^2}{2(\sigma_0^2+\sigma^2)}\right) \\
&= \frac{2\xi_i\sigma\sigma_0}{\sigma^2+\sigma_0^2} exp\left(\frac{-\|\boldsymbol{t}-\boldsymbol{t_{0i}}\|^2}{2(\sigma^2+\sigma_0^2)}\right).
\end{aligned} \tag{26}$$

So the white noise was first smoothed with the Gaussian kernel and then the two "smoothed" signals were added. Then the same procedure mentioned in the

previous subsection was used to obtain Y_{max} under the alternative hypothesis. For each condition, 5000 replications were run, and in each condition a percentage of Y_{max} exceeding the critical value was calculated and reported as the power at α of .05 and .01, respectively.

Comparison of the Power between Y_{max} and X^2_{max}

In order to compare the power between Y_{max} and X^2_{max}, one also needs to know the thresholds of X^2_{max}. In the current study, the author used simulation to approximate the probability of X^2_{max} to keep consistency with Y_{max} of which the probability is also approximated by simulation, as opposed to employing the Euler characteristic formula for χ^2 field. Another reason is that the resolution of the simulated images was fairly low (64×64); thus it is more appropriate to approximate the probability of X^2_{max} using simulation, rather than the Euler characteristic formula. When σ was fixed, Equation (14) was reduced to

$$X^2_{max} = \max_{\boldsymbol{t}} X^2(\boldsymbol{t}), \tag{27}$$

and was used to obtain X^2_{max}. To examine the power of X^2_{max}, as in the case of Y_{max}, images under the same null and alternative hypotheses were simulated. First the empirical distribution of X^2_{max} was generated with 5000 replications for each kernel width, and both the 95^{th} as well as the 99^{th} percentile were obtained from the distribution to serve as thresholds. Then X^2_{max} were calculated from the images under the alternative hypothesis and compared to the critical values for power.

Likelihood Ratio Test Statistic for Testing Two Signals in Scale Space

According to the matched filter theorem, one should use a kernel with the width matched to the extent of the signal so the signal will be best detected. However, in practice the scale of the signal is not known beforehand and then of course one do not know how much smoothing should be applied to an image. Therefore, it is not very useful to smooth images with fixed kernel, which is the motivation of employing the scale space method (Poline & Mazoyer, 1994). In scale space, the search region includes not only the locations but also the width of the kernel. Thus, the likelihood ratio test statistic Y_{max} in scale space for testing two signals can be extended from Equation (17) to:

$$Y_{max} = \max_{\boldsymbol{t_1},\boldsymbol{t_2},\sigma} \left\{ \frac{1}{1 - R^2(\boldsymbol{t_1},\boldsymbol{t_2},\sigma)} \left[X^2(\boldsymbol{t_1},\sigma) + X^2(\boldsymbol{t_2},\sigma) - 2X(\boldsymbol{t_1},\sigma)X(\boldsymbol{t_2},\sigma)R(\boldsymbol{t_1},\boldsymbol{t_2},\sigma) \right] \right\} \tag{28}$$

where

$$X(\boldsymbol{t},\sigma) = \mu(\boldsymbol{t},\sigma;\xi,\boldsymbol{t_0},\sigma_0) + W^*(\boldsymbol{t},\sigma), \tag{29}$$

and $R(\boldsymbol{t_1},\boldsymbol{t_2},\sigma)$ is the correlation matrix as the one in Equation (19). As mentioned in the previous subsection, the kernel width was searched over the interval of $[log(0.02), log(0.2)]$ and ten kernels were selected where they were equally spaced on the log scale. These ten numbers are 0.02, 0.0258, 0.0334, 0.0431, 0.0557, 0.0719, 0.0928, 0.1199, 0.1549, 0.2. One reason to have only ten kernels out of this range was to reduce computation time. The other reason is that Worsley et al. (1996)

mentioned the scale dimension in the scale space field is usually pretty smooth; thus only a few samples are required.

The empirical distribution of Y_{max} under the null hypothesis was generated using Equation (28). The simulation procedure was basically the same as before. Nevertheless, because the kernel was not fixed, in each replication the program looped through all ten σ values and for each σ, Y_{max} was calculated. In the end only the maximum of the ten Y_{max}'s was retained. After 5000 replications, eventually 5000 maxima were generated and an empirical distribution was formed. The 95^{th} and 99^{th} percentile were calculated and served as the critical thresholds at α of .05 and .01, respectively.

To assess the power of scale space Y_{max}, images were generated based on the 36 schemes described in Table 1. Using Equation (28), the same simulation procedure illustrated in the previous paragraph was carried out to obtain the values of Y_{max} under the alternative hypothesis and its power was calculated. And to compare the power of Y_{max} with the one of X^2_{max}, Equation (14) was used to obtain the scale space X^2_{max} under the null and alternative hypotheses in simulation.

Finally, though not mentioned in the research questions and really not the focus of this study, X_{max} was also simulated. Equation (13) was used for scale space, and when σ was fixed, the following equation was used,

$$X_{max} = \max_{\boldsymbol{t}} X(\boldsymbol{t}, \sigma). \tag{30}$$

The critical values as well as the power of X_{max} were reported in Chapter IV but just for the readers' information.

CHAPTER IV

RESULTS

This chapter begins with a discussion about the critical values of Y_{max} obtained from the empirical distribution using Monte Carlo simulation. Then it follows with an investigation of the power of Y_{max} under different conditions. The author was also concerned about the behavior of Y_{max} when the width of smoothing kernel, σ, was fixed or not, and the corresponding results were presented respectively. To facilitate the explanation, $Y_{max}95$ and $X^2_{max}95$ are used as the notations for the test statistics Y_{max} and X^2_{max} obtained at $\alpha = .05$, respectively, whereas $Y_{max}99$ and $X^2_{max}99$ are the ones obtained at $\alpha = .01$.

Results for Fixed Smoothing Kernel

As a visualization of the empirical distributions, Figure 6 shows the histograms and density plots of $Y_{max}95$ under different widths of smoothing kernel. The critical values of Y_{max} and X^2_{max} under 10 different σ values at α of .05 and .01, respectively, are displayed in Table 2. The kernel width and the threshold values were negatively correlated that a smaller smoothing parameter resulted in a higher critical value for the test statistics. That is, when applying a narrower kernel width, the threshold was more stringent.

The critical values were then used to calculate the power. As illustrated in Chapter III, there are 36 data schemes manipulating different factors of two simulated signals, including distance, amplitude, and scale. Because the width of smoothing kernel was fixed, all 36 schemes were coupled with one of the 10 σ values, resulting in 360 conditions. The power of Y_{max} as well as X^2_{max} were reported in a table format, see Table 3 - Table 22. To further explore the effect of distance, amplitude, and scale on the power of Y_{max}, line plots were drawn as visual examples for certain conditions. More details about the results are presented in the following sections. Given that at $\alpha = .01$ the power was of course smaller than at $\alpha = .05$, to avoid lengthy discussion, the following results are only focused on the case at $\alpha = .05$.

Table 2

Critical Values When σ Is Fixed

	$\alpha = .05$			$\alpha = .01$		
σ	Y_{max}	X^2_{max}	X_{max}	Y_{max}	X^2_{max}	X_{max}
0.02	31.49	18.19	4.08	35.65	21.35	4.49
0.0258	30.03	17.78	4.00	34.45	20.82	4.37
0.0334	28.56	16.74	3.88	32.69	19.88	4.32
0.0431	26.73	15.48	3.77	31.14	19.15	4.16
0.0557	25.22	14.55	3.65	30.49	18.12	4.09
0.0719	22.97	13.91	3.46	28.43	17.31	3.88
0.0928	21.19	12.65	3.30	26.20	15.99	3.84
0.1199	19.48	11.30	3.18	26.00	14.70	3.64
0.1549	18.01	10.16	2.89	23.60	13.80	3.45
0.2	16.27	9.22	2.75	22.67	12.88	3.37

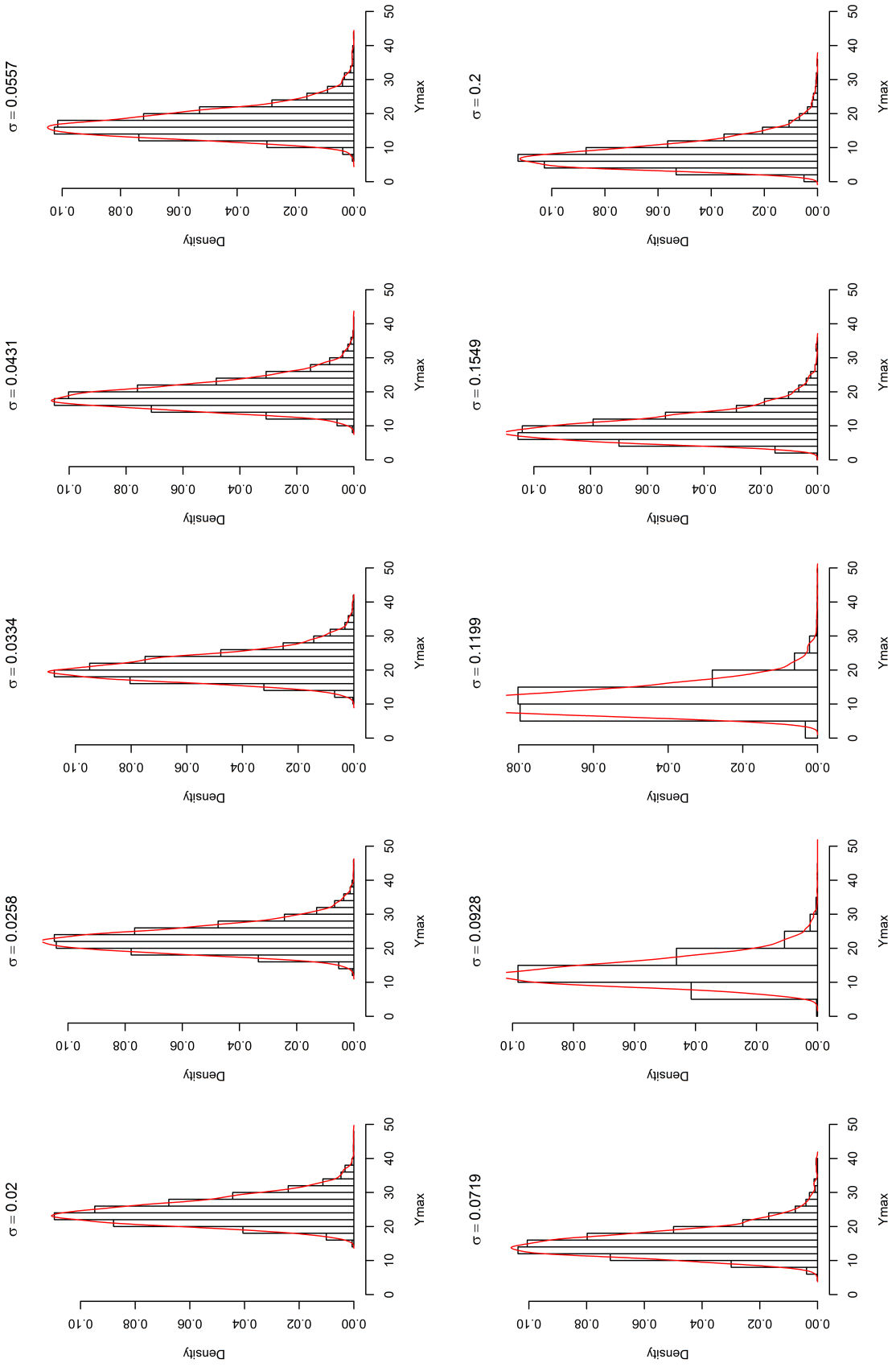


Figure 6. The empirical distribution of Y_{max} for different σ values at $\alpha = .05$

Table 3

Power Table When $\sigma = 0.02$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0546**	0.0526	0.0538
		Middle	0.0502	**0.0536**	0.0466
		Far	0.0534	**0.0556**	0.0578
	(0.5,2)	Near	**0.0774**	0.0684	0.0930
		Middle	**0.0790**	0.0712	0.0882
		Far	**0.0770**	0.0710	0.0866
	(2,2)	Near	**0.1010**	0.0938	0.1296
		Middle	**0.0974**	0.0894	0.1148
		Far	**0.0974**	0.0944	0.1200
	(0.5,4)	Near	0.4740	**0.4898**	0.5610
		Middle	0.4616	**0.4764**	0.5406
		Far	0.4750	**0.4840**	0.5718
	(2,4)	Near	0.4792	**0.4978**	0.5680
		Middle	**0.5062**	0.5048	0.5790
		Far	0.4876	**0.5010**	0.5764
	(4,4)	Near	**0.7590**	0.7258	0.8056
		Middle	**0.7584**	0.7188	0.8090
		Far	**0.7432**	0.7290	0.7918
0.04	(0.5,0.5)	Near	**0.0580**	0.0468	0.0600
		Middle	0.0506	**0.0528**	0.0556
		Far	0.0470	**0.0474**	0.0506
	(0.5,2)	Near	0.0696	**0.0720**	0.0816
		Middle	**0.0690**	0.0638	0.0860
		Far	**0.0752**	0.0712	0.0826
	(2,2)	Near	0.0868	**0.0874**	0.1164
		Middle	**0.0990**	0.0836	0.1166
		Far	**0.1006**	0.0912	0.1234
	(0.5,4)	Near	**0.3482**	0.3370	0.4318
		Middle	0.3436	**0.3538**	0.4218
		Far	**0.3586**	0.3446	0.4332
	(2,4)	Near	**0.3686**	0.3524	0.4390
		Middle	**0.3654**	0.3516	0.4584
		Far	**0.3788**	0.3474	0.4586
	(4,4)	Near	**0.6006**	0.5570	0.6570
		Middle	**0.5982**	0.5480	0.6468
		Far	**0.6030**	0.5554	0.6534

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 4

Power Table When $\sigma = 0.02$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0116	**0.0130**	0.0096
		Middle	**0.0104**	0.0090	0.0104
		Far	**0.0108**	0.0100	0.0096
	(0.5,2)	Near	**0.0182**	0.0180	0.0216
		Middle	0.0166	**0.0196**	0.0220
		Far	**0.0202**	0.0182	0.0264
	(2,2)	Near	**0.0282**	0.0258	0.0378
		Middle	**0.0306**	0.0236	0.0370
		Far	0.0264	0.0264	0.0370
	(0.5,4)	Near	0.2830	**0.3194**	0.3660
		Middle	0.2752	**0.3172**	0.3642
		Far	0.2902	**0.3200**	0.3712
	(2,4)	Near	0.2934	**0.3292**	0.3718
		Middle	0.3038	**0.3424**	0.3794
		Far	0.3036	**0.3288**	0.3758
	(4,4)	Near	**0.5936**	0.5380	0.6186
		Middle	**0.5848**	0.5302	0.6080
		Far	**0.5796**	0.5540	0.6036
0.04	(0.5,0.5)	Near	**0.0116**	0.0100	0.0116
		Middle	**0.0114**	0.0110	0.0116
		Far	**0.0096**	0.0090	0.0084
	(0.5,2)	Near	0.0160	**0.0186**	0.0192
		Middle	**0.0178**	0.0156	0.0218
		Far	0.0154	**0.0190**	0.0176
	(2,2)	Near	0.0218	**0.0226**	0.0318
		Middle	**0.0284**	0.0236	0.0372
		Far	**0.0288**	0.0234	0.0328
	(0.5,4)	Near	0.1724	**0.1910**	0.2378
		Middle	0.1762	**0.2004**	0.2344
		Far	0.1772	**0.1834**	0.2390
	(2,4)	Near	**0.1900**	0.1890	0.2368
		Middle	0.1822	**0.1880**	0.2396
		Far	**0.1928**	0.1922	0.2410
	(4,4)	Near	**0.3812**	0.3444	0.4024
		Middle	**0.3820**	0.3316	0.4008
		Far	**0.3858**	0.3368	0.4148

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 5

Power Table When $\sigma = 0.0258$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0588**	0.0470	0.0536
		Middle	**0.0598**	0.0386	0.0536
		Far	**0.051**4	0.0428	0.0482
	(0.5,2)	Near	**0.0814**	0.0590	0.0946
		Middle	**0.0750**	0.0658	0.0844
		Far	**0.0754**	0.0668	0.0884
	(2,2)	Near	**0.1166**	0.0892	0.1362
		Middle	**0.0968**	0.0834	0.1140
		Far	**0.1036**	0.0808	0.1202
	(0.5,4)	Near	0.4336	**0.4562**	0.5162
		Middle	**0.4390**	0.4384	0.5196
		Far	**0.4580**	0.4550	0.5402
	(2,4)	Near	**0.4888**	0.4710	0.5690
		Middle	**0.4712**	0.4634	0.5484
		Far	**0.4692**	0.4628	0.5520
	(4,4)	Near	**0.7332**	0.7150	0.7896
		Middle	**0.7308**	0.6832	0.7716
		Far	**0.7218**	0.6694	0.7670
0.04	(0.5,0.5)	Near	**0.0548**	0.0456	0.0526
		Middle	**0.0566**	0.0436	0.0520
		Far	**0.0594**	0.0420	0.0594
	(0.5,2)	Near	**0.0826**	0.0668	0.0982
		Middle	**0.0798**	0.0676	0.0968
		Far	**0.0858**	0.0626	0.1018
	(2,2)	Near	**0.1178**	0.0954	0.1410
		Middle	**0.1144**	0.1026	0.1424
		Far	**0.1234**	0.1004	0.1422
	(0.5,4)	Near	0.4644	**0.4730**	0.5574
		Middle	0.4542	**0.4558**	0.5542
		Far	**0.4732**	0.4580	0.5616
	(2,4)	Near	**0.5088**	0.4916	0.5898
		Middle	**0.5044**	0.4918	0.5784
		Far	**0.4958**	0.4810	0.5750
	(4,4)	Near	**0.7678**	0.6926	0.8086
		Middle	**0.7608**	0.7000	0.7914
		Far	**0.7548**	0.7086	0.7954

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 6

Power Table When $\sigma = 0.0258$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	**0.0124**	0.0082	0.0144
		Middle	**0.0112**	0.0108	0.0126
		Far	0.0102	**0.0112**	0.0122
	(0.5,2)	Near	**0.0190**	0.0178	0.0286
		Middle	**0.0232**	0.0178	0.0272
		Far	0.0150	**0.0220**	0.0226
	(2,2)	Near	0.0292	**0.0310**	0.0472
		Middle	0.0252	**0.0278**	0.0368
		Far	**0.0300**	0.0280	0.0412
	(0.5,4)	Near	0.2494	**0.2940**	0.3522
		Middle	0.2484	**0.2834**	0.3610
		Far	0.2630	**0.2998**	0.3768
	(2,4)	Near	0.2850	**0.3084**	0.4006
		Middle	0.2724	**0.3024**	0.3732
		Far	0.2802	**0.3032**	0.3718
	(4,4)	Near	**0.5398**	0.5302	0.6200
		Middle	**0.5372**	0.5050	0.5914
		Far	**0.5306**	0.4852	0.5904
0.04	(0.5,0.5)	Near	0.0100	**0.0114**	0.0140
		Middle	**0.0112**	0.0092	0.0104
		Far	**0.0100**	0.0092	0.0156
	(0.5,2)	Near	**0.0198**	0.0194	0.0316
		Middle	0.0218	**0.0224**	0.0304
		Far	**0.0230**	0.0210	0.0318
	(2,2)	Near	**0.0320**	0.0280	0.0494
		Middle	**0.0322**	0.0304	0.0468
		Far	**0.0340**	0.0314	0.0504
	(0.5,4)	Near	0.2658	**0.3120**	0.3736
		Middle	0.2538	**0.2990**	0.3632
		Far	0.2610	**0.2930**	0.3760
	(2,4)	Near	0.3050	**0.3172**	0.3996
		Middle	0.2974	**0.3158**	0.3928
		Far	0.2872	**0.3160**	0.3844
	(4,4)	Near	**0.5686**	0.5058	0.6162
		Middle	**0.5508**	0.5028	0.6018
		Far	**0.5604**	0.5004	0.6128

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 7

Power Table When $\sigma = 0.0334$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0518**	0.0480	0.0510
		Middle	**0.0460**	0.0458	0.0522
		Far	**0.0510**	0.0480	0.0580
	(0.5,2)	Near	**0.0694**	0.0612	0.0866
		Middle	**0.0714**	0.0652	0.0852
		Far	**0.0688**	0.0670	0.0830
	(2,2)	Near	**0.1050**	0.0962	0.1388
		Middle	**0.0890**	0.0828	0.1150
		Far	**0.0912**	0.0846	0.1166
	(0.5,4)	Near	0.3750	**0.3794**	0.4588
		Middle	0.3602	**0.3714**	0.4572
		Far	0.3540	**0.3628**	0.4528
	(2,4)	Near	0.4228	**0.4382**	0.5268
		Middle	**0.3922**	0.3766	0.4792
		Far	**0.3778**	0.3714	0.4702
	(4,4)	Near	0.6644	**0.6938**	0.7686
		Middle	**0.6100**	0.5788	0.6810
		Far	**0.6128**	0.5926	0.6812
0.04	(0.5,0.5)	Near	**0.0508**	0.0472	0.0528
		Middle	**0.0492**	0.0410	0.0568
		Far	**0.0558**	0.0500	0.0584
	(0.5,2)	Near	**0.0952**	0.0856	0.1224
		Middle	**0.0898**	0.0836	0.1214
		Far	**0.0986**	0.0868	0.1302
	(2,2)	Near	**0.1384**	0.1302	0.1838
		Middle	**0.1322**	0.1222	0.1774
		Far	**0.1432**	0.1266	0.1820
	(0.5,4)	Near	0.5600	**0.5694**	0.6694
		Middle	0.5598	**0.5812**	0.6794
		Far	0.5568	**0.5846**	0.6668
	(2,4)	Near	0.5946	**0.5986**	0.6894
		Middle	0.5902	**0.5922**	0.6890
		Far	0.5796	**0.5960**	0.6836
	(4,4)	Near	**0.8394**	0.8180	0.8828
		Middle	**0.8346**	0.8148	0.8752
		Far	**0.8472**	0.8100	0.8880

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 8

Power Table When $\sigma = 0.0334$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	**0.0154**	0.0128	0.0104
		Middle	**0.0112**	0.0094	0.0096
		Far	**0.0128**	0.0122	0.0104
	(0.5,2)	Near	**0.0184**	0.0162	0.0184
		Middle	0.0190	**0.0196**	0.0186
		Far	0.0166	**0.0184**	0.0184
	(2,2)	Near	**0.0328**	0.0318	0.0410
		Middle	0.0234	**0.0242**	0.0292
		Far	**0.0266**	0.0250	0.0302
	(0.5,4)	Near	0.2030	**0.2238**	0.2706
		Middle	0.2014	**0.2240**	0.2628
		Far	0.1958	**0.2264**	0.2606
	(2,4)	Near	0.2470	**0.2768**	0.3220
		Middle	0.2130	**0.2228**	0.2818
		Far	0.2126	**0.2278**	0.2696
	(4,4)	Near	0.4776	**0.5178**	0.5678
		Middle	**0.4184**	0.3834	0.4530
		Far	**0.4264**	0.3910	0.4538
0.04	(0.5,0.5)	Near	0.0104	**0.0108**	0.0102
		Middle	0.0074	**0.0082**	0.0114
		Far	0.0118	**0.0126**	0.0130
	(0.5,2)	Near	**0.0276**	0.0262	0.0344
		Middle	0.0254	**0.0268**	0.0332
		Far	0.0284	**0.0300**	0.0342
	(2,2)	Near	0.0472	**0.0486**	0.0664
		Middle	**0.0438**	0.0412	0.0510
		Far	**0.0474**	0.0422	0.0624
	(0.5,4)	Near	0.3622	**0.4032**	0.4734
		Middle	0.3572	**0.4110**	0.4682
		Far	0.3616	**0.4114**	0.4660
	(2,4)	Near	0.3954	**0.4210**	0.4774
		Middle	0.3896	**0.4198**	0.4744
		Far	0.3870	0.4146	0.4730
	(4,4)	Near	**0.6966**	0.6588	0.7072
		Middle	**0.6896**	0.6444	0.7010
		Far	**0.7088**	0.6492	0.7188

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 9

Power Table When $\sigma = 0.0431$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	0.0520	**0.0590**	0.0536
		Middle	0.0518	**0.0582**	0.0574
		Far	0.0500	**0.0552**	0.0494
	(0.5,2)	Near	0.0652	**0.0738**	0.0802
		Middle	**0.0712**	0.0686	0.0808
		Far	0.0672	**0.0710**	0.0752
	(2,2)	Near	0.1064	**0.1144**	0.1390
		Middle	0.0774	**0.0868**	0.0954
		Far	0.0834	**0.0850**	0.0970
	(0.5,4)	Near	0.2732	**0.2972**	0.3410
		Middle	0.2650	**0.2786**	0.3344
		Far	0.2542	**0.2788**	0.3224
	(2,4)	Near	0.3634	**0.3984**	0.4582
		Middle	0.2618	**0.2918**	0.3290
		Far	0.2682	**0.2918**	0.3246
	(4,4)	Near	0.6282	**0.6706**	0.7340
		Middle	**0.4612**	0.4536	0.5278
		Far	**0.4622**	0.4468	0.5266
0.04	(0.5,0.5)	Near	0.0520	**0.0576**	0.0554
		Middle	0.0512	**0.0572**	0.0586
		Far	0.0490	**0.0578**	0.0532
	(0.5,2)	Near	0.0998	**0.1130**	0.1266
		Middle	0.0948	**0.0994**	0.1270
		Far	0.1026	**0.1066**	0.1228
	(2,2)	Near	0.1492	**0.1648**	0.1968
		Middle	0.1474	**0.1578**	0.1924
		Far	0.1552	**0.1560**	0.1938
	(0.5,4)	Near	0.5804	**0.6266**	0.6942
		Middle	0.5744	**0.6336**	0.6976
		Far	0.5796	**0.6422**	0.6892
	(2,4)	Near	0.6160	**0.6584**	0.7132
		Middle	0.6138	**0.6690**	0.7028
		Far	0.6110	**0.6514**	0.7092
	(4,4)	Near	**0.8726**	0.8690	0.9060
		Middle	**0.8798**	0.8550	0.9068
		Far	**0.8706**	0.8586	0.9054

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 10

Power Table When $\sigma = 0.0431$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0106	**0.0108**	0.0136
		Middle	**0.0136**	0.0100	0.0158
		Far	**0.0108**	0.0092	0.0134
	(0.5,2)	Near	**0.0168**	0.0150	0.0236
		Middle	**0.0196**	0.0156	0.0266
		Far	**0.0174**	0.0154	0.0258
	(2,2)	Near	0.0308	**0.0358**	0.0544
		Middle	**0.0176**	0.0170	0.0288
		Far	**0.0218**	0.0202	0.0322
	(0.5,4)	Near	0.1280	**0.1424**	0.1948
		Middle	0.1188	**0.1302**	0.1896
		Far	0.1144	**0.1270**	0.1830
	(2,4)	Near	0.1768	**0.2116**	0.2848
		Middle	0.1190	**0.1392**	0.1856
		Far	0.1210	**0.1438**	0.1818
	(4,4)	Near	0.4076	**0.4642**	0.5706
		Middle	**0.2578**	0.2334	0.3236
		Far	**0.2604**	0.2376	0.3278
0.04	(0.5,0.5)	Near	**0.0132**	0.0112	0.0144
		Middle	**0.0116**	0.0088	0.0126
		Far	**0.0104**	0.0092	0.0102
	(0.5,2)	Near	0.0300	**0.0322**	0.0474
		Middle	0.0250	**0.0254**	0.0436
		Far	**0.0320**	0.0312	0.0472
	(2,2)	Near	**0.0496**	0.0482	0.0756
		Middle	**0.0494**	0.0484	0.0774
		Far	**0.0510**	0.0498	0.0804
	(0.5,4)	Near	0.3710	**0.4264**	0.5240
		Middle	0.3812	**0.4316**	0.5238
		Far	0.3698	**0.4454**	0.5218
	(2,4)	Near	0.4268	**0.4418**	0.5424
		Middle	0.4108	**0.4638**	0.5356
		Far	0.4062	**0.4432**	0.5330
	(4,4)	Near	**0.7262**	0.6920	0.7804
		Middle	**0.7438**	0.6808	0.7814
		Far	**0.7314**	0.6760	0.7778

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 11

Power Table When $\sigma = 0.0557$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	0.0450	**0.0566**	0.0492
		Middle	0.0456	**0.0560**	0.0474
		Far	0.0464	**0.0492**	0.0490
	(0.5,2)	Near	0.0680	**0.0690**	0.0778
		Middle	0.0564	**0.0618**	0.0676
		Far	0.0542	**0.0662**	0.0634
	(2,2)	Near	0.0856	**0.1118**	0.1210
		Middle	0.0618	**0.0734**	0.0804
		Far	0.0692	**0.0804**	0.0790
	(0.5,4)	Near	0.1866	**0.2140**	0.2640
		Middle	0.1694	**0.1840**	0.2280
		Far	0.1574	**0.1808**	0.2152
	(2,4)	Near	0.2786	**0.3266**	0.3788
		Middle	0.1626	**0.1906**	0.2196
		Far	0.1656	**0.1876**	0.2342
	(4,4)	Near	0.5576	**0.6120**	0.6794
		Middle	0.2906	**0.3022**	0.3566
		Far	0.2882	**0.2912**	0.3586
0.04	(0.5,0.5)	Near	0.0486	**0.0562**	0.0522
		Middle	0.0438	**0.0574**	0.0524
		Far	0.0474	**0.0548**	0.0548
	(0.5,2)	Near	0.0820	**0.0992**	0.1218
		Middle	0.0820	**0.1040**	0.1194
		Far	0.0914	**0.0972**	0.1216
	(2,2)	Near	0.1474	**0.1666**	0.2078
		Middle	0.1438	**0.1600**	0.1930
		Far	0.1394	**0.1442**	0.1942
	(0.5,4)	Near	0.5278	**0.6120**	0.6498
		Middle	0.5194	**0.5946**	0.6526
		Far	0.5200	**0.5838**	0.6450
	(2,4)	Near	0.5800	**0.6320**	0.6980
		Middle	0.5722	**0.6088**	0.6836
		Far	0.5702	**0.6050**	0.6846
	(4,4)	Near	0.8340	**0.8486**	0.8900
		Middle	**0.8334**	0.8162	0.8816
		Far	0.8224	**0.8240**	0.8722

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 12

Power Table When $\sigma = 0.0557$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0076	**0.0094**	0.0108
		Middle	0.0080	**0.0104**	0.0096
		Far	**0.0092**	0.0074	0.0144
	(0.5,2)	Near	0.0136	**0.0142**	0.0202
		Middle	0.0098	**0.0116**	0.0170
		Far	0.0112	**0.0134**	0.0154
	(2,2)	Near	0.0190	**0.0296**	0.0418
		Middle	0.0138	**0.0168**	0.0230
		Far	0.0130	**0.0174**	0.0184
	(0.5,4)	Near	0.0610	**0.0862**	0.1238
		Middle	0.0540	**0.0738**	0.1068
		Far	0.0460	**0.0754**	0.0898
	(2,4)	Near	0.1146	**0.1678**	0.2096
		Middle	0.0488	**0.0758**	0.0954
		Far	0.0514	**0.0680**	0.1006
	(4,4)	Near	0.3106	**0.4184**	0.4896
		Middle	0.1106	**0.1352**	0.1782
		Far	0.1186	**0.1278**	0.1744
0.04	(0.5,0.5)	Near	0.0080	**0.0112**	0.0106
		Middle	0.0068	**0.0108**	0.0092
		Far	0.0080	**0.0112**	0.0106
	(0.5,2)	Near	0.0182	**0.0294**	0.0400
		Middle	0.0190	**0.0270**	0.0394
		Far	0.0250	**0.0276**	0.0440
	(2,2)	Near	0.0378	**0.0512**	0.0786
		Middle	0.0352	**0.0514**	0.0736
		Far	0.0368	**0.0440**	0.0716
	(0.5,4)	Near	0.2990	**0.4104**	0.4676
		Middle	0.2816	**0.3948**	0.4604
		Far	0.2826	**0.3932**	0.4618
	(2,4)	Near	0.3342	**0.4296**	0.5064
		Middle	0.3250	**0.4062**	0.4836
		Far	0.3314	**0.3990**	0.4824
	(4,4)	Near	0.6344	**0.6726**	0.7438
		Middle	**0.6470**	0.6220	0.7200
		Far	**0.6362**	0.6176	0.7058

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 13

Power Table When $\sigma = 0.0719$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0610**	0.0400	0.0642
		Middle	**0.0560**	0.0404	0.0580
		Far	**0.0516**	0.0398	0.0540
	(0.5,2)	Near	**0.0628**	0.0562	0.0768
		Middle	**0.0606**	0.0532	0.0728
		Far	0.0546	**0.0560**	0.0604
	(2,2)	Near	**0.0916**	0.0746	0.1274
		Middle	**0.0670**	0.0590	0.0856
		Far	**0.0720**	0.0584	0.0854
	(0.5,4)	Near	**0.1416**	0.1312	0.1954
		Middle	**0.1198**	0.1108	0.1722
		Far	**0.1252**	0.1072	0.1660
	(2,4)	Near	**0.2314**	0.2278	0.3238
		Middle	**0.1368**	0.1160	0.1858
		Far	**0.1354**	0.1226	0.1746
	(4,4)	Near	0.4726	**0.4772**	0.5928
		Middle	**0.2058**	0.1726	0.2672
		Far	**0.2042**	0.1640	0.2568
0.04	(0.5,0.5)	Near	**0.0582**	0.0440	0.0590
		Middle	**0.0544**	0.0458	0.0620
		Far	**0.0540**	0.0458	0.0584
	(0.5,2)	Near	0.0920	0.0920	0.1322
		Middle	**0.0946**	0.0890	0.1274
		Far	**0.0912**	0.0738	0.1246
	(2,2)	Near	**0.1524**	0.1476	0.2276
		Middle	**0.1322**	0.1204	0.1874
		Far	**0.1258**	0.1184	0.1820
	(0.5,4)	Near	0.4674	**0.4836**	0.5844
		Middle	0.4652	**0.4792**	0.5842
		Far	0.4566	**0.4634**	0.5708
	(2,4)	Near	0.5274	**0.5498**	0.6562
		Middle	**0.4910**	0.4888	0.6026
		Far	**0.4990**	0.4806	0.6024
	(4,4)	Near	**0.7764**	0.7678	0.8694
		Middle	**0.7516**	0.6866	0.8150
		Far	**0.7506**	0.6930	0.8084

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 14

Power Table When $\sigma = 0.0719$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0116**	0.0086	0.0166
		Middle	**0.0108**	0.0082	0.0176
		Far	**0.0090**	0.0064	0.0112
	(0.5,2)	Near	0.0110	**0.0132**	0.0214
		Middle	0.0100	**0.0122**	0.0182
		Far	0.0094	**0.0134**	0.0150
	(2,2)	Near	**0.0200**	0.0198	0.0456
		Middle	**0.0138**	0.0128	0.0242
		Far	**0.0168**	0.0138	0.0252
	(0.5,4)	Near	0.0398	**0.0498**	0.0860
		Middle	0.0276	**0.0364**	0.0660
		Far	0.0276	**0.0326**	0.0682
	(2,4)	Near	0.0820	**0.1030**	0.1754
		Middle	0.0352	**0.0418**	0.0726
		Far	0.0346	**0.0414**	0.0706
	(4,4)	Near	0.2350	**0.2988**	0.4120
		Middle	**0.0634**	0.0598	0.1214
		Far	**0.0670**	0.0610	0.1156
0.04	(0.5,0.5)	Near	0.0096	**0.0104**	0.0144
		Middle	**0.0116**	0.0088	0.0164
		Far	0.0104	**0.0120**	0.0146
	(0.5,2)	Near	0.0218	**0.0276**	0.0502
		Middle	0.0194	**0.0238**	0.0434
		Far	0.0214	**0.0222**	0.0438
	(2,2)	Near	0.0386	**0.0548**	0.0980
		Middle	0.0322	**0.0426**	0.0730
		Far	0.0316	**0.0376**	0.0656
	(0.5,4)	Near	0.2348	**0.3062**	0.3994
		Middle	0.2246	**0.2966**	0.4028
		Far	0.2252	**0.2898**	0.3954
	(2,4)	Near	0.2752	**0.3486**	0.4742
		Middle	0.2434	**0.2948**	0.4034
		Far	0.2506	**0.2892**	0.4082
	(4,4)	Near	0.5218	**0.5896**	0.7234
		Middle	**0.4976**	0.4848	0.6258
		Far	**0.4976**	0.4746	0.6290

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 15

Power Table When $\sigma = 0.0928$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0522**	0.0492	0.0600
		Middle	**0.0550**	0.0484	0.0554
		Far	**0.0532**	0.0488	0.0540
	(0.5,2)	Near	**0.0578**	0.0500	0.0744
		Middle	**0.0616**	0.0554	0.0700
		Far	**0.0578**	0.0542	0.0676
	(2,2)	Near	**0.0866**	0.0798	0.1160
		Middle	0.0582	**0.0588**	0.0768
		Far	**0.0624**	0.0528	0.0744
	(0.5,4)	Near	0.1010	**0.1048**	0.1472
		Middle	**0.0934**	0.0810	0.1240
		Far	**0.0958**	0.0878	0.1298
	(2,4)	Near	0.1708	**0.1822**	0.2524
		Middle	**0.1066**	0.0960	0.1392
		Far	**0.0978**	0.0874	0.1280
	(4,4)	Near	0.3558	**0.3752**	0.4570
		Middle	**0.1394**	0.1356	0.1980
		Far	**0.1424**	0.1238	0.1808
0.04	(0.5,0.5)	Near	**0.0570**	0.0478	0.0600
		Middle	0.0490	**0.0494**	0.0534
		Far	**0.0558**	0.0488	0.0550
	(0.5,2)	Near	**0.0886**	0.0844	0.1268
		Middle	**0.0860**	0.0818	0.1198
		Far	**0.0852**	0.0826	0.1022
	(2,2)	Near	0.1436	**0.1496**	0.2216
		Middle	**0.1040**	0.1026	0.1580
		Far	**0.1148**	0.1114	0.1586
	(0.5,4)	Near	0.3542	**0.3766**	0.4756
		Middle	0.3390	**0.3526**	0.4530
		Far	0.3362	**0.3500**	0.4460
	(2,4)	Near	0.4414	**0.5022**	0.6008
		Middle	**0.3838**	0.3798	0.4910
		Far	0.3804	**0.3890**	0.4946
	(4,4)	Near	0.7008	**0.7532**	0.8268
		Middle	**0.6066**	0.5652	0.6930
		Far	**0.5824**	0.5542	0.6572

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 16

Power Table When $\sigma = 0.0928$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	**0.0120**	0.0086	0.0102
		Middle	0.0130	**0.0132**	0.0100
		Far	0.0104	**0.0106**	0.0084
	(0.5,2)	Near	**0.0142**	0.0130	0.0162
		Middle	**0.0146**	0.0140	0.0152
		Far	**0.0148**	0.0144	0.0138
	(2,2)	Near	0.0182	**0.0216**	0.0302
		Middle	**0.0126**	0.0122	0.0148
		Far	**0.0164**	0.0134	0.0154
	(0.5,4)	Near	0.0292	**0.0332**	0.0410
		Middle	**0.0270**	0.0222	0.0388
		Far	0.0264	**0.0268**	0.0340
	(2,4)	Near	0.0510	**0.0706**	0.0956
		Middle	0.0288	**0.0290**	0.0444
		Far	**0.0286**	0.0242	0.0358
	(4,4)	Near	0.1594	**0.2048**	0.2468
		Middle	0.0424	**0.0442**	0.0618
		Far	**0.0470**	0.0364	0.0552
0.04	(0.5,0.5)	Near	**0.0124**	0.0086	0.0108
		Middle	**0.0134**	0.0102	0.0098
		Far	**0.0152**	0.0104	0.0094
	(0.5,2)	Near	0.0214	**0.0240**	0.0346
		Middle	0.0204	**0.0234**	0.0322
		Far	0.0216	**0.0236**	0.0294
	(2,2)	Near	0.0394	**0.0554**	0.0784
		Middle	0.0280	**0.0348**	0.0462
		Far	0.0300	**0.0346**	0.0444
	(0.5,4)	Near	0.1576	**0.2144**	0.2576
		Middle	0.1578	**0.1926**	0.2400
		Far	0.1452	**0.1860**	0.2416
	(2,4)	Near	0.2148	**0.3102**	0.3576
		Middle	0.1712	**0.2034**	0.2644
		Far	0.1716	**0.2042**	0.2560
	(4,4)	Near	0.4548	**0.5674**	0.6352
		Middle	0.3406	**0.3410**	0.4226
		Far	**0.3338**	0.3322	0.3930

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 17

Power Table When $\sigma = 0.1199$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	0.0500	**0.0564**	0.0442
		Middle	0.0506	**0.0528**	0.0456
		Far	**0.0522**	0.0514	0.0454
	(0.5,2)	Near	0.0566	**0.0614**	0.0602
		Middle	0.0514	**0.0620**	0.0588
		Far	**0.0642**	0.0570	0.0582
	(2,2)	Near	0.0706	**0.0820**	0.0884
		Middle	0.0620	**0.0656**	0.0722
		Far	0.0656	**0.0658**	0.0674
	(0.5,4)	Near	0.0872	**0.0970**	0.1088
		Middle	0.0820	**0.0838**	0.0966
		Far	0.0802	**0.0804**	0.0940
	(2,4)	Near	0.1344	**0.1434**	0.1792
		Middle	0.0830	**0.0940**	0.1114
		Far	0.0798	**0.0906**	0.0906
	(4,4)	Near	0.2414	**0.2682**	0.3244
		Middle	0.1058	**0.1198**	0.1468
		Far	0.1076	**0.1094**	0.1194
0.04	(0.5,0.5)	Near	**0.0604**	0.0602	0.0564
		Middle	**0.0568**	0.0470	0.0568
		Far	0.0544	**0.0564**	0.0520
	(0.5,2)	Near	0.0808	**0.0942**	0.1054
		Middle	**0.0786**	0.0746	0.0982
		Far	0.0772	**0.0830**	0.0942
	(2,2)	Near	0.1324	**0.1574**	0.2006
		Middle	0.0942	**0.1080**	0.1276
		Far	**0.0998**	0.0972	0.1226
	(0.5,4)	Near	0.2692	**0.2942**	0.3558
		Middle	0.2368	**0.2608**	0.3160
		Far	0.2390	**0.2626**	0.3196
	(2,4)	Near	0.3726	**0.4328**	0.5036
		Middle	0.2698	**0.2794**	0.3500
		Far	0.2540	**0.2720**	0.3242
	(4,4)	Near	0.6518	**0.7146**	0.7764
		Middle	0.4272	**0.4374**	0.5142
		Far	**0.4080**	0.3996	0.4686

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 18

Power Table When $\sigma = 0.1199$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0074	**0.0078**	0.0106
		Middle	0.0076	**0.0134**	0.0130
		Far	0.0072	**0.0124**	0.0088
	(0.5,2)	Near	0.0096	**0.0164**	0.0162
		Middle	0.0080	**0.0124**	0.0136
		Far	**0.0108**	0.0096	0.0156
	(2,2)	Near	0.0122	**0.0240**	0.0282
		Middle	0.0082	**0.0152**	0.0186
		Far	0.0118	**0.0152**	0.0200
	(0.5,4)	Near	0.0152	**0.0266**	0.0370
		Middle	0.0130	**0.0250**	0.0294
		Far	0.0146	**0.0198**	0.0292
	(2,4)	Near	0.0238	**0.0470**	0.0754
		Middle	0.0134	**0.0250**	0.0350
		Far	0.0170	**0.0238**	0.0300
	(4,4)	Near	0.0612	**0.1146**	0.1702
		Middle	0.0200	**0.0356**	0.0496
		Far	0.0228	**0.0346**	0.0388
0.04	(0.5,0.5)	Near	0.0108	**0.0118**	0.0174
		Middle	**0.0104**	0.0090	0.0174
		Far	0.0086	**0.0128**	0.0126
	(0.5,2)	Near	0.0112	**0.0280**	0.0364
		Middle	0.0146	**0.0194**	0.0286
		Far	0.0142	**0.0230**	0.0296
	(2,2)	Near	0.0290	**0.0522**	0.0806
		Middle	0.0152	**0.0318**	0.0474
		Far	0.0238	**0.0256**	0.0436
	(0.5,4)	Near	0.0754	**0.1426**	0.1896
		Middle	0.0624	**0.1186**	0.1592
		Far	0.0618	**0.1232**	0.1658
	(2,4)	Near	0.1266	**0.2466**	0.3056
		Middle	0.0686	**0.1224**	0.1724
		Far	0.0712	**0.1242**	0.1614
	(4,4)	Near	0.3452	**0.5202**	0.5992
		Middle	0.1460	**0.2208**	0.2872
		Far	0.1596	**0.1962**	0.2640

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 19

Power Table When $\sigma = 0.1549$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	0.0488	**0.0556**	0.0642
		Middle	**0.0540**	0.0522	0.0638
		Far	0.0476	**0.0544**	0.0608
	(0.5,2)	Near	**0.0612**	0.0600	0.0850
		Middle	0.0530	**0.0562**	0.0774
		Far	0.0562	**0.0628**	0.0696
	(2,2)	Near	0.0640	**0.0752**	0.1062
		Middle	0.0578	**0.0662**	0.0904
		Far	**0.0622**	0.0608	0.0786
	(0.5,4)	Near	0.0726	**0.0890**	0.1248
		Middle	0.0646	**0.0918**	0.1136
		Far	0.0660	**0.0772**	0.1038
	(2,4)	Near	0.0906	**0.1144**	0.1690
		Middle	0.0758	**0.0796**	0.1302
		Far	0.0756	**0.0802**	0.1130
	(4,4)	Near	0.1756	**0.1850**	0.2964
		Middle	0.0806	**0.0992**	0.1590
		Far	0.0928	**0.0938**	0.1370
0.04	(0.5,0.5)	Near	0.0522	**0.0614**	0.0692
		Middle	0.0532	**0.0572**	0.0632
		Far	0.0528	0.0528	0.0698
	(0.5,2)	Near	0.0764	**0.0892**	0.1292
		Middle	0.0676	**0.0724**	0.1140
		Far	0.0664	**0.0788**	0.1052
	(2,2)	Near	0.1036	**0.1486**	0.2168
		Middle	0.0814	**0.0952**	0.1514
		Far	0.0820	**0.0910**	0.1358
	(0.5,4)	Near	0.1936	**0.2282**	0.3194
		Middle	0.1652	**0.1948**	0.2796
		Far	0.1522	**0.1896**	0.2702
	(2,4)	Near	0.2810	**0.3588**	0.4612
		Middle	0.1908	**0.2170**	0.3208
		Far	0.1790	**0.1944**	0.2994
	(4,4)	Near	0.5502	**0.6190**	0.7300
		Middle	0.2850	**0.3164**	0.4538
		Far	0.2642	**0.2786**	0.3844

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 20

Power Table When $\sigma = 0.1549$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0114	**0.0120**	0.0156
		Middle	**0.0132**	0.0088	0.0148
		Far	0.0088	**0.0090**	0.0100
	(0.5,2)	Near	**0.0136**	0.0114	0.0186
		Middle	0.0108	**0.0110**	0.0164
		Far	**0.0118**	0.0098	0.0158
	(2,2)	Near	0.0128	**0.0216**	0.0284
		Middle	0.0132	**0.0134**	0.0218
		Far	0.0138	**0.0140**	0.0204
	(0.5,4)	Near	0.0140	**0.0180**	0.0356
		Middle	0.0104	**0.0188**	0.0290
		Far	0.0160	**0.0174**	0.0246
	(2,4)	Near	0.0218	**0.0350**	0.0562
		Middle	0.0156	**0.0204**	0.0368
		Far	**0.0218**	0.0192	0.0312
	(4,4)	Near	0.0496	**0.0642**	0.1202
		Middle	0.0162	**0.0254**	0.0444
		Far	**0.0268**	0.0250	0.0396
0.04	(0.5,0.5)	Near	0.0110	**0.0126**	0.0148
		Middle	**0.0110**	0.0102	0.0142
		Far	**0.0120**	0.0096	0.0156
	(0.5,2)	Near	0.0182	**0.0236**	0.0362
		Middle	**0.0158**	0.0144	0.0294
		Far	0.0146	**0.0188**	0.0250
	(2,2)	Near	0.0238	**0.0520**	0.0678
		Middle	0.0180	**0.0276**	0.0428
		Far	**0.0232**	0.0208	0.0378
	(0.5,4)	Near	0.0552	**0.0878**	0.1374
		Middle	0.0434	**0.0714**	0.1128
		Far	0.0436	**0.0652**	0.1122
	(2,4)	Near	0.0918	**0.1696**	0.2470
		Middle	0.0540	**0.0792**	0.1294
		Far	0.0582	**0.0704**	0.1188
	(4,4)	Near	0.2818	**0.3988**	0.5088
		Middle	0.0908	**0.1292**	0.2022
		Far	0.0998	**0.1170**	0.1710

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 21

Power Table When $\sigma = 0.2$ *at* $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	**0.0576**	0.0492	0.0618
		Middle	**0.0528**	0.0472	0.0538
		Far	**0.0542**	0.0486	0.0534
	(0.5,2)	Near	**0.0610**	0.0476	0.0702
		Middle	0.0492	**0.0530**	0.0614
		Far	**0.0582**	0.0542	0.0628
	(2,2)	Near	0.0672	**0.0698**	0.0966
		Middle	**0.0630**	0.0578	0.0862
		Far	0.0570	**0.0620**	0.0714
	(0.5,4)	Near	0.0678	**0.0736**	0.1078
		Middle	**0.0724**	0.0712	0.1002
		Far	0.0598	**0.0652**	0.0774
	(2,4)	Near	0.0860	**0.0910**	0.1386
		Middle	**0.0772**	0.0766	0.1148
		Far	**0.0710**	0.0646	0.0878
	(4,4)	Near	0.1330	**0.1436**	0.2164
		Middle	0.0818	**0.0980**	0.1380
		Far	**0.0828**	0.0736	0.1096
0.04	(0.5,0.5)	Near	0.0538	**0.0554**	0.0646
		Middle	**0.0570**	0.0486	0.0640
		Far	0.0542	**0.0546**	0.0558
	(0.5,2)	Near	**0.0744**	0.0742	0.1068
		Middle	**0.0702**	0.0638	0.1048
		Far	0.0674	0.0674	0.0864
	(2,2)	Near	0.1028	**0.1158**	0.1724
		Middle	0.0818	**0.0856**	0.1376
		Far	**0.0824**	0.0796	0.1050
	(0.5,4)	Near	0.1576	**0.1616**	0.2386
		Middle	0.1284	**0.1402**	0.2078
		Far	0.1248	**0.1446**	0.1904
	(2,4)	Near	0.2238	**0.2680**	0.3668
		Middle	0.1406	**0.1648**	0.2412
		Far	0.1400	**0.1510**	0.2090
	(4,4)	Near	0.4436	**0.4784**	0.5970
		Middle	0.2184	**0.2308**	0.3514
		Far	**0.2088**	0.1938	0.2804

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 22

Power Table When $\sigma = 0.2$ *at* $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	0.0090	**0.0092**	0.0122
		Middle	0.0064	**0.0076**	0.0112
		Far	**0.0090**	0.0070	0.0102
	(0.5,2)	Near	**0.0120**	0.0096	0.0168
		Middle	**0.0106**	0.0088	0.0102
		Far	**0.0116**	0.0086	0.0148
	(2,2)	Near	0.0130	**0.0176**	0.0214
		Middle	**0.0122**	0.0090	0.0192
		Far	0.0094	**0.0124**	0.0136
	(0.5,4)	Near	0.0110	**0.0140**	0.0244
		Middle	0.0126	**0.0184**	0.0252
		Far	0.0102	**0.0132**	0.0152
	(2,4)	Near	0.0126	**0.0230**	0.0384
		Middle	0.0118	**0.0184**	0.0294
		Far	**0.0142**	0.0108	0.0192
	(4,4)	Near	0.0268	**0.0428**	0.0682
		Middle	0.0136	**0.0260**	0.0358
		Far	0.0174	**0.0182**	0.0236
0.04	(0.5,0.5)	Near	0.0096	**0.0098**	0.0128
		Middle	**0.0116**	0.0096	0.0146
		Far	0.0102	**0.0106**	0.0106
	(0.5,2)	Near	0.0110	**0.0176**	0.0254
		Middle	**0.0118**	0.0116	0.0236
		Far	0.0120	**0.0154**	0.0190
	(2,2)	Near	0.0176	**0.0340**	0.0520
		Middle	0.0156	**0.0208**	0.0330
		Far	0.0142	**0.0152**	0.0230
	(0.5,4)	Near	0.0290	**0.0564**	0.0844
		Middle	0.0276	**0.0446**	0.0656
		Far	0.0258	**0.0434**	0.0610
	(2,4)	Near	0.0548	**0.1096**	0.1532
		Middle	0.0332	**0.0516**	0.0778
		Far	0.0342	**0.0462**	0.0654
	(4,4)	Near	0.1642	**0.2688**	0.3472
		Middle	0.0536	**0.0830**	0.1376
		Far	0.0662	**0.0672**	0.1044

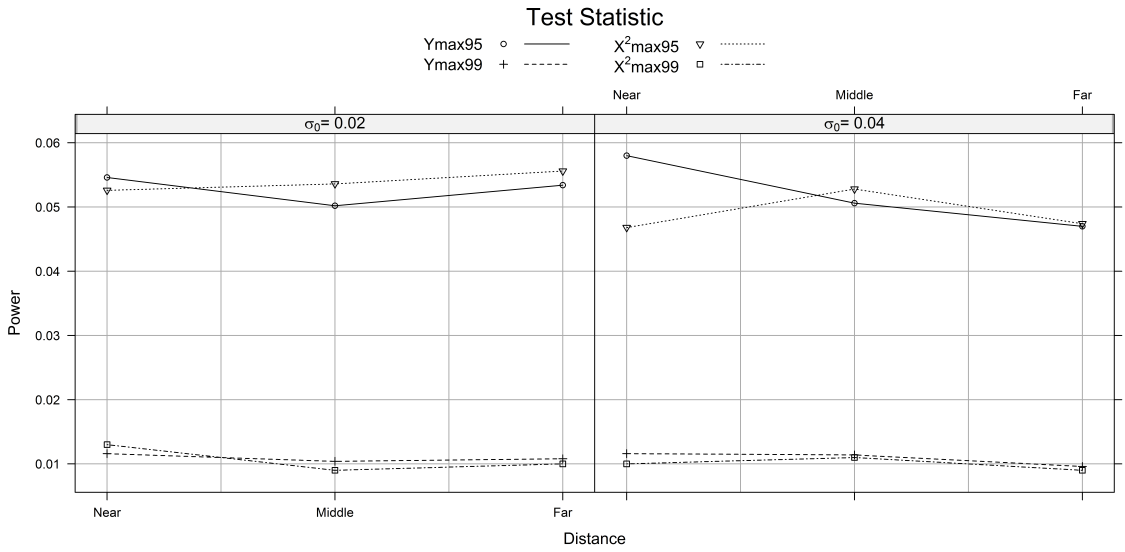
Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

The Effect of Distance on the Power of Y_{max}

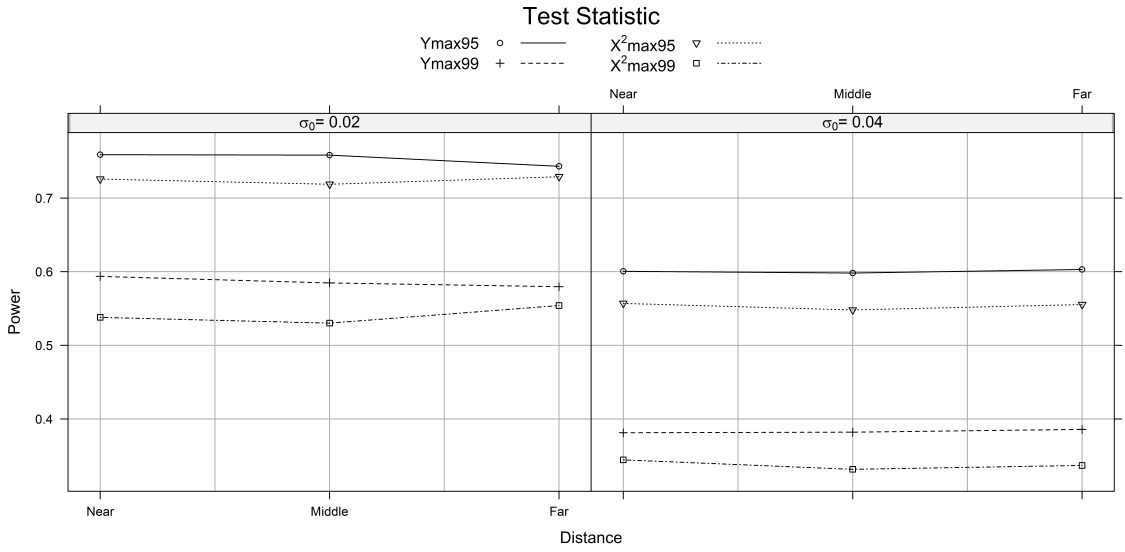
Figure 7 displays four line plots as a visual example illustrating the change of the power of Y_{max} at α of .05 and .01, respectively, under three different conditions of "distance" when other factors were held constant. When the two amplitudes of the signals were the smallest (0.5), the power of $Y_{max}95$ was low (around .05 or so) across all three distance conditions regardless the scale or the kernel width. Nonetheless, when the amplitudes increased, the impact of distance on the power became more related to the other factors. Take Figure 7 (b) and (d) for example. In (b), when the width of kernel and scale were close to each other, the effect of distance was small. However, in (d) as the width of the smoothing kernel considerably deviated from the scale, Y_{max} had greater power when the two signals were close to each other nearly twice as much as when the signals were far away from each other.

The Effect of Amplitude on the Power of Y_{max}

Figure 8 shows the line plots as visual examples of how the power of Y_{max} was related to the amplitudes ξ_1 and ξ_2. Compared to the other factors, the impact of amplitude was substantial. In the present study the power of Y_{max} increased in the following order: (0.5, 0.5), (0.5, 2), (2, 2), (0.5, 4), (2, 4), and (4, 4), when the other factors were held constant. That is, the bigger the amplitude, the greater the power of Y_{max}. As shown in Figure 8, the trend of the lines was all going upward. However, how much the power increased still depended on the other factors. For

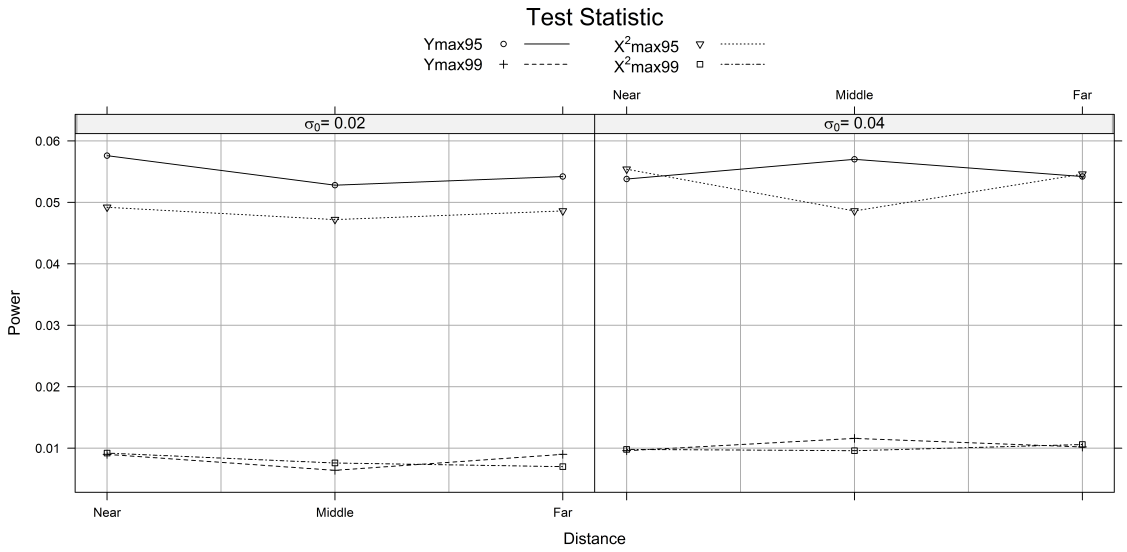


(a) $\sigma = 0.02$ & amplitude $= (0.5, 0.5)$

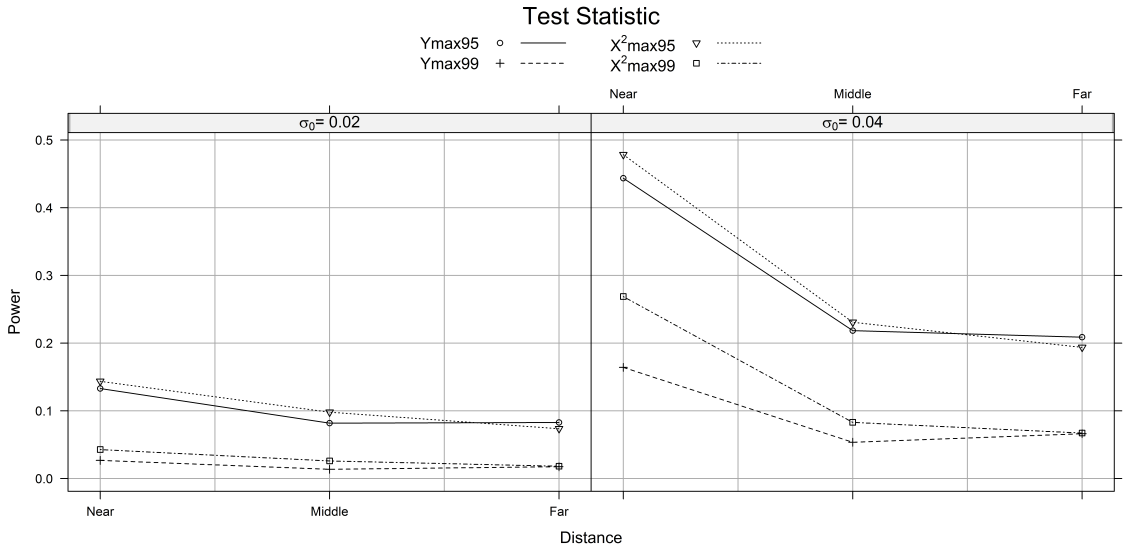


(b) $\sigma = 0.02$ & amplitude $= (4, 4)$

Figure 7. Simulated power vs. distance between the two signals.



(c) $\sigma = 0.2$ & amplitude $= (0.5, 0.5)$



(d) $\sigma = 0.2$ & amplitude $= (4, 4)$

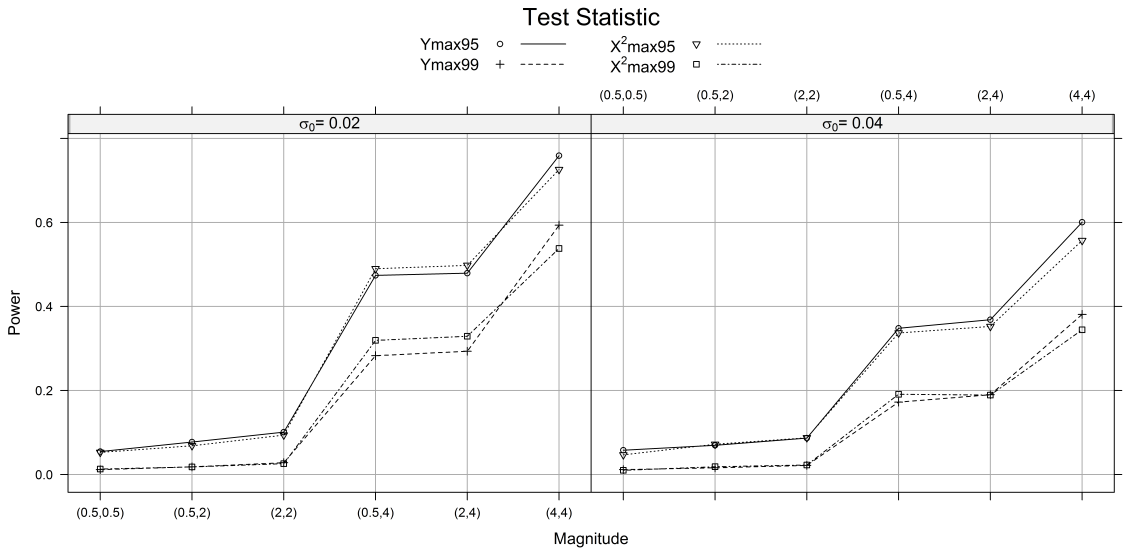
Figure 7. Simulated power vs. distance between the two signals.

example, in Figure 8 (a) and (b) the power was greater when $\sigma_0 = 0.02$ than $\sigma_0 = 0.04$ because of the matched filter theorem ($\sigma = \sigma_0 = 0.02$), whilst in (c) and (d) Y_{max} had low power when the image was over-smoothed using a kernel of 0.2.

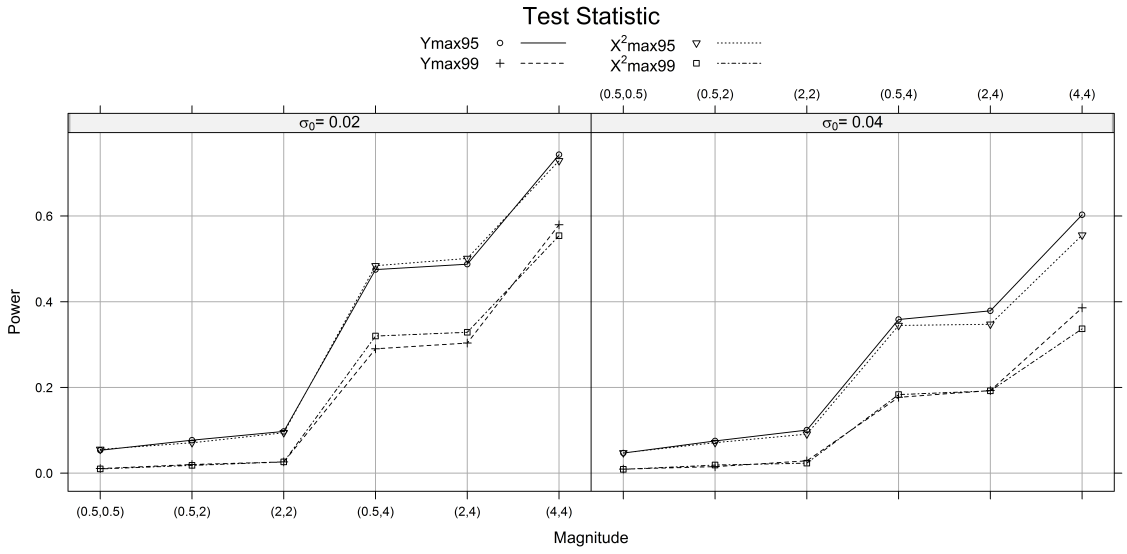
The Effect of Scale and σ on the Power of Y_{max}

Both the scale and the width of smoothing kernel are discussed together. Figure 9 shows the line plots of the power of Y_{max} against σ under certain conditions. When the signals had small amplitudes, it was more difficult to see if the smoothing kernel had any influence on the power of Y_{max}. From Figure 9 (a) and (c), there seemed no particular relationship between the power and the width of smoothing kernel. But as the amplitude increased, the matched filter theorem appeared to play a much more important role that Y_{max} had the greatest power when the scale and smoothing kernel were matched. This could be easily seen in Figure 9 (b) and (d) that when $\sigma_0 = 0.02$ the greatest power happened at $\sigma = 0.02$. Likewise when $\sigma_0 = 0.04$ the power was the largest at $\sigma = 0.0431$.

As for the impact of the scale, regardless of the effect of the matched filter theorem, in general when everything else was held the same, Y_{max} had greater power at $\sigma_0 = 0.04$ than $\sigma_0 = 0.02$. This makes sense because signals with larger scale should be easier to detect.

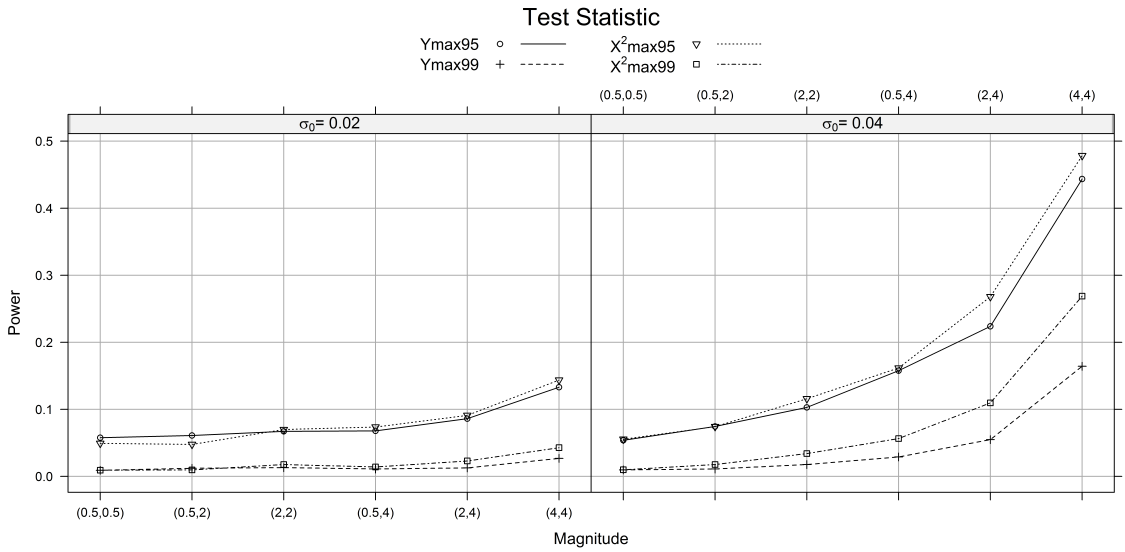


(a) $\sigma = 0.02$ & distance = Near

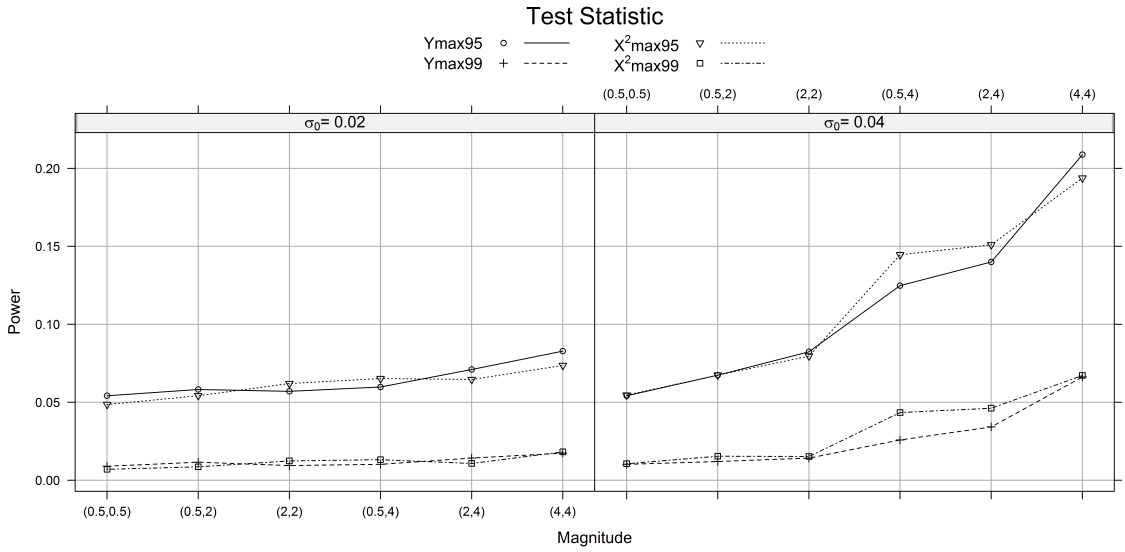


(b) $\sigma = 0.02$ & distance = Far

Figure 8. Simulated power vs. amplitude of the two signals.

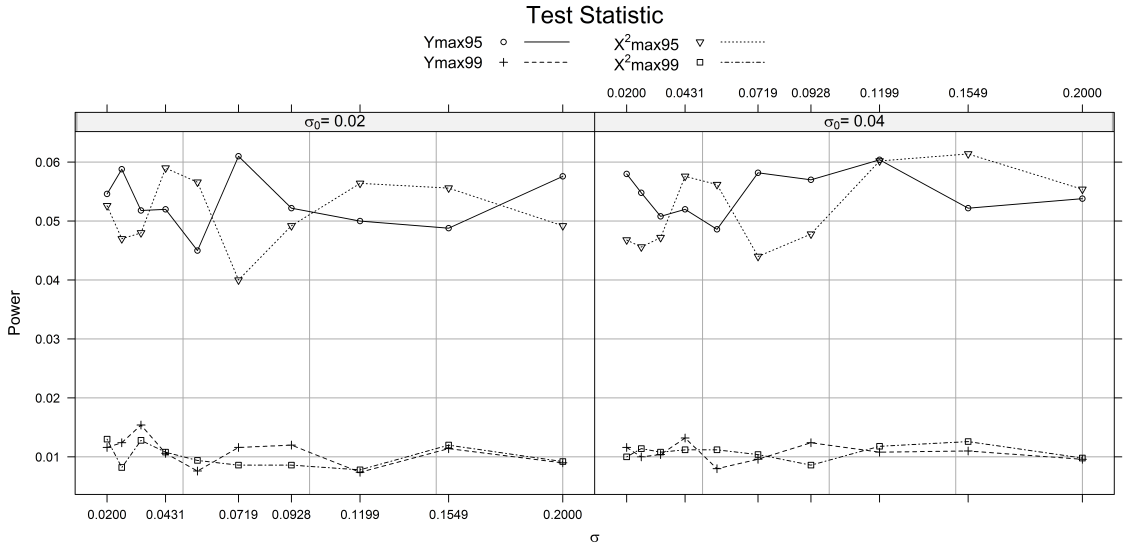


(c) $\sigma = 0.2$ & distance = Near

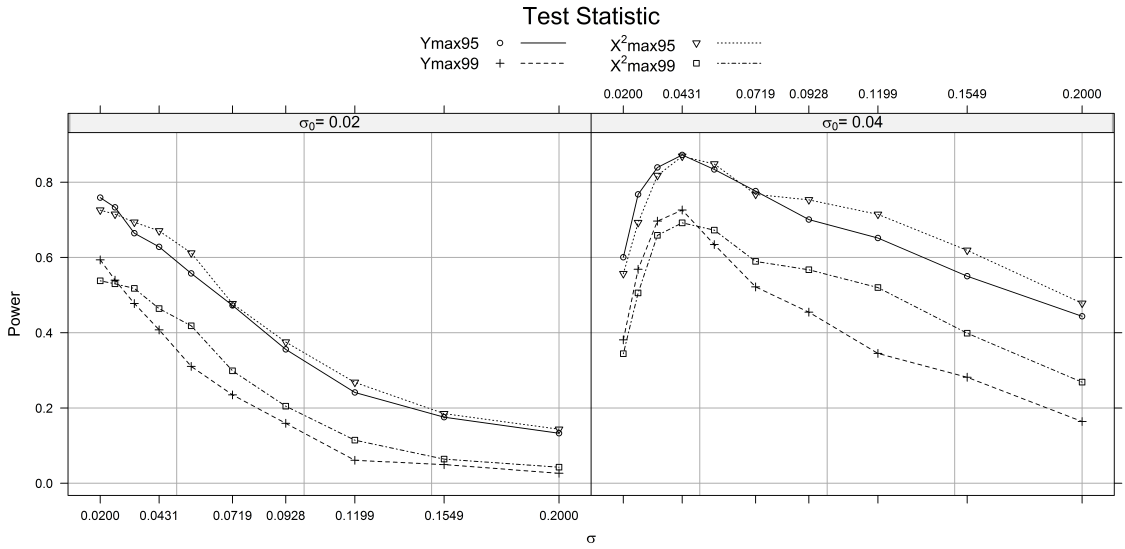


(d) $\sigma = 0.2$ & distance = Far

Figure 8. Simulated power vs. amplitude of the two signals.

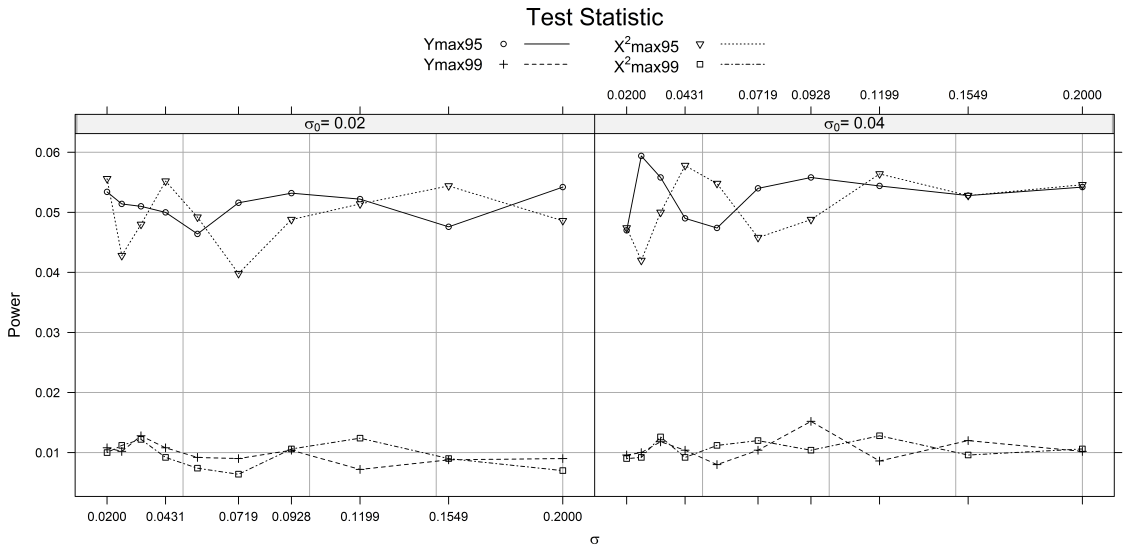


(a) Distance = Near & amplitude = (0.5,0.5)

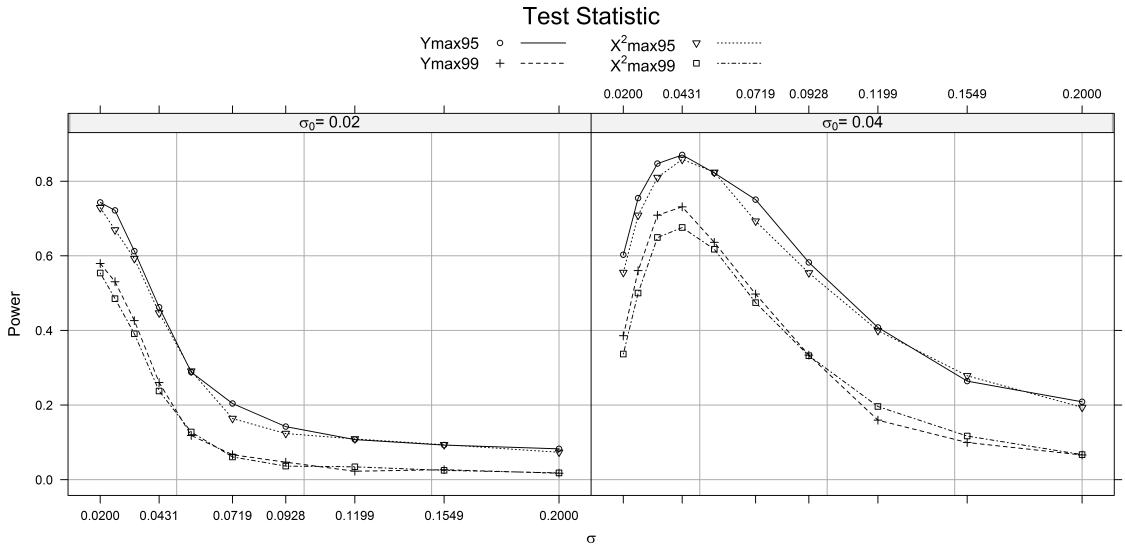


(b) Distance = Near & amplitude = (4,4)

Figure 9. Simulated power vs. σ.



(c) Distance = Far & amplitude = (0.5,0.5)



(d) Distance = Far & amplitude = (4,4)

Figure 9. Simulated power vs. σ.

Power vs. The Norm of the Signals

To further investigate the impact of distance, amplitude, and scale on the power of Y_{max}, the "norm" was used. The norm is defined as the distance between the null and alternative hypothesis. Recall in the null hypothesis there is no signal whereas the alternative assumes there are two signals. If m represents the alternative with the following format,

$$m = \xi_1 f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{01}}{\sigma_0}\right) + \xi_2 f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{02}}{\sigma_0}\right). \tag{31}$$

Then the norm between the null and alternative is

$$\begin{aligned}\|m\|^2 &= \int m^2(\boldsymbol{t})d\boldsymbol{t} \\ &= \underbrace{\int \xi_1^2 f^2\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{01}}{\sigma_0}\right)d\boldsymbol{t}}_{(1)} + \underbrace{\int \xi_2^2 f^2\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{02}}{\sigma_0}\right)d\boldsymbol{t}}_{(2)} + 2\underbrace{\int \xi_1\xi_2 f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{01}}{\sigma_0}\right) f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{02}}{\sigma_0}\right)d\boldsymbol{t}}_{(3)}.\end{aligned} \tag{32}$$

Equation (32) is broken into three parts and is discussed individually as follows.

(1) According to the matched filter theorem we know $\int k^2(\boldsymbol{t})d\boldsymbol{t} = 1$.

Let $\boldsymbol{u} = \frac{\boldsymbol{t}-\boldsymbol{t}_{01}}{\sigma_0}$, so $\int k^2(\boldsymbol{u})d\boldsymbol{u} = 1$ and $\boldsymbol{t} = \sigma_0\boldsymbol{u} + \boldsymbol{t}_{01}$.

Thus the Jacobian $\boldsymbol{J}$ is

$$\boldsymbol{J} = \left|\frac{\partial(\sigma_0\boldsymbol{u}+\boldsymbol{t}_{01})}{\partial\boldsymbol{u}}\right| = \sigma_0\boldsymbol{I},$$

and the Jacobian determinant is

$$|\boldsymbol{J}| = |\sigma_0 \boldsymbol{I}| = \sigma_0^2.$$

Now, the first part of Equation (32) is

$$\begin{aligned} & \int \xi_1^2 f^2 \left(\frac{\boldsymbol{t} - \boldsymbol{t}_{01}}{\sigma_0} \right) d\boldsymbol{t} \\ &= \xi_1^2 \int f^2 \left(\frac{\boldsymbol{t}_{01} + \sigma_0 \boldsymbol{u} - \boldsymbol{t}_{01}}{\sigma_0} \right) \sigma_0^2 d\boldsymbol{u} \\ &= \xi_1^2 \sigma_0^2 \underbrace{\int f^2(\boldsymbol{u}) d\boldsymbol{u}}_{1} \\ &= \xi_1^2 \sigma_0^2. \end{aligned} \tag{33}$$

(2) Likewise, the second part of Equation (32) is

$$\int \xi_2^2 f^2 \left(\frac{\boldsymbol{t} - \boldsymbol{t}_{02}}{\sigma_0} \right) d\boldsymbol{t} = \xi_2^2 \sigma_0^2.$$

(3) Recall the Gaussian signal has the following form,

$$f(\boldsymbol{t}) = \pi^{(-1/2)} exp \left(-\frac{1}{2} \|\boldsymbol{t}\|^2 \right). \tag{34}$$

We first assume $\boldsymbol{t}_{02}|\boldsymbol{t} \sim N(\boldsymbol{t}, \sigma_0^2 \boldsymbol{I})$, which has the probability density function

$$f(\boldsymbol{t}_{02}|\boldsymbol{t}) = \frac{1}{2\pi\sigma_0^2} exp \left(-\frac{\|\boldsymbol{t}_{02} - \boldsymbol{t}\|^2}{2\sigma_0^2} \right),$$

and $\boldsymbol{t} \sim N(\boldsymbol{t}_{01}, \sigma_0^2 \boldsymbol{I})$, which has the probability density function

$$f(\boldsymbol{t}) = \frac{1}{2\pi\sigma_0^2} exp\left(-\frac{\|\boldsymbol{t}-\boldsymbol{t}_{01}\|^2}{2\sigma_0^2}\right).$$

Further assume $\boldsymbol{t}_{02} = \boldsymbol{t} + \boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon} \sim N(0, \sigma_0^2 \boldsymbol{I}) \Rightarrow \boldsymbol{t}_{02} \sim N(\boldsymbol{t}_{01}, 2\sigma_0^2 \boldsymbol{I})$.

The last part of Equation (32) becomes

$$\begin{aligned}
&2\int \xi_1\xi_2 f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{01}}{\sigma_0}\right) f\left(\frac{\boldsymbol{t}-\boldsymbol{t}_{02}}{\sigma_0}\right) d\boldsymbol{t} \\
&= 2\xi_1\xi_2(2\pi^{\frac{1}{2}})^2(\sigma_0^2)^2 \underbrace{\int \frac{1}{2\pi\sigma_0^2} exp\left(\frac{-\|\boldsymbol{t}-\boldsymbol{t}_{01}\|^2}{2\sigma_0^2}\right) \times \frac{1}{2\pi\sigma_0^2} exp\left(\frac{-\|\boldsymbol{t}_{02}-\boldsymbol{t}\|^2}{2\sigma_0^2}\right) d\boldsymbol{t}}_{\boldsymbol{t}_{02}\sim N(\boldsymbol{t}_{01}, 2\sigma_0^2\boldsymbol{I})} \\
&= 2 \times \xi_1 \times \xi_2 \times 4\pi \times (\sigma_0^2)^2 \times \frac{1}{4\pi\sigma_0^2} exp\left(\frac{-\|\boldsymbol{t}_{02}-\boldsymbol{t}_{01}\|^2}{4\sigma_0^2}\right) \\
&= 2 \times \xi_1 \times \xi_2 \times \sigma_0^2 \times exp\left(\frac{-\|\boldsymbol{t}_{02}-\boldsymbol{t}_{01}\|^2}{4\sigma_0^2}\right).
\end{aligned} \tag{35}$$

As a result, the norm is expressed as follows

$$\|m\|^2 = \xi_1^2\sigma_0^2 + \xi_2^2\sigma_0^2 + 2\xi_1\xi_2\sigma_0^2 exp\left(\frac{-\|\boldsymbol{t}_{02}-\boldsymbol{t}_{01}\|^2}{4\sigma_0^2}\right). \tag{36}$$

Equation (36) was then used to calculate the norm for all 36 schemes shown in Table 1 (not 360 conditions because Equation (36) does not involve the width of smoothing kernel σ.). Plots of the power of $Y_{max}95$ against the norm were drawn for selected σ values to describe the relationship (see Figure 10). Figure 10 (b), (c) and (d) clearly show the dots were forming a nearly straight line trending left to right upward, meaning the dots which had bigger norm values also had greater power. In

other words, the conditions that yielded bigger norm were easier for Y_{max} to detect the signals. This is intuitive because the norm represents the distance between the null and alternative hypothesis. Bigger norm means the alternative is farther away from the null, and consequently it should result in greater power.

Nonetheless, the relationship between the norm and power was less clear from Figure 10 (a). This is so because of the matched filter theorem — when the smoothing kernel was 0.02, the power of $Y_{max}95$ would become the greatest at $\sigma_0 = 0.02$. In Figure 11 the observations were grouped by the scale and it is clear the slopes of the lines were positive after taking the matched filter theorem into consideration.

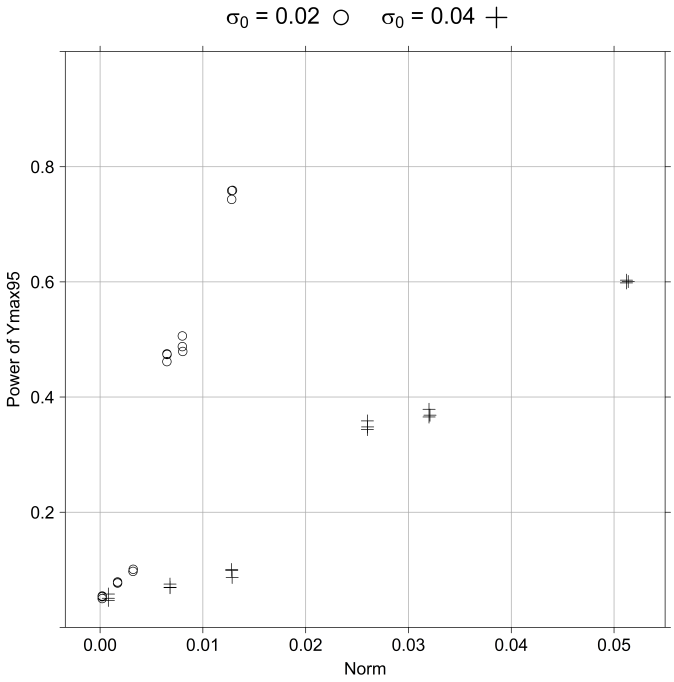
To better understand the nature of the result, the observations resulted in greater power and higher norm values were singled out, and their corresponding conditions were examined. In summary, both the norm and power were higher when at least one of the amplitude parameters, ξ_1 and ξ_2, was four (i.e., the highest value for amplitude in this study), and when the two signals were close to each other. The scale of the signals only appeared to affect the power when the condition of the matched filter theorem was met. By looking at the expression of Equation (36), it is clear that the value of norm was dominated by the squared values of ξ_1 and ξ_2. On the other hand, the scale had smaller influence because the values were set to be too small (only 0.02 and 0.04). Likewise, two of the three distance conditions in the current study, “middle” and “far”, did not have much an impact on the norm

because after taking the exponential the value of the last term in Equation 36 was close to zero.

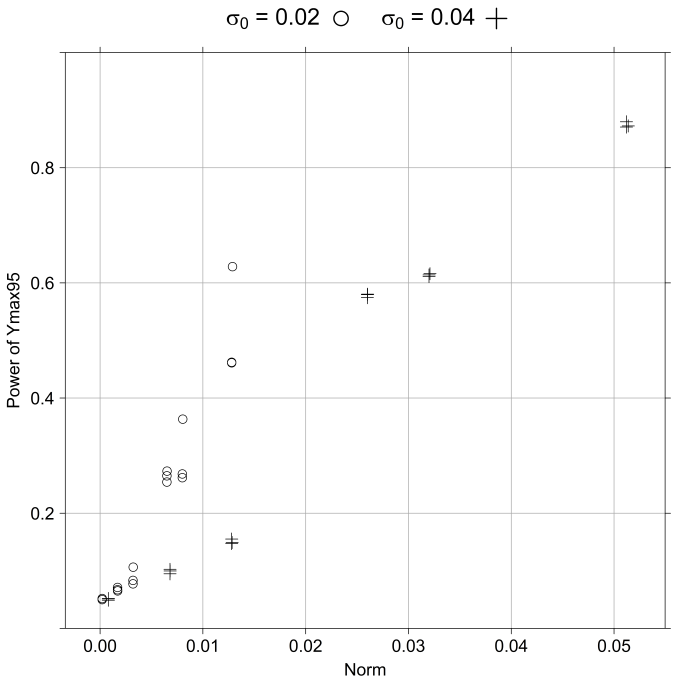
Comparison of the Power between Y_{max} and X^2_{max}

In Table 3 - Table 22, any Y_{max} whose value was larger than X^2_{max} was made in boldface. Throughout all the conditions simulated in the current study, when $\alpha = .05$ there were 167 (46.38 %) out of 360 conditions in which Y_{max} outperformed X^2_{max} whilst when $\alpha = .01$ there were 117 (32.5 %) conditions. However, based on Table 3 - Table 22, Figure 7, Figure 8, and Figure 9, it was difficult to determine under what conditions Y_{max} performed better than X^2_{max} — sometimes Y_{max} did better than X^2_{max} but sometimes it did not.

Figure 12 displays the line plots of the norm against the power of Y_{max} and X^2_{max} with selected smoothing kernels were also drawn. However, the lines were close to each other that they did not particularly show which test statistic was doing better than the other. Further, though the power of Y_{max} was greater than X^2_{max} under certain cases, in general the differences were not much. When $\alpha = .05$ at most the power of Y_{max} was only .0752 greater than X^2_{max} whereas when $\alpha = .01$ at most the power of Y_{max} was .063 better than X^2_{max}. Based on the results obtained in the present study, Y_{max} did not perform as well as expected when comparing to the other conventional test statistics. However, note the maximum value for norm was only 0.0514, which is extremely small. Thus with only limited information, it is too early to make any conclusion about the performance of Y_{max}.

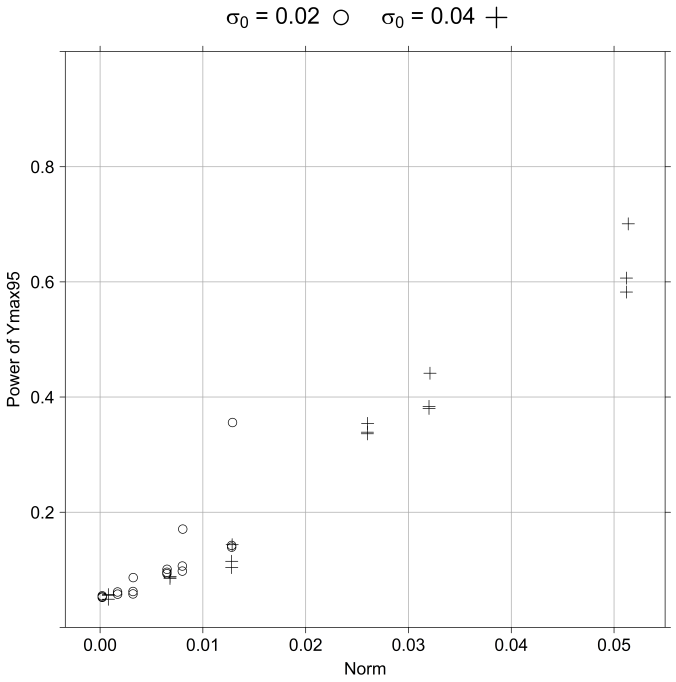


(a) $\sigma = 0.02$

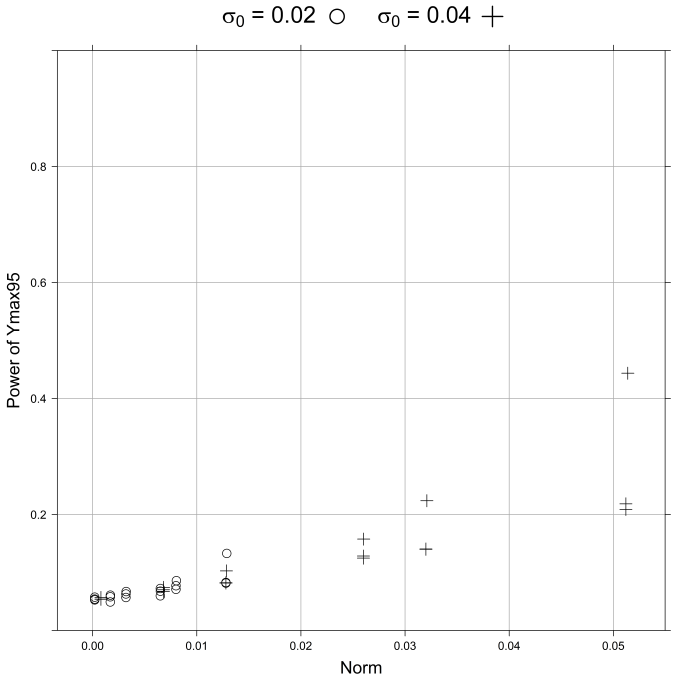


(b) $\sigma = 0.0431$

Figure 10. Simulated power of Y_{max} vs. norm for selected σ values at $\alpha = .05$.



(c) $\sigma = 0.0928$



(d) $\sigma = 0.2$

Figure 10. Simulated power of Y_{max} vs. norm for selected σ values at $\alpha = .05$

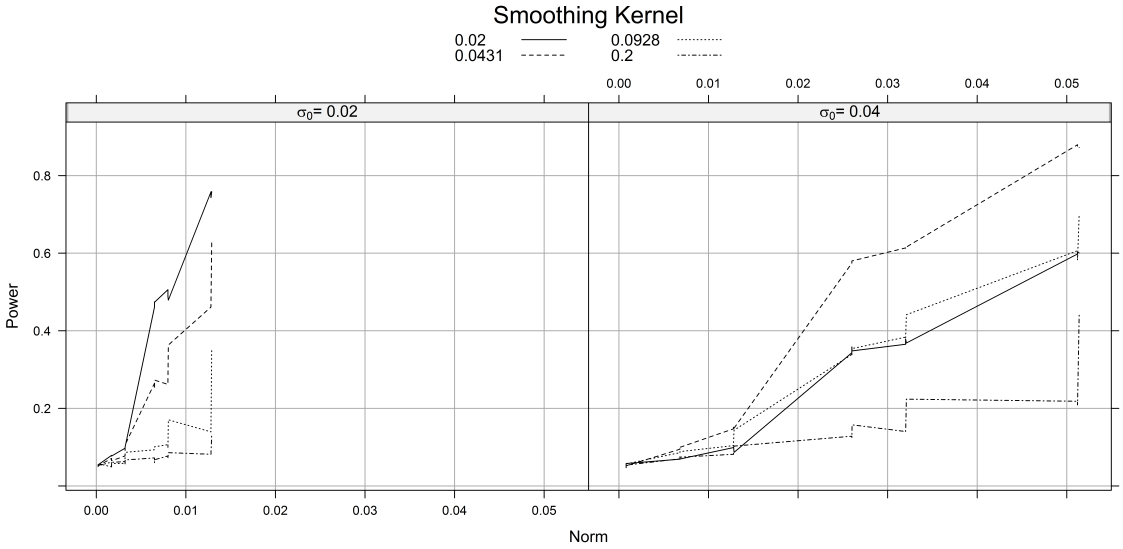
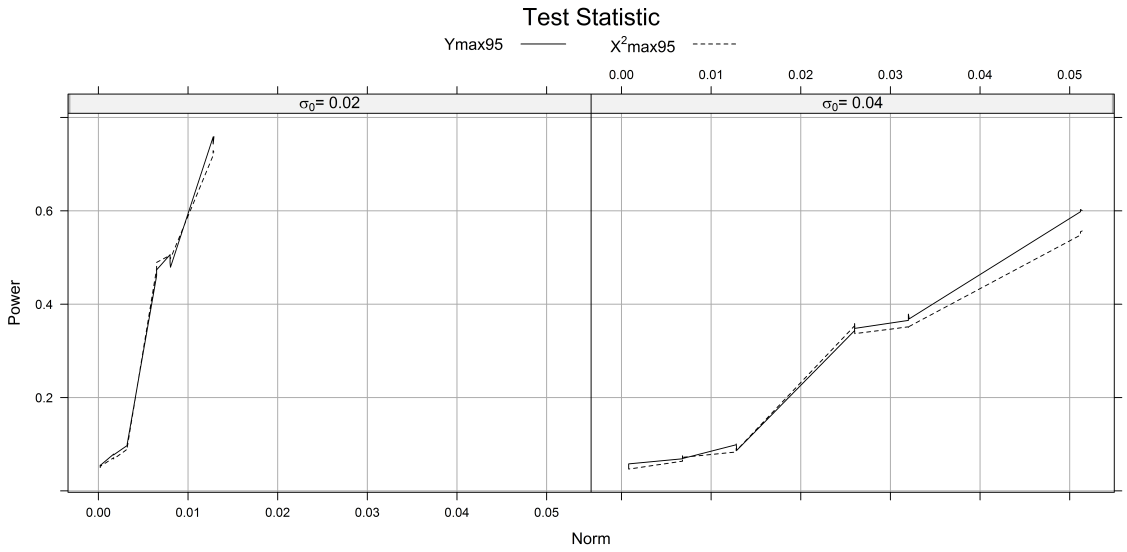
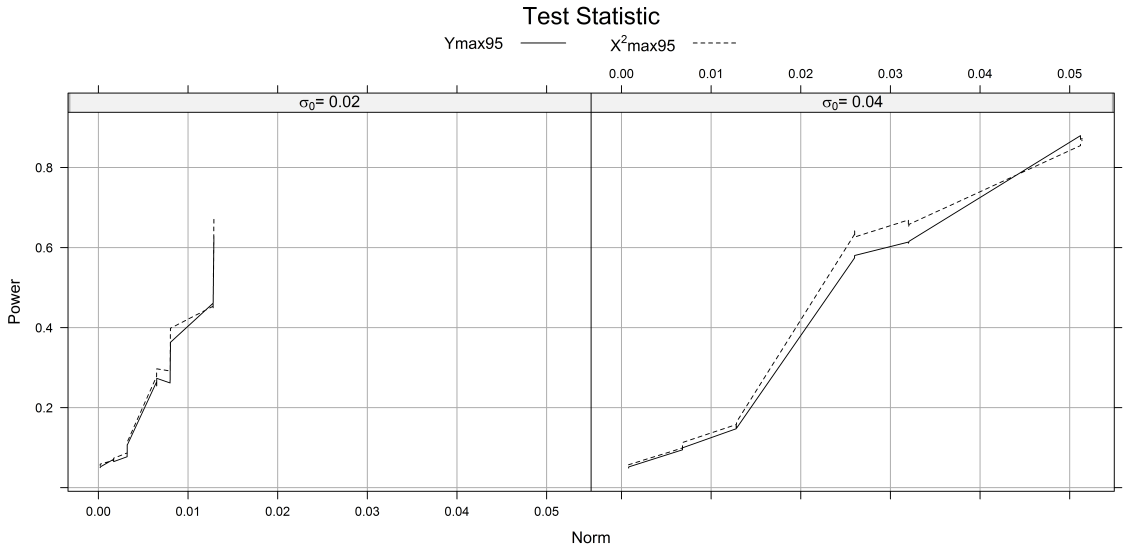


Figure 11. Simulated power of $Y_{max}95$ vs. norm grouped by scale

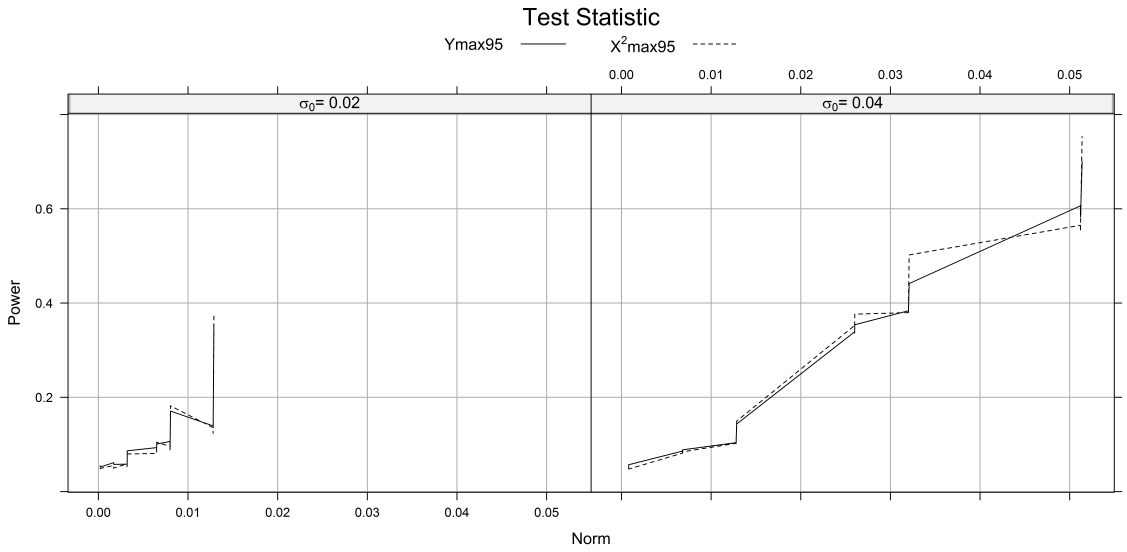


(a) $\sigma = 0.02$

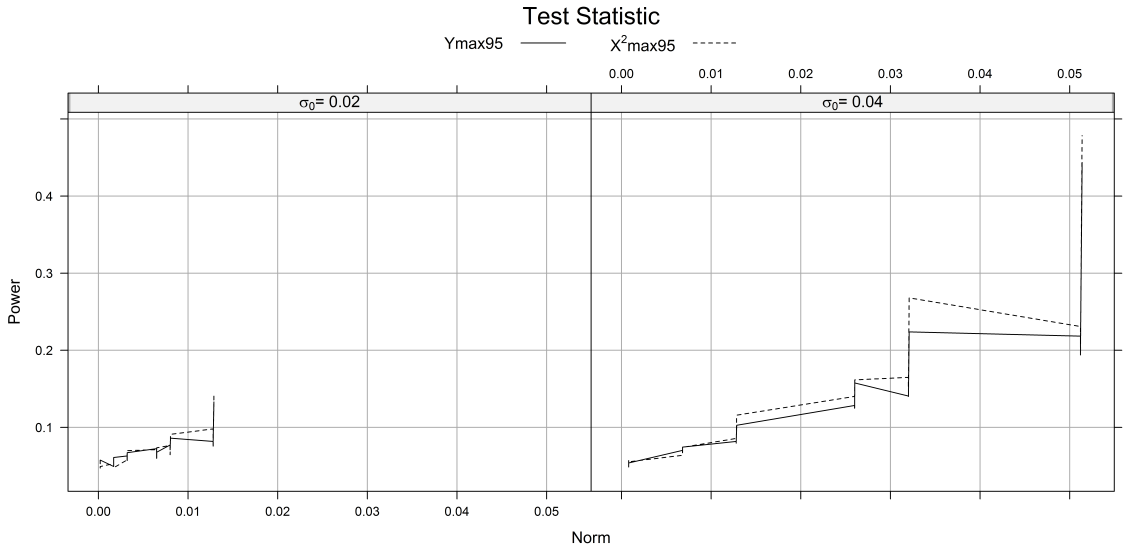


(b) $\sigma = 0.0431$

Figure 12. Comparison between Y_{max} and X^2_{max} on power against norm at $\alpha = .05$



(c) $\sigma = 0.0928$



(d) $\sigma = 0.2$

Figure 12. Comparison between Y_{max} and X^2_{max} on power against norm at $\alpha = .05$

Results for the Scale Space

This section focuses on the results obtained in the scale space where the smoothing kernel was treated as a variable, instead of a fixed parameter. The critical values of Y_{max} and X^2_{max} at α of .05, and .01, respectively, were presented in Table 23. A visualization of the empirical distribution of Y_{max} at $\alpha = .05$ along with the density plot were shown in Figure 13. It is worth noting that the critical values were all higher than the ones obtained when the smoothing parameter was fixed.

Table 23

Critical Values in the Scale Space

$\alpha = .05$			$\alpha = .01$		
Y_{max}	X^2_{max}	X_{max}	Y_{max}	X^2_{max}	X_{max}
34.0949	20.3229	4.3615	38.5596	23.5224	4.7064

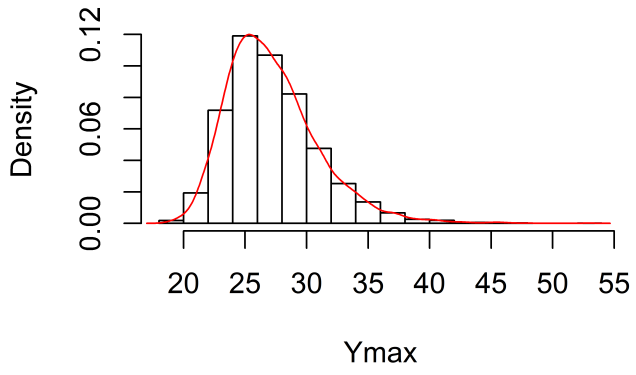


Figure 13. The empirical distribution of scale space $Y_{max}95$

The power of Y_{max} and X^2_{max} at α of .05 and .01 were reported in Table 24 and Table 25, respectively. As mentioned earlier, in scale space the critical values

were more stringent. However, it came with a reward that the power was greater than when the smoothing parameter was fixed. To compare, in Table 26 and Table 27, the power of scale space Y_{max} were presented along with the ones when the smoothing parameter was fixed at selected widths — four different σ values were chosen and they are 0.02, 0.0431, 0.0928, and 0.2. In both tables the values of scale space Y_{max} were in boldface if the power was greater than any of the ones with fixed smoothing parameters. At $\alpha = .05$, the scale space Y_{max} has greater power in 25 out of 36 conditions (69.4%) whereas at $\alpha = .01$ it was 20 out of 36 (55.56%). Overall, in scale space Y_{max} had greater power than the ones at fixed kernel width. This was especially true when the amplitude of the signals was large. For example, when $(\xi_1, \xi_2) = (4, 4)$, the power of scale space Y_{max} was over .90, which was much higher than the case of fixed kernel width. It is important to note that the power of scale space Y_{max} was generally greater than the ones which met the condition of the matched filter theorem. For example, for a signal with a scale of 0.02, most of the time the power of scale space Y_{max} was higher than the power of Y_{max} with a fix kernel width of .02.

Table 24

Power Table in Scale Space at $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}95$	$X^2_{max}95$	$X_{max}95$
0.02	(0.5,0.5)	Near	0.0526	**0.0534**	0.0588
		Middle	**0.0484**	0.0476	0.0494
		Far	0.0500	**0.0518**	0.0464
	(0.5,2)	Near	0.0774	**0.0888**	0.1004
		Middle	**0.0740**	0.0732	0.0970
		Far	0.0790	**0.0912**	0.0966
	(2,2)	Near	0.1338	**0.1522**	0.2006
		Middle	0.1054	**0.1074**	0.1470
		Far	0.1028	**0.1134**	0.1378
	(0.5,4)	Near	0.6354	**0.7154**	0.8008
		Middle	0.6324	**0.7022**	0.7740
		Far	0.6414	**0.7026**	0.7854
	(2,4)	Near	0.7044	**0.7910**	0.8578
		Middle	0.6642	**0.7194**	0.8006
		Far	0.6632	**0.7180**	0.7886
	(4,4)	Near	0.9586	**0.9772**	0.9916
		Middle	**0.9260**	0.9064	0.9502
		Far	**0.9096**	0.9066	0.9448
0.04	(0.5,0.5)	Near	0.0526	**0.0544**	0.0518
		Middle	**0.0562**	0.0520	0.0582
		Far	0.0484	**0.0534**	0.0506
	(0.5,2)	Near	0.1012	**0.1230**	0.1560
		Middle	0.0952	**0.1212**	0.1482
		Far	0.1046	**0.1132**	0.1522
	(2,2)	Near	0.1526	**0.2264**	0.2700
		Middle	0.1452	**0.1768**	0.2368
		Far	0.1506	**0.1754**	0.2388
	(0.5,4)	Near	0.7914	**0.9050**	0.9384
		Middle	0.7968	**0.9040**	0.9376
		Far	0.8028	**0.8994**	0.9408
	(2,4)	Near	0.8416	**0.9294**	0.9642
		Middle	0.8334	**0.9050**	0.9438
		Far	0.8404	**0.8940**	0.9478
	(4,4)	Near	0.9938	**0.9972**	0.9988
		Middle	**0.9928**	0.9890	0.9972
		Far	0.9908	**0.9912**	0.9966

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 25

Power Table in Scale Space at $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	$Y_{max}99$	$X^2_{max}99$	$X_{max}99$
0.02	(0.5,0.5)	Near	**0.0128**	0.0094	0.0130
		Middle	0.0094	0.0106	0.0102
		Far	0.0094	0.0114	0.0096
	(0.5,2)	Near	0.0178	0.0214	0.0324
		Middle	0.0190	0.0192	0.0300
		Far	0.0186	0.0256	0.0282
	(2,2)	Near	0.0368	0.0470	0.0766
		Middle	0.0238	0.0356	0.0434
		Far	0.0294	0.0386	0.0462
	(0.5,4)	Near	0.3820	0.4942	0.6010
		Middle	0.3812	0.4950	0.5704
		Far	0.3826	0.4972	0.5842
	(2,4)	Near	0.4392	0.5874	0.6692
		Middle	0.3980	0.4978	0.5880
		Far	0.4026	0.5008	0.5848
	(4,4)	Near	0.8204	0.8890	0.9424
		Middle	**0.7614**	0.7374	0.8312
		Far	**0.7502**	0.7252	0.8076
0.04	(0.5,0.5)	Near	**0.0110**	0.0106	0.0096
		Middle	**0.0106**	0.0102	0.0124
		Far	**0.0110**	0.0092	0.0080
	(0.5,2)	Near	0.0224	0.0382	0.0530
		Middle	0.0236	0.0382	0.0520
		Far	0.0246	0.0332	0.0538
	(2,2)	Near	0.0398	0.0806	0.1018
		Middle	0.0354	0.0570	0.0866
		Far	0.0406	0.0644	0.0874
	(0.5,4)	Near	0.5204	0.7434	0.8118
		Middle	0.5210	0.7390	0.8070
		Far	0.5372	0.7280	0.8096
	(2,4)	Near	0.5896	0.7728	0.8488
		Middle	0.5678	0.7326	0.8138
		Far	0.5814	0.7260	0.8200
	(4,4)	Near	0.9452	0.9584	0.9840
		Middle	**0.9330**	0.9284	0.9628
		Far	**0.9368**	0.9244	0.9674

Note. Any Y_{max} larger than X^2_{max} is in boldface; likewise, any X^2_{max} larger than Y_{max} is in boldface.

Table 26

Comparison of the Power of Y_{max}'s at $\alpha = .05$

σ_0	(ξ_1, ξ_2)	Distance	Y_{max} (Scale space)	Y_{max} ($\sigma = 0.02$)	Y_{max} ($\sigma = 0.0431$)	Y_{max} ($\sigma = 0.0928$)	Y_{max} ($\sigma = 0.2$)
0.02	(0.5,0.5)	Near	0.0526	0.0546	0.0520	0.0522	0.0576
		Middle	0.0484	0.0502	0.0518	0.0550	0.0528
		Far	0.0500	0.0534	0.0500	0.0532	0.0542
	(0.5,2)	Near	0.0774	0.0774	0.0652	0.0578	0.0610
		Middle	0.0740	0.0790	0.0712	0.0616	0.0492
		Far	**0.0790**	0.0770	0.0672	0.0578	0.0582
	(2,2)	Near	**0.1338**	0.1010	0.1064	0.0866	0.0672
		Middle	**0.1054**	0.0974	0.0774	0.0582	0.0630
		Far	**0.1028**	0.0974	0.0834	0.0624	0.0570
	(0.5,4)	Near	**0.6354**	0.4740	0.2732	0.1010	0.0678
		Middle	**0.6324**	0.4616	0.2650	0.0934	0.0724
		Far	**0.6414**	0.4750	0.2542	0.0958	0.0598
	(2,4)	Near	**0.7044**	0.4792	0.3634	0.1708	0.0860
		Middle	**0.6642**	0.5062	0.2618	0.1066	0.0772
		Far	**0.6632**	0.4876	0.2682	0.0978	0.0710
	(4,4)	Near	**0.9586**	0.7590	0.6282	0.3558	0.1330
		Middle	**0.9260**	0.7584	0.4612	0.1394	0.0818
		Far	**0.9096**	0.7432	0.4622	0.1424	0.0828
0.04	(0.5,0.5)	Near	0.0526	0.0580	0.0520	0.0570	0.0538
		Middle	**0.0562**	0.0506	0.0512	0.0490	0.0570
		Far	0.0484	0.0470	0.0490	0.0558	0.0542
	(0.5,2)	Near	**0.1012**	0.0696	0.0998	0.0886	0.0744
		Middle	**0.0952**	0.0690	0.0948	0.0860	0.0702
		Far	**0.1046**	0.0752	0.1026	0.0852	0.0674
	(2,2)	Near	**0.1526**	0.0868	0.1492	0.1436	0.1028
		Middle	0.1452	0.0990	0.1474	0.1040	0.0818
		Far	0.1506	0.1006	0.1552	0.1148	0.0824
	(0.5,4)	Near	**0.7914**	0.3482	0.5804	0.3542	0.1576
		Middle	**0.7968**	0.3436	0.5744	0.3390	0.1284
		Far	**0.8028**	0.3586	0.5796	0.3362	0.1248
	(2,4)	Near	**0.8416**	0.3686	0.6160	0.4414	0.2238
		Middle	**0.8334**	0.3654	0.6138	0.3838	0.1406
		Far	**0.8404**	0.3788	0.6110	0.3804	0.1400
	(4,4)	Near	**0.9938**	0.6006	0.8726	0.7008	0.4436
		Middle	**0.9928**	0.5982	0.8798	0.6066	0.2184
		Far	**0.9908**	0.6030	0.8706	0.5824	0.2088

Note. Any scale space Y_{max} larger than the Y_{max} at fixed σ is in boldface.

Table 27

Comparison of the Power of Y_{max}'s at $\alpha = .01$

σ_0	(ξ_1, ξ_2)	Distance	Y_{max} (Scale space)	Y_{max} ($\sigma = 0.02$)	Y_{max} ($\sigma = 0.0431$)	Y_{max} ($\sigma = 0.0928$)	Y_{max} ($\sigma = 0.2$)
0.02	(0.5,0.5)	Near	**0.0128**	0.0116	0.0106	0.0120	0.0090
		Middle	0.0094	0.0104	0.0136	0.0130	0.0064
		Far	0.0094	0.0108	0.0108	0.0104	0.0090
	(0.5,2)	Near	0.0178	0.0182	0.0168	0.0142	0.0120
		Middle	0.0190	0.0166	0.0196	0.0146	0.0106
		Far	0.0186	0.0202	0.0174	0.0148	0.0116
	(2,2)	Near	**0.0368**	0.0282	0.0308	0.0182	0.0130
		Middle	0.0238	0.0306	0.0176	0.0126	0.0122
		Far	**0.0294**	0.0264	0.0218	0.0164	0.0094
	(0.5,4)	Near	**0.3820**	0.2830	0.1280	0.0292	0.0110
		Middle	**0.3812**	0.2752	0.1188	0.0270	0.0126
		Far	**0.3826**	0.2902	0.1144	0.0264	0.0102
	(2,4)	Near	**0.4392**	0.2934	0.1768	0.0510	0.0126
		Middle	**0.3980**	0.3038	0.1190	0.0288	0.0118
		Far	**0.4026**	0.3036	0.1210	0.0286	0.0142
	(4,4)	Near	**0.8204**	0.5936	0.4076	0.1594	0.0268
		Middle	**0.7614**	0.5848	0.2578	0.0424	0.0136
		Far	**0.7502**	0.5796	0.2604	0.0470	0.0174
0.04	(0.5,0.5)	Near	0.0110	0.0116	0.0132	0.0124	0.0096
		Middle	0.0106	0.0114	0.0116	0.0134	0.0116
		Far	0.0110	0.0096	0.0104	0.0152	0.0102
	(0.5,2)	Near	0.0224	0.0160	0.0300	0.0214	0.0110
		Middle	0.0236	0.0178	0.0250	0.0204	0.0118
		Far	0.0246	0.0154	0.0320	0.0216	0.0120
	(2,2)	Near	0.0398	0.0218	0.0496	0.0394	0.0176
		Middle	0.0354	0.0284	0.0494	0.0280	0.0156
		Far	0.0406	0.0288	0.0510	0.0300	0.0142
	(0.5,4)	Near	**0.5204**	0.1724	0.3710	0.1576	0.0290
		Middle	**0.5210**	0.1762	0.3812	0.1578	0.0276
		Far	**0.5372**	0.1772	0.3698	0.1452	0.0258
	(2,4)	Near	**0.5896**	0.1900	0.4268	0.2148	0.0548
		Middle	**0.5678**	0.1822	0.4108	0.1712	0.0332
		Far	**0.5814**	0.1928	0.4062	0.1716	0.0342
	(4,4)	Near	**0.9452**	0.3812	0.7262	0.4548	0.1642
		Middle	**0.9330**	0.3820	0.7438	0.3406	0.0536
		Far	**0.9368**	0.3858	0.7314	0.3338	0.0662

Note. Any scale space Y_{max} larger than the Y_{max} at fixed σ is in boldface.

The Effect of Distance on the Power of Scale Space Y_{max}

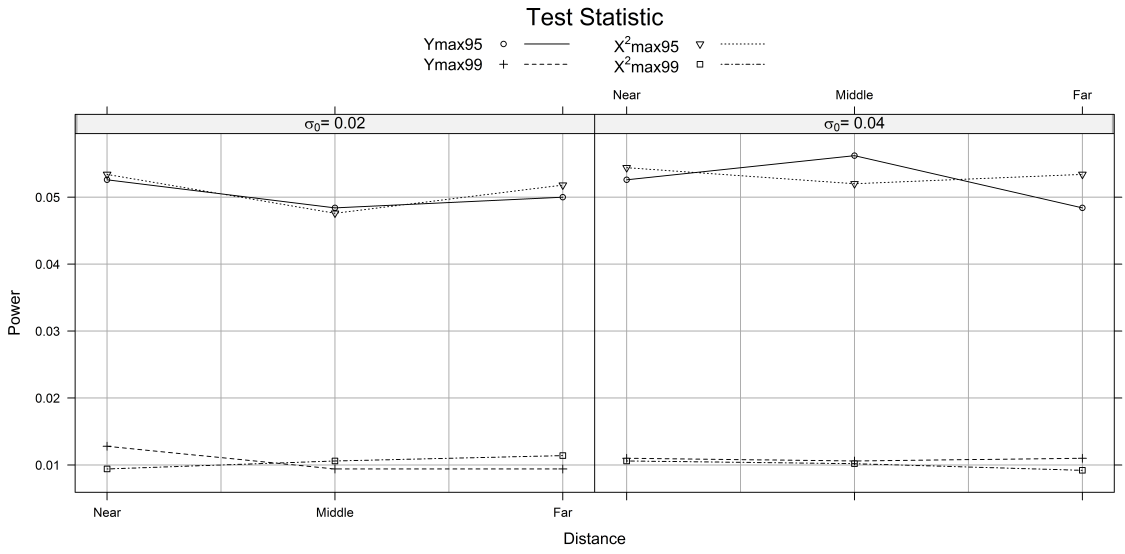
The impact of distance on the power of Y_{max} was not substantial. The power was only a bit greater when the signals were close together if $\sigma_0 = 0.02$ coupled with larger amplitude. Nonetheless, when $\sigma_0 = 0.04$ it seemed the distance did not affect on the power at all (see Figure 14).

The Effect of Amplitude on the Power of Scale Space Y_{max}

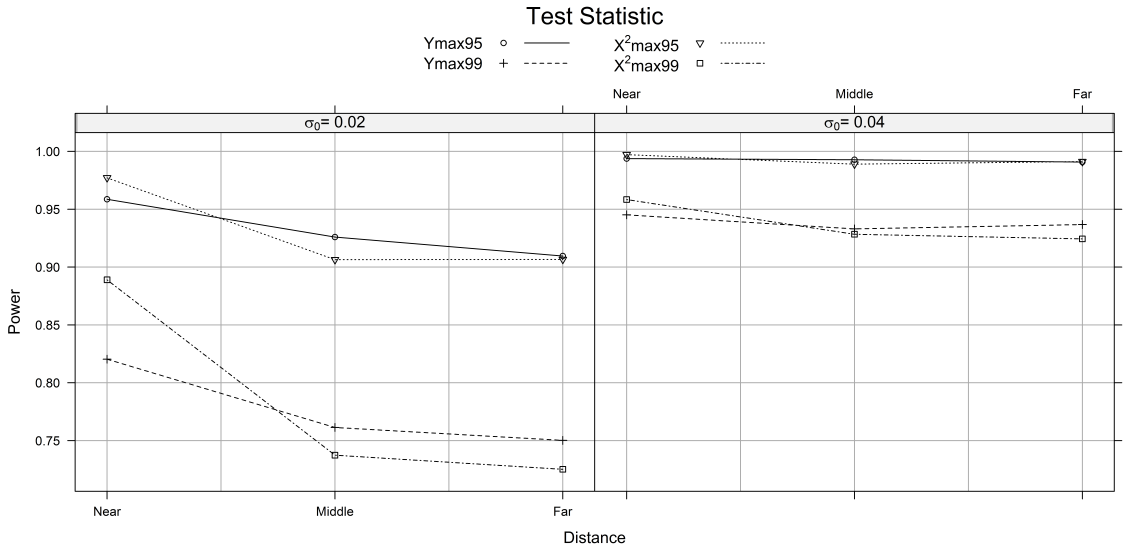
For the effect of amplitude, it was positively related to the power of Y_{max}. This is intuitive that with greater amplitude it is easier for the test statistics to detect the signals. And as shown in Figure 15, the power of Y_{max} increased in the following sequence: (0.5, 0.5), (0.5, 2), (2, 2), (0.5, 4), (2, 4), and (4, 4) when everything else was held the same.

The Effect of Scale on the Power of Y_{max}

In general, when everything else was held the same, at $\sigma_0 = 0.04$ the power of Y_{max} was greater than at $\sigma_0 = 0.02$. For example, under the case of $(\xi_1, \xi_2) = (0.5, 4)$ coupled with the "near" condition, the power of $Y_{max}95$ was .7914 when $\sigma_0 = 0.04$ whereas the power was only .6354 when $\sigma_0 = 0.02$. The relationship between the scale and power could also be observed from Figure 14 and Figure 15.

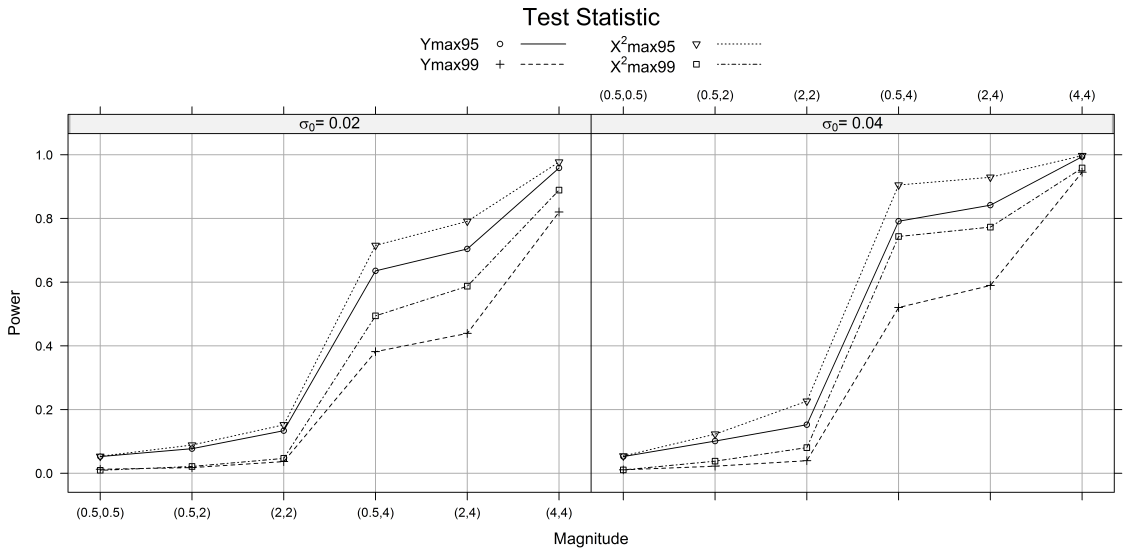


(a) Amplitude = (0.5, 0.5)

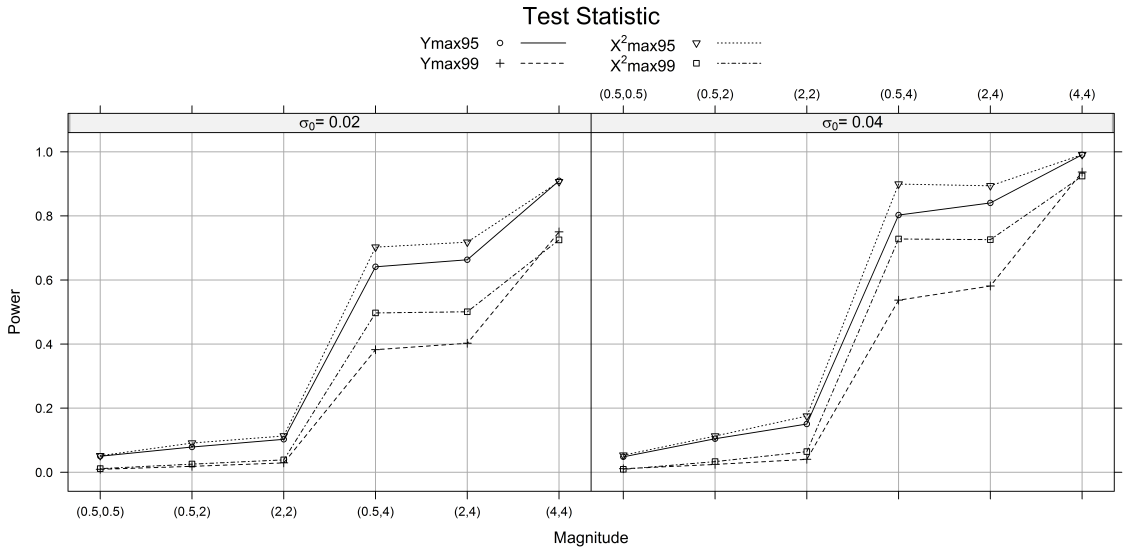


(b) Amplitude = (4, 4)

Figure 14. Simulated power vs. distance in scale space



(a) Distance = Near



(b) Distance = Far

Figure 15. Simulated power vs. amplitude in scale space

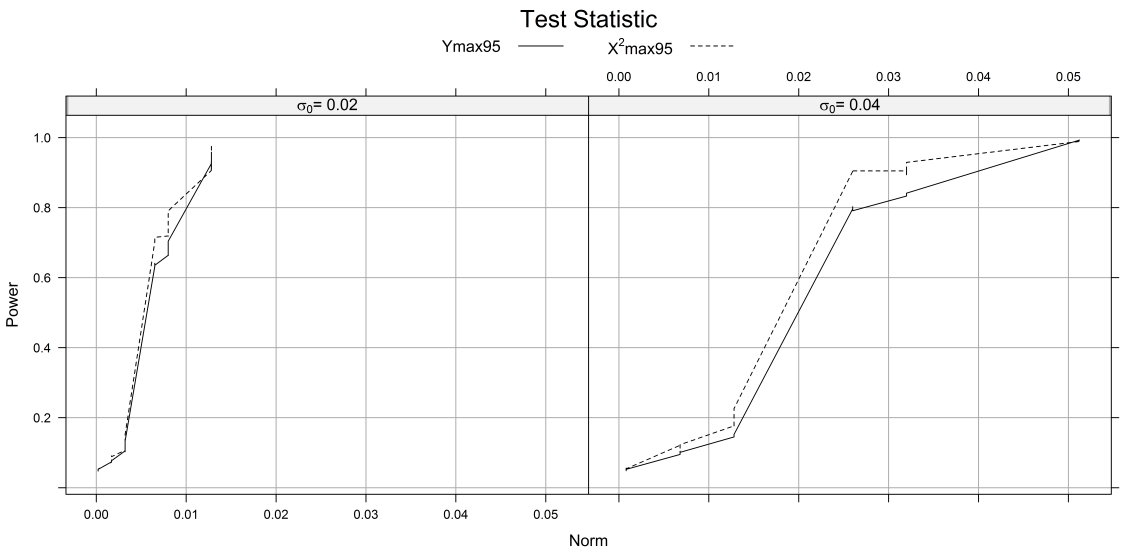


Figure 16. Simulated power vs. norm in scale space at $\alpha = .05$

Power vs. The Norm of the Signals

Figure 16 shows the line plots of the norm values against the power of Y_{max} at $\alpha = .05$. As expected, the norm and the power were positively correlated that the larger the norm the greater the power.

Comparison of the Power between Y_{max} and X^2_{max}

For any scale space Y_{max} whose power was greater than X^2_{max}, the values were made in boldface in Table 24 and Table 25. Throughout all the 36 schemes simulated in the current study, at $\alpha = .05$ there were six (16.67%) conditions in which Y_{max} outperformed X^2_{max} and at $\alpha = .01$, eight (22.22%) conditions in which Y_{max} had greater power. Further, for the cases when Y_{max} surpassed X^2_{max}, the differences were very small. At $\alpha = .05$, at most the power of Y_{max} was 0.0196 greater than X^2_{max} whilst at $\alpha = .01$ at most the power of Y_{max} was .025 greater

than X^2_{max}. Figure 16 also presents the line plots for the power of $Y_{max}95$ and X^2_{max} against the norm respectively. By looking at Figure 16, $X^2_{max}95$ almost always did better than $Y_{max}95$ under the same norm.

CHAPTER V

CONCLUSIONS

The purpose of the current study was to propose a likelihood ratio test statistic, Y_{max}, for testing two signals simultaneously in a two-dimensional image. This test statistic was developed by Shafie (2014) based on Gaussian random field theory. Unlike other commonly-used fields, there is no closed form solutions to approximate the probability of Y_{max}. Hence, this study used Monte Carlo simulation to obtain the critical values as thresholds for signal detection. The power of Y_{max} was also assessed under various factors and compared to other conventional test statistics.

Findings

The author was able to use simulation to approximate the probability of Y_{max} under the null as well as alternative hypothesis. However, it was time-consuming to generate the correlation matrices and to compute the field $Y(\boldsymbol{t_1}, \boldsymbol{t_2})$, which is why the biggest image resolution employed in this study was limited to 64×64. The R code used for simulation is included in Appendix A.

Discussion of the Findings on Fixed Smoothing Kernel

The simulated critical values show that with a smaller smoothing kernel, the resulting threshold was higher. Likewise, when applying a wider kernel width to the

image the threshold was smaller. This is consistent to the results from multiple previous studies (Poline & Mazoyer, 1994; Worsley et al., 1996). The finding should be intuitive because with a smaller smoothing parameter, the image is rougher, and in turn the likelihood of a local maximum exceeding the threshold increases, which causes a higher false positive rate (Friston et al., 1991; Worsley et al., 1996).

After obtaining the critical values, images with two signals plus white noise were simulated under different conditions based on factors including distance, amplitude, and scale. Then the critical values were used for thresholding and the power of Y_{max} was calculated. With fixed kernel width, the results revealed that to some extent all three factors had impact on the power of Y_{max}. For the effect of distance, if the kernel width matched the scale of the signals then the distance between the signals had little effect on the power. However, if the width of the kernel differed from the scale considerably, then it was easier for Y_{max} to detect the signals when they were close to each other. For amplitude, there was a strong and positive relationship between the power and amplitude that the power went up when the signals had greater amplitude. In the current study, the power increased with the amplitudes, (ξ_1, ξ_2), in the following order: (0.5, 0.5), (0.5, 2), (2, 2), (0.5, 4), (2, 4), and (4, 4). Note that the impact of the amplitude on the power was not linear but quadratic, which is shown in the norm formula (36). The impact of the scale is also straightforward that larger scale increased the power. Nonetheless, when the kernel width was fixed, the impact of the scale on power would go hand in hand with the kernel width because of the matched filter theorem, stating that the signal

is best detected if the kernel matched the signal. Hence, in certain cases in this study the power of Y_{max} was greater even when the scale of the signals was smaller.

This study also compared the power of Y_{max} to the one of X^2_{max}. Y_{max} did not seem to outperform X^2_{max} on power across all conditions. It was not clear which test statistic was the winner in this simulation study, and it was also unclear under what conditions Y_{max} was more sensitive than X^2_{max}. This is not satisfying because when testing multiple signals Y_{max} has been shown to be the likelihood ratio test statistic (Shafie, 2014) whereas there is no theoretical reason to employ X^2_{max} in this case. The author suspects the reason might be the conditions incorporated in this study were limited, and thus the result could not provide a full picture of the behavior of Y_{max}.

Discussion of the Findings on the Scale Space Y_{max}

In scale space, the resulting critical values were all higher than the ones simulated with fixed kernel width, which is in line with the previous literature such as Poline and Mazoyer (1994), Worsley (2001), and Worsley et al. (1996). When searching over an extra dimension (σ; the width of smoothing kernel), it comes with the cost of an increase in the critical values of the test statistic.

As for the factors impacting on the sensitivity of detecting signals, both the amplitude and scale were positively correlated with the power of Y_{max}. What was confusing is the impact of the distance. It seems the distance only came into play when the amplitude was large coupled with smaller scale. Under this situation, the

power increased if the signals were close to each other. It is important to note that this does not necessarily mean Y_{max} is actually detecting two distinct signals — because later when looking for the actual locations of the signals, they might be detected as one broader signal instead of two separate signals. Previous literature points out that the key shortcoming of employing the scale space (and rotation space) method is that signals which are near each other often get mistakenly detected as a single signal rather than separate signals (Shafie et al., 2003; Worsley, 2001; Worsley et al., 1996). Further investigation is required to determine if Y_{max} share the same concern as the literature suggests.

Between the scale space Y_{max} and the Y_{max} with fixed smoothing parameter, it shows the scale space Y_{max} was more powerful than the one with fixed kernel width for more than 2/3 of the simulated conditions. This is intriguing because several research studies on scale space (and rotation space) pointed out when the scale of the signal is known the matched filter theorem guarantees the signals will be best detected using the kernel that matches the signal (Shafie et al., 2003; Siegmund & Worsley, 1995; Worsley et al., 1996). That is, if the scale is known then the fixed kernel test is more powerful than the scale space method. In the present study, however, when $\sigma_0 = 0.02$, the power of the scale space Y_{max} was greater than the power of Y_{max} with the fixed kernel width at 0.02 in general. The inconsistency from the literature requires further investigation.

When comparing the power of Y_{max} to the one of X^2_{max}, it shows the scale space Y_{max} performed worse than the cases of fixed kernel width given that the

sensitivity of Y_{max} was almost always lower than X^2_{max}. Nevertheless, since the conditions being simulated in the current study were limited it is too early to make any conclusion about the scale space Y_{max}.

Limitations and Suggestions for Future Research

One of the biggest limitations is that the algorithm used to obtain Y_{max} in the current study is not efficient. As mentioned in the previous section, when simulating a 128×128 image, R could not generate the correlation function $R(\boldsymbol{t_1}, \boldsymbol{t_2})$, let alone the $Y(\boldsymbol{t_1}, \boldsymbol{t_2})$ field. Thus, this algorithm may not be very useful for practitioners if the image size is larger than 64×64, where in fMRI it is not uncommon to have images with resolution of 128×128 or 256×256. Further, even when the image size decreased to 64×64, it still took the algorithm nearly 7 seconds to produced Y_{max} when the kernel width was fixed whereas it took 26 seconds for the scale space method on an Intel©Core i5-3210 2.50 GHz/8 GB RAM-based system.

Another issue arose with the approximation of the tail probability. Recall in this study the thresholds (based on the 95^{th} and 99^{th} percentiles) were approximated from the tail of the empirical distribution. However, these two percentiles may have larger margin of error because of the erratic behavior on the variance of the estimated tail probability (Brown, Cai, & DasGupta, 2001). With that being said, the simulated thresholds might not be as precise and reliable as expected (that if one ran another 5000 simulations, the thresholds might be different from, say, the 2^{nd} decimal place digit of the ones obtained in this study), and in turn the power might get either underestimated or overestimated. In the current study, to ensure

the results are reliable, the author used the Euler characteristic solution for X_{max} derived by Siegmund and Worsley (1995) to approximate the thresholds of scale space X_{max} and compared to the simulated one — the results were close. Nonetheless, it should be kept in mind that the simulation and Euler characteristic formula are both approximations, rather than exact solutions. Several studies published in recent years were trying to address the issue of computing the tail probability. For example, Adler, Blanchet, and Liu (2012) proposed an algorithm for efficient Monte Carlo simulation using importance sampling, hoping to obtain relatively accurate estimates; nonetheless, their method is only for the random field whose distribution is known. More importantly, they also noticed that as the resolution of the image increases, the burden of computation increases as well. Note that this problem is inevitable given that when correcting the multiple comparisons in fMRI using Gaussian random theory, the images are broken into small grids and treated as continuous fields. Thus, discretization makes the computation of the correlation function challenging and more efficient simulation algorithm is definitely needed when dealing with larger lattice sizes.

As mentioned before, the performance of Y_{max} on power was not clear where the author suspected the reason could be the limited conditions being manipulated in the current study. If using the norm as a criterion since it measures the "distance" between the null and alternative hypotheses, the biggest norm value was only 0.05 in this study indicating the alternative hypotheses were not very far away from the null. Thus the future research should try to manipulate Equation (36) to see what

condition may result in a bigger norm. However, if the alternative is obviously too different from the null (e.g., a signal with huge amplitude, say, 100), then the power for sure will go up to one, which is not very meaningful. Thus it is important to strike a balance between appropriate alternative hypotheses and their practicality. Further, it should be noted that in the present study the values computed from the exponential part of last term of Equation (36) were extremely small, which could be the reason the distance did not affect the power of Y_{max} as much as the other effect. Since it is obvious that the amplitude plays a dominant role in determining the norm, the future research can focus on the other parameters, such as $\boldsymbol{t_1}$, $\boldsymbol{t_2}$, and σ_0, and see what combination will yield larger values for the exponential term.

The current study only dealt with two signal in two-dimensional images. Future study may consider generalize the result to higher dimensions. Or it may generalize the result to more than two signals. As mentioned in Chapter III, Shafie (2014) derived the form of the likelihood ratio test statistic for testing n signals. The current study only concentrated on two signals. Future research may study the behavior and power of the likelihood ratio test statistic when testing n signals.

Closing Remarks

The research presented here provides some insights about the new test statistic Y_{max} when detecting two signals. As the first simulation study attempting to address the sensitivity of Y_{max}, Y_{max} shows promising properties. Although when comparing to other conventional test statistics, such as X^2_{max} or X_{max}, the limited evidence presented in this study remains inconclusive regarding the advantages of

using Y_{max} for signal detection. Another concern is the computation burden. Nonetheless, before making any conclusion about this new, theory-based likelihood ratio test statistic, further investigation is required. Future research may use different algorithms to approximate the probability of Y_{max}, and employ other conditions to generate signals to examine the power of Y_{max}.

REFERENCES

Adler, R. J. (1981). *The geometry of random fields* (1st ed.). New York: Wiley.

Adler, R. J., Bartz, K., & Kou, S. C. (2011). *Estimating thresholding levels for random fields via Euler characteristics.* Unpublished manuscript.

Adler, R. J., Blanchet, J. H., & Liu, J. (2012). Efficient Monte Carlo for high excursions of Gaussian random fields. *The Annals of Applied Probability*, *22*(3), 1167–1214.

Ashby, F. G. (2011). *Statistical analysis of fMRI data.* Cambridge, MA: MIT Press.

Banich, M. (2010). Magnetic resonance imaging. In *The corsini encyclopedia of psychology* (pp. 1–2). Wiley.

Bender, R., & Lange, S. (2001). Adjusting for multiple testing: When and how? *Journal of Clinical Epidemiology*, *54*, 343–349.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Bennett, C. M., Baird, A. A., Miller, M. B., & Wolford, G. L. (2010). Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: An argument for proper multiple comparisons correction. *Journal of Serendipitous and Unexpected Results*, *1*(1), 1–5.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience*, *4*(4), 417–422. doi: 10.1093/scan/nsp053

Brett, M., Penny, W., & Kiebel, S. (2003). An introduction to random field theory. In J. Ashburner, K. Friston, & W. Penny (Eds.), *Human brain function* (pp. 1–23). London, UK: Academic Press.

Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical science*, *16*(2), 101–117.

Cao, J., & Worsley, K. J. (1999). The detection of local shape changes via the geometry of Hotelling's T^2 fields. *The Annals of Statistics*, *27*(3), 925–942.

Cao, J., & Worsley, K. J. (2001). Applications of random fields in human brain mapping. *Spatial Statistics: Methodological Aspects and Applications*, *159*, 170–182.

Friston, K. J., Frith, C., Liddle, P., & Frackowiak, R. (1991). Comparing functional (PET) images: The assessment of significant change. *Journal of Cerebral Blood Flow & Metabolism*, *11*(4), 690–699.

Friston, K. J., Holmes, A., Poline, J.-B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: levels of inference and power. *NeuroImage*, *4*(3), 223–235.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*(2), 189–211. doi: 10.1080/19345747.2011.618213

Hartvig, N. (1999). *A stochastic geometry model for fMRI data* (Tech. Rep.). Department of Theoretical Statistics, University of Aarlus.

Huettel, S. A., Song, A. W., & McCarthy, G. (2008). *Functional magnetic resonance imaging* (2nd ed.). Sunderland, MA: Sinauer Associates.

Lazar, N. (2008). *The statistical analysis of functional MRI data*. New York: Springer.

Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1003–1037. doi: 10.1098/rstb.2002.1114

Longuet-Higgins, M. S. (1952). On the statistical distributions of sea waves. *Journal of Marine Research*, *11*(3), 245–266.

Marchini, J., & Presanis, A. (2004). Comparing methods of analyzing fMRI statistical parametric maps. *NeuroImage*, *22*(3), 1203–1213. doi: 10.1016/j.neuroimage.2004.03.030

Maydeu-Olivares, A., & Brown, G. (2013). Modeling fMRI data: Challenges and opportunities. *Psychometrika*, *78*(2), 240–242.

McIntosh, A. R. (2010). Moving between functional and effective connectivity. In *Analysis and function of large-scale brain networks* (pp. 15–24). Washington, DC: Society for Neuroscience.

Moran, J. M., & Zaki, J. (2013). Functional neuroimaging and psychology: What have you done for me lately? *Journal of cognitive neuroscience*, *25*(6), 834–842.

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, *12*(5), 419–446.

Oehlert, G. W. (2000). *A first course in design and analysis of experiments.* New York: Freeman and Company.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(24), 9868–9872.

Pauling, L., & Coryell, C. D. (1936). The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxyhemoglobin. *Proceedings of the National Academy of Sciences*, *22*(4), 210–216.

Petersson, K. M., Nichols, T. E., Poline, J.-b., & Holmes, A. P. (1999). Statistical limitations in functional neuroimaging ii. signal detection and statistical inference. *Philosophical Transactions: Biological Sciences*, *354*(1387), 1261–1281.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis.* New York, NY: Cambridge University Press.

Poline, J.-B., & Mazoyer, B. M. (1994). Enhanced detection in brain activation maps using a multifiltering approach. *Journal of cerebral blood flow and metabolis*, *14*(4), 639–642.

Rodenfeld, A., & Kak, A. (1982). *Digital picture processing.* New York: Academic Press.

Rohani, M. F. (2003). *Bayesian approach to the Gaussian scale space random fields.* Unpublished doctoral dissertation. Shahid Beheshti University, Tehran.

Rohani, M. F., Shafie, K., & Noorbaloochi, S. (2006). A Bayesian signal detection procedure for scale-space random fields. *The Canadian Journal of Statistics*, *34*(2), 311–325.

Schweitzer, J. (2010). Neuroimaging. In *The corsini encyclopedia of psychology* (pp. 1–3). Wiley.

Shafie, K. (2014). *Likelihood ratio test for the multiple signals model for signal detection from fMRI brain images* (Tech. Rep.). Department of Applied Statistics and Researcch Methods, University of Northern Colorado.

Shafie, K., Sigal, B., Siegmund, D., & Worsley, K. J. (2003). Rotation space random fields with an application to fmri data. *The Annals of Statistics*, *31*(6), 1732–1771.

Siegmund, D., & Worsley, K. (1995). Testing for a signal with unknown location and scale in a stationary guassian random field. *The Annals of Statistics*, *23*(2), 608–639.

Stroman, P. W. (2011). *Essentials of functional MRI.* Boca Raton, FL: CRC Press.

Worsley, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2, f and t fields. *Advances in Applied Probability*, *26*(1), 13–42.

Worsley, K. J. (1996). The geometry of random images. *Chance*, *9*(1), 27–40.

Worsley, K. J. (2001). Testing for signals with unknown location and scale in a χ^2 random field, with an application to fMRI. *Advances in Applied Probability*, 773–793.

Worsley, K. J. (2002). Random field, Gaussian. In *Encyclopedia of environmetrics* (1st ed., pp. 1674–1676). Wiley.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of cerebral blood flow and metabolism*, *12*(6), 900–918.

Worsley, K. J., Marrett, S., Neelin, P., & Evans, A. (1996). Searching scale space for activation in PET images. *Human brain mapping*, *4*(1), 74–90.

Yaglom, A. (1957). Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory of Probability and Its Applications*, *2*(3), 273–320.

APPENDIXA

R CODE FOR MONTE CARLO SIMULATION

Code for Simulating Critical Values with Fixed Kernel

```
##### Parameters

N <- 64
rep <- 5000
s <- round(exp(seq(log(0.02),log(.2),length.out=10)), digits=4)
aa <- 1    # Control for sigma values (1-10)

##### FUNCTION(noise.generate) ######

noise.generate <- function(N,sigma){

  noise <- matrix(rnorm(N*N),N,N)  # white noise
  x <- ((row(noise)-1)/(N-1))-.5   # Range [0,1] & center the kernel
  y <- ((col(noise)-1)/(N-1))-.5
  Q <- exp(-0.5*(x^2+y^2)/sigma^2) # smoothing kernel
  filter <- Q/sqrt(sum(Q^2))      # normalizing
  ffilter <- Mod(fft(filter))      # FFT
  fdata <- fft(noise)
  sz <- Re(fft(fdata*ffilter,inverse=T))/N/N # Smoothing

  return(sz)

  }

##### FUNCTION(covariance) #####

covariance <- function(N,sigma){
  mm <- matrix(0,N,N)
  x <- ((row(mm)-1)/(N-1))
  y <- ((col(mm)-1)/(N-1))
  cor1 <- exp(-.25/sigma^2*outer(x, x, FUN="-")^2)
  cor2 <- exp(-.25/sigma^2*outer(y, y, FUN="-")^2)
  corr <- cor1*cor2

  return(corr)

 }

##### FUNCTION(replicate)

replicate <- function(rep,N,sigma){
  out <- matrix(0,nrow=rep,ncol=3)
  for (i in 1:rep){
```

```
    sz <- noise.generate(N,sigma)
    Xmax <- max(sz)
    X2max <- max(sz^2)
    test <- (1/(1-R^2))*(outer(sz^2,sz^2,FUN="+")
            -2*outer(sz,sz,FUN="*")*R)
    Ymax <- max(test[is.finite(test)], na.rm=T)
    out[i,1] <- Ymax
    out[i,2] <- X2max
    out[i,3] <- Xmax
 }
  return(out)
}

#####

result <- matrix(0,nrow=1,ncol=7)
colnames(result) <- c("sigma", "Ymax95", "Ymax99",
                "X2max95", "X2max99", "Xmax95", "Xmax99")

sigma <- s[aa]
R <- covariance(N, sigma) # Correlation matrix
output <- replicate(rep, N, sigma)
result[,2:3] <- quantile(output[,1], c(.95,.99))  # Ymax
result[,4:5] <- quantile(output[,2], c(.95,.99))  # X2max
result[,6:7] <- quantile(output[,3], c(.95,.99))  # Xmax
result[,1] <- sigma
```

Code for Generating Power with Fixed Kernel

```
install.packages('iterpc')
install.packages('reshape')
library(iterpc)
library(reshape)

##### Parameters

N <- 64 # Size of the image
rep <- 5000
aa <- 1     # Control for sigma values (1-10)
bb <- c(1,2,3)     # Control for condition number (1-36)

##### Create a matrix for 36 schemes (p.29)

xis <- c(.5,2,4) # Levels of magnitude (xi's)
mag <- getall(iterpc(table(xis), 2, replace=TRUE)) # Combinations of xi's
locale1 <- matrix(c(.04,0,-.05,0,.17,-.17,-.17,.17,.45,-.45,-.45,.45)
                  ,3,4, byrow=T)
locale2 <- matrix(c(.09,0,-.1,0,.2,-.2,-.2,.2,.41,-.41,-.41,.41)
                  ,3,4, byrow=T)
sig2 <- expand.grid.df(data.frame(xi=mag),data.frame(t=locale1))
sig5 <- expand.grid.df(data.frame(xi=mag),data.frame(t=locale2))
condition <- rbind(sig2,sig5)
sigma0 <- rep(c(.02,.04),each=18)
ID <- c(1:36)
condition <- cbind(condition,sigma0,ID)

##### Create a matrix with sigma & critical values for Ymax & Xmax

s <- round(exp(seq(log(0.02),log(.2),length.out=10)), digits=4)
Y.cv95 <- c(31.48850633,30.03230504,28.56379799,26.73200706,25.22117191,
 22.96604318,21.1920527,19.47512097,18.00737539,16.26957716)
Y.cv99 <- c(35.64601899,34.45080549,32.68962761,31.13986208,30.49064484,
 28.43327152,26.19818472,26.00416054,23.59683928,22.67088214)
X2.cv95 <- c(18.18687578,17.77512965,16.74117914,15.47643558,14.55024733,
 13.90980734,12.65096897,11.29799853,10.16492417,9.223819999)
X2.cv99 <- c(21.34550049,20.8186419,19.87925651,19.15277943,18.11621129,
 17.30650789,15.98955291,14.70326492,13.80137539,12.88002099)
X.cv95 <- c(4.078552141,3.999744587,3.877193811,3.774215661,3.646633203,
 3.460923328,3.304718338,3.181804798,2.894145748,2.748313766)
X.cv99 <- c(4.485598187,4.368285249,4.324698613,4.159435617,4.085898287,
 3.882812423,3.835473701,3.642527524,3.449664211,3.36678417)
cv <- cbind(s,Y.cv95,X2.cv95,X.cv95,Y.cv99,X2.cv99,X.cv99)
```

```
##### Function(signal.generate)

signal.generate <- function(N,xi1,xi2,t11,t12,t21,t22,sigma0,sigma){
  noise <- matrix(rnorm(N*N),N,N)  # white noise
  x <- ((row(noise)-1)/(N-1))-.5   # Range [0,1] & center the kernel
  y <- ((col(noise)-1)/(N-1))-.5
  # 1st signal (already smoothed)
  mu1 <- (2*xi1*sigma*sigma0)/(sigma^2+sigma0^2)*
          exp(-0.5*((x-t11)^2+(y-t12)^2)/(sigma^2+sigma0^2))
  # 2nd signal (already smoothed)
  mu2 <- (2*xi2*sigma*sigma0)/(sigma^2+sigma0^2)*
          exp(-0.5*((x-t21)^2+(y-t22)^2)/(sigma^2+sigma0^2))
  Q <- exp(-0.5*(x^2+y^2)/sigma^2)   # smoothing kernel
  filter <- Q/sqrt(sum(Q^2))      # normalizing
  ffilter <- Mod(fft(filter))      # FFT
  fdata <- fft(noise)
  sz <- Re(fft(fdata*ffilter,inverse=T))/N/N # Smoothing
  sz <- sz + mu1 + mu2

  return(sz)
}

##### FUNCTION(covariance) #####

covariance <- function(N,sigma){
  mm <- matrix(0,N,N)
  x <- ((row(mm)-1)/(N-1))
  y <- ((col(mm)-1)/(N-1))
  cor1 <- exp(-.25/sigma^2*outer(x, x, FUN="-")^2)
  cor2 <- exp(-.25/sigma^2*outer(y, y, FUN="-")^2)
  corr <- cor1*cor2

  return(corr)
}

##### Function(replicate) #####
replicate <- function(rep,N,xi1,xi2,t11,t12,t21,t22,sigma0,sigma){
  out <- matrix(0,nrow=rep,ncol=3)
  for (i in 1:rep){
    sz <- signal.generate(N,xi1,xi2,t11,t12,t21,t22,sigma0,sigma)
    Xmax <- max(sz)
    X2max <- max(sz^2)
    test <- (1/(1-R^2))*(outer(sz^2,sz^2,FUN="+")
            -2*outer(sz,sz,FUN="*")*R)
```

```
    Ymax <- max(test[is.finite(test)], na.rm=T)
    out[i,1] <- Ymax
    out[i,2] <- X2max
    out[i,3] <- Xmax
  }
   return(out)
}

##### Final step #####

output <- matrix(0,nrow=length(bb),ncol=15)
colnames(output) <- c("sigma","xi1","xi2","t11","t12","t21","t22",
 "sigma0","YmaxPower95","X2maxPower95","XmaxPower95",
 "YmaxPower99","X2maxPower99","XmaxPower99","ID")

sigma <- cv[aa,1]          # Define sigma
R <- covariance(N, sigma)  # Calculate the covariance matrix

for (j in 1:length(bb)){  # Loop for the condition(s)
  xi1 <- condition[bb[j],1]
  xi2 <- condition[bb[j],2]
  t11 <- condition[bb[j],3]
  t12 <- condition[bb[j],4]
  t21 <- condition[bb[j],5]
  t22 <- condition[bb[j],6]
  sigma0 <- condition[bb[j],7]

  out <- replicate(rep,N,xi1,xi2,t11,t12,t21,t22,sigma0,sigma)

  output[j,1] <- sigma
  output[j,2] <- xi1
  output[j,3] <- xi2
  output[j,4] <- t11
  output[j,5] <- t12
  output[j,6] <- t21
  output[j,7] <- t22
  output[j,8] <- sigma0
  output[j,9] <- sum(out[,1] > cv[aa,2])/nrow(out)
  output[j,10] <- sum(out[,2] > cv[aa,3])/nrow(out)
  output[j,11] <- sum(out[,3] > cv[aa,4])/nrow(out)
  output[j,12] <- sum(out[,1] > cv[aa,5])/nrow(out)
  output[j,13] <- sum(out[,2] > cv[aa,6])/nrow(out)
  output[j,14] <- sum(out[,3] > cv[aa,7])/nrow(out)
  output[j,15] <- condition[bb[j],8]
 }   #The end of loop(j)
```

Code for Simulating Critical Values in Scale Space

```
##### Parameters

N <- 64
rep <- 5000
s <- round(exp(seq(log(0.02),log(.2),length.out=10)), digits=4);

##### Create the Gaussian kernel for all sigmas

dd <- matrix(0,N,N)
x <- ((row(dd)-1)/(N-1))-.5        # Range [0,1] & center the kernel
y <- ((col(dd)-1)/(N-1))-.5

kernel <- array(0, c(length(s),N,N))      # Save the kernels
for (i in 1:length(s)){
  Q <- exp(-0.5*(x^2+y^2)/s[i]^2)  # smoothing kernel
  filter <- Q/sqrt(sum(Q^2))      # normalizing
  ffilter <- Mod(fft(filter))      # FFT

  kernel[i, , ] <- ffilter
}

##### FUNCTION(covariance) #####

covariance <- function(N,sigma){
  mm <- matrix(0,N,N)
  xx <- ((row(mm)-1)/(N-1))
  yy <- ((col(mm)-1)/(N-1))
  cor1 <- exp(-.25/sigma^2*outer(xx, xx, FUN="-")^2)
  cor2 <- exp(-.25/sigma^2*outer(yy, yy, FUN="-")^2)
  corr <- cor1*cor2

  return(corr)

  }

##### Create the correlation matrix for all sigma

  covar <- array(0, c(length(s),N,N,N,N))
  for (i in 1:length(s)){
    covar[i, , , , ] <- covariance(N, sigma=s[i])
```

```
  }

##### FUNCTION(noise.generate) ######

  noise.generate <- function(N, ffilter){
    noise <- matrix(rnorm(N*N),N,N)  # white noise
    fdata <- fft(noise)
    sz <- Re(fft(fdata*ffilter,inverse=T))/N/N # Smoothing

    return(sz)

  }

##### FUNCTION(maximum)

output <- matrix(0, nrow=rep, ncol=6)
colnames(output) <- c("sigmaY","Ymax","sigmaX2",
                      "X2max","sigmaX","Xmax")

for (j in 1:rep){
  loop <- matrix(0, nrow=length(s), ncol=4)

  for (i in 1:length(s)){
    sz <- noise.generate(N, ffilter=kernel[i, , ])
    Xmax <- max(sz)
    X2max <- max(sz^2)
    R <- covar[i, , , , ]
    test <- (1/(1-R^2))*(outer(sz^2,sz^2,FUN="+")
             -2*outer(sz,sz,FUN="*")*R)
    Ymax <- max(test[is.finite(test)], na.rm=T)

    loop[i,1] <- s[i]
    loop[i,2] <- Ymax
    loop[i,3] <- X2max
    loop[i,4] <- Xmax
    }  # The end of loop (i)

  output[j,1:2] <- loop[which.max(loop[,2]),c(-3,-4)] # Maximum of Ymax
  output[j,3:4] <- loop[which.max(loop[,3]),c(-2,-4)] # Maximum of X2max
  output[j,5:6] <- loop[which.max(loop[,4]),c(-2,-3)] # Maximum of Xmax

  } # The end of loop (j)

result <- matrix(0, nrow=1, ncol=6)
colnames(result) <- c("Ymax95","Ymax99","X2max95",
```

```
                      "X2max99","Xmax95","Xmax99")
result[,1:2] <- quantile(output[,2], c(.95,.99))
result[,3:4] <- quantile(output[,4], c(.95,.99))
result[,5:6] <- quantile(output[,6], c(.95,.99))
```

Code for Generating Power in Scale Space

```
install.packages('iterpc')
install.packages('reshape')
library(iterpc)
library(reshape)

##### Parameters

N <- 64 # Size of the image
rep <- 5
s <- round(exp(seq(log(0.02),log(.2),length.out=10)), digits=4)
bb <- 1     # Control for condition number (1-36)

##### Create a matrix for 36 schemes (p.29)

xis <- c(.5,2,4)
mag <- getall(iterpc(table(xis), 2, replace=TRUE))
locale1 <- matrix(c(.04,0,-.05,0,.17,-.17,-.17,.17,.45,-.45,-.45,.45)
           ,3,4, byrow=T)
locale2 <- matrix(c(.09,0,-.1,0,.2,-.2,-.2,.2,.41,-.41,-.41,.41)
           ,3,4, byrow=T)
sig2 <- expand.grid.df(data.frame(xi=mag),data.frame(t=locale1))
sig5 <- expand.grid.df(data.frame(xi=mag),data.frame(t=locale2))
condition <- rbind(sig2,sig5)
sigma0 <- rep(c(.02,.04),each=18)
ID <- c(1:36)
condition <- cbind(condition,sigma0,ID)

##### Create a vector for critical values
cv <- c(34.09486474,20.32286359,4.361468703,
        38.5595987,23.5223987,4.706276293)
colnames(cv) <- c("Y.cv95","X2.cv95","X.cv95",
                  "Y.cv99","X2.cv99","X.cv99")

##### Create the Gaussian kernel for all sigmas

dd <- matrix(0,N,N)          # Create an N X N matrix for x and y
x <- ((row(dd)-1)/(N-1))-.5  # Range [0,1] & center the kernel
y <- ((col(dd)-1)/(N-1))-.5

kernel <- array(0, c(length(s),N,N))   # Save the kernels
for (i in 1:length(s)){
  Q <- exp(-0.5*(x^2+y^2)/s[i]^2)  # smoothing kernel
  filter <- Q/sqrt(sum(Q^2))    # normalizing
```

```
  ffilter <- Mod(fft(filter))    # FFT

  kernel[i, , ] <- ffilter
}

##### Function(signal.noise)

signal.noise<-function(N,ffilter,xi1,xi2,t11,t12,t21,t22,sigma0,sigma){
  # 1st signal (already smoothed)
  mu1 <- (2*xi1*sigma*sigma0)/(sigma^2+sigma0^2)
          *exp(-0.5*((x-t11)^2+(y-t12)^2)/(sigma^2+sigma0^2))
  # 2nd signal (already smoothed)
  mu2 <- (2*xi2*sigma*sigma0)/(sigma^2+sigma0^2)
          *exp(-0.5*((x-t21)^2+(y-t22)^2)/(sigma^2+sigma0^2))
  mu <- mu1+mu2

  noise <- matrix(rnorm(N*N),N,N)  # white noise
  fdata <- fft(noise)
  sz <- Re(fft(fdata*ffilter,inverse=T))/N/N # Smoothing
  sz <- sz + mu

  return(sz)
}

##### FUNCTION(covariance) #####

covariance <- function(N,sigma){
  mm <- matrix(0,N,N)
  xx <- ((row(mm)-1)/(N-1))
  yy <- ((col(mm)-1)/(N-1))
  cor1 <- exp(-.25/sigma^2*outer(xx, xx, FUN="-")^2)
  cor2 <- exp(-.25/sigma^2*outer(yy, yy, FUN="-")^2)
  corr <- cor1*cor2

  return(corr)
}

##### Create the correlation matrix for all sigma

covar <- array(0, c(length(s),N,N,N,N))
for (i in 1:length(s)){
  covar[i, , , , ] <- covariance(N, sigma=s[i])
}

##### Final step #####
```

```
xi1 <- condition[bb,1] # Define the parameters
xi2 <- condition[bb,2]
t11 <- condition[bb,3]
t12 <- condition[bb,4]
t21 <- condition[bb,5]
t22 <- condition[bb,6]
sigma0 <- condition[bb,7]

out <- matrix(0, nrow=rep, ncol=6)
colnames(out) <- c("sigmaY","Ymax","sigmaX2","X2max","sigmaX","Xmax")

for (i in 1:rep){
  loop <- matrix(0, nrow=length(s), ncol=4)
  for (j in 1:length(s)){

    sigma <- s[j]
    ffilter <- kernel[j, , ]
    R <- covar[j, , , , ]

    sz <- signal.noise(N,ffilter,xi1,xi2,t11,t12,t21,t22,sigma0,sigma)
    Xmax <- max(sz)
    X2max <- max(sz^2)
    test <- (1/(1-R^2))*(outer(sz^2,sz^2,FUN="+")
            -2*outer(sz,sz,FUN="*")*R)
    Ymax <- max(test[is.finite(test)], na.rm=T)

    loop[j,1] <- sigma
    loop[j,2] <- Ymax
    loop[j,3] <- X2max
    loop[j,4] <- Xmax
  } # The end of loop(j)

  out[i,1:2] <- loop[which.max(loop[,2]),c(-3,-4)] # Maximum of Ymax
  out[i,3:4] <- loop[which.max(loop[,3]),c(-2,-4)] # Maximum of X2max
  out[i,5:6] <- loop[which.max(loop[,4]),c(-2,-3)] # Maximum of Xmax

} # The end of loop(i)

result <- matrix(0,nrow=length(bb),ncol=14)
colnames(result) <- c("xi1","xi2","t11","t12","t21","t22","sigma0",
  "YmaxP95","X2maxP95","XmaxP95","YmaxP99","X2maxP99","XmaxP99","ID")

result[ ,1] <- xi1                          # xi1
result[ ,2] <- xi2                          # xi2
```

```
result[ ,3] <- t11                          # t11
result[ ,4] <- t12                          # t12
result[ ,5] <- t21                          # t21
result[ ,6] <- t22                          # t22
result[ ,7] <- sigma0                       # sigma0
result[ ,8] <- sum(out[,2] > cv[1])/rep
result[ ,9] <- sum(out[,4] > cv[2])/rep
result[ ,10] <- sum(out[,6] > cv[3])/rep
result[ ,11] <- sum(out[,2] > cv[4])/rep
result[ ,12] <- sum(out[,4] > cv[5])/rep
result[ ,13] <- sum(out[,6] > cv[6])/rep
result[ ,14] <- condition[bb,8]             # ID
```