

DISSERTATION

CONFIDENCE REGIONS FOR LEVEL CURVES AND A LIMIT THEOREM  
FOR THE MAXIMA OF GAUSSIAN RANDOM FIELDS

Submitted by

Joshua French

Department of Statistics

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2009

UMI Number: 3385155

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3385155

Copyright 2009 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

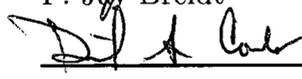
COLORADO STATE UNIVERSITY

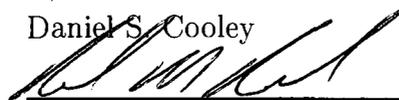
July 9, 2009

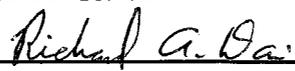
WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY JOSHUA FRENCH ENTITLED CONFIDENCE REGIONS FOR LEVEL CURVES AND A LIMIT THEOREM FOR THE MAXIMA OF GAUSSIAN RANDOM FIELDS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

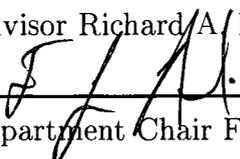
Committee on Graduate Work

  
\_\_\_\_\_  
F. Jay Breidt

  
\_\_\_\_\_  
Daniel S. Cooley

  
\_\_\_\_\_  
Robin M. Reich

  
\_\_\_\_\_  
Advisor Richard A. Davis

  
\_\_\_\_\_  
Department Chair F. Jay Breidt

## ABSTRACT OF DISSERTATION

### CONFIDENCE REGIONS FOR LEVEL CURVES AND A LIMIT THEOREM FOR THE MAXIMA OF GAUSSIAN RANDOM FIELDS

One of the most common display tools used to represent spatial data is the contour plot. Informally, a contour plot is created by taking a “slice” of a three-dimensional surface at a certain level of the response variable and projecting the slice onto the two-dimensional coordinate-plane. The “slice” at each level is known as a level curve.

Consider a Gaussian random field  $\{Z(s) : s \in \mathcal{D}\}$ , where  $s$  is the location in the continuous two-dimensional region of interest  $\mathcal{D} \subset \mathbb{R}^2$ . The level curve for the process  $Z$  at level  $u$  is defined to be  $I_u = \{s : Z(s) = u\}$ . From the observed data  $z(s_1), \dots, z(s_N)$ , one can predict  $\hat{Z}(s)$  for any location  $s \in \mathcal{D}$  using kriging or some other procedure, and then construct the estimated level curve  $\hat{I}_u = \{s : \hat{Z}(s) = u\}$  as an approximation of  $I_u$ . We present two methods for constructing confidence regions for the level curves of a contour plot.

The first method proposed is an extension of Lindgren and Rychlik (1995) and Wameling (2003b) based on level crossings. A series of rectangular confidence regions are constructed along  $\hat{I}_u$  which should individually intersect the true level curve with high confidence. The boxes extend in directions perpendicular to the estimated level curve and the widths of the boxes are chosen so that the edge of each box touches the neighboring box and there are no gaps between the boxes along the estimated level curve. The heights of the boxes are chosen by simulating realizations of

$Z(s)|z(s_1), \dots, z(s_N)$ , and then taking the appropriate quantiles of the set of nearest level crossings for the realizations. The heights of the boxes give insight into the approximate distance between the estimated level curve and the true level curve.

The second method constructs a confidence region for  $I_u$  through hypothesis testing, adjusting the critical value to control the simultaneous Type I error rate. Our goal is to construct a confidence region  $S$  for the true level curve such that  $\mathbb{P}(I_u \subseteq S) \geq 1 - \alpha$ . Instead of finding  $S$  directly, we adopt a different approach and find a set  $R$  which does not intersect  $I_u$  with high confidence, so that  $\mathbb{P}(\{R \cap I_u\} = \emptyset) \geq 1 - \alpha$ . Consequently, the set  $S = R^c$  will satisfy our goal since  $\mathbb{P}(I_u \subseteq R^c) \geq 1 - \alpha$ . The region  $R$  is constructed by testing  $H_0 : Z(s) = u$  versus  $H_a : Z(s) \neq u$ , and taking  $R$  to be the union of all  $s$  for which we conclude that  $Z(s) \neq u$ . Using kriging, we construct a test statistic which has a standard normal distribution. The critical value is adjusted to control the simultaneous Type I error rate through empirical simulation of the test statistic.

We conclude by introducing a limit theorem for the distribution of the maxima of a triangular sequence of stationary Gaussian random fields on an  $n \times n$  lattice. The result is an extension of the work presented by Hsing et al. (1996) to two dimensions. The result was motivated by the desire to control the simultaneous Type I error rate of hypothesis tests at locations on an  $n \times n$  lattice where the test statistics are Gaussian and correlated. Under certain dependence and limiting conditions we show that the maximum of the random fields exhibits extremal clustering in the limit. Consideration is then given to the use of this result in approximating  $\mathbb{P}(\max_{1 \leq i, j \leq n} Z_{i, j} \leq u)$ , where  $\{Z_{i, j}\}$  is a stationary Gaussian random field on a square  $n \times n$  lattice of equally spaced locations.

Joshua French  
 Department of Statistics  
 Colorado State University  
 Fort Collins, Colorado 80523  
 Summer 2009

## ACKNOWLEDGEMENTS

There are many individuals who deserve recognition for their help in completing this dissertation:

Richard Davis for advising me and guiding me through the dissertation process. I have been the beneficiary of a great deal of wisdom, advice, and instruction as a result. I certainly could not have completed this dissertation without your guidance and assistance.

N. Scott Urquhart for advising me while completing my Master's thesis and supporting me through the STARMAP program. He also provided me with the idea for the subject of this dissertation.

Jay Breidt, Dan Cooley, and Robin Reich for serving on my committee and for their patience, questions, and input.

The STARMAP program and the U.S. Environmental Protection Agency (EPA) for partially supporting this research under the STAR Research Assistance Agreement CR-829095 awarded by the EPA to Colorado State University. This dissertation has not been formally reviewed by the EPA. The views expressed here are solely those of the author and STARMAP, the program they represent. EPA does not endorse any products or commercial services mentioned in this dissertation.

The Swiss Federal Institute of Technology at Lausanne for providing the Swiss Jura data analyzed in Chapter 4.

Last, and most importantly, Julia and Asher for being such faithful supporters during this process. Julia, you have helped me become a better man. Asher, you have made me a very happy father. I love you both.

## DEDICATION

To Julia, Asher, and the child on the way.

## CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Defining a Contour Plot . . . . .	1
1.2	Uncertainty in Contour Plots . . . . .	3
1.3	Existing Methods for Constructing Confidence Regions for Level Curves	5
1.3.1	Confidence Bands for Level Curves of Smooth Gaussian Fields	5
1.3.2	Confidence Bands for a Broader Range of Processes . . . . .	8
1.3.3	Measuring accuracy via individual confidence levels . . . . .	11
1.3.4	Credible Bands for Level Curves . . . . .	12
1.3.5	Related Work . . . . .	13
1.4	Outline of Dissertation . . . . .	14
<b>2</b>	<b>Confidence Boxes</b>	<b>16</b>
2.1	Introduction . . . . .	16
2.2	Interpolation of Spatial Processes . . . . .	18
2.2.1	Characterizing Spatial Variation . . . . .	18
2.2.2	Kriging . . . . .	21
2.3	Details of Constructing Confidence Boxes . . . . .	22
2.4	Simulation Study . . . . .	27
2.4.1	General Description of Simulation Procedure . . . . .	28
2.4.2	Specifics of Simulation Procedure . . . . .	28
2.4.3	Results . . . . .	30
<b>3</b>	<b>Hypothesis Testing</b>	<b>32</b>
3.1	Introduction . . . . .	32
3.2	Proposed Method of Simulation for Estimating $C_\alpha$ . . . . .	36
3.3	Validation of Proposed Method . . . . .	39
3.3.1	General Description of Validation Procedure . . . . .	39
3.3.2	Specifics of Validation Procedure . . . . .	40
3.3.3	Simulation Results for Hypothesis Testing . . . . .	41
3.4	Use of Adler (2008) to Estimate $C_\alpha$ . . . . .	48
3.5	Future Alternative Error Criteria for Consideration . . . . .	51
<b>4</b>	<b>Application and Comparison of Methods</b>	<b>55</b>
4.1	Introduction . . . . .	55
4.2	Application to Simulated Data . . . . .	55
4.3	Application to Jura Data . . . . .	57
4.4	Comparison of the Two Methods . . . . .	63

<b>5</b>	<b>The Asymptotic Distribution of the Maxima of a Triangular Sequence of Gaussian Random Fields</b>	<b>65</b>
5.1	Introduction . . . . .	65
5.2	Motivating Example . . . . .	67
5.3	Setup . . . . .	67
5.4	Main Results . . . . .	68
5.5	Approximation of Maximum on a Fixed Domain . . . . .	73
5.5.1	Structure of Simulation . . . . .	73
5.5.2	Estimating $\tilde{\vartheta}$ . . . . .	75
5.5.3	Covariance Functions and Other Details . . . . .	75
5.5.4	Summary of results . . . . .	78
5.5.5	A Brief Comparison with Adler (2008) . . . . .	79
5.5.6	Overall results . . . . .	82
5.6	Proof of Theorem 5.1 and Complements . . . . .	83
5.6.1	Preliminaries . . . . .	83
5.6.2	Bounding $\mathbb{P}(M_n \leq u_{n^2})$ . . . . .	85
5.6.3	Convergence Results . . . . .	97
5.6.4	Proof of Theorem 6.1 . . . . .	106
<b>6</b>	<b>Conclusion and Future Work</b>	<b>110</b>
<b>A</b>	<b>Appendix: Figures for Chapter 5 Simulations</b>	<b>112</b>
	<b>References</b>	<b>126</b>

## LIST OF TABLES

2.1	Empirical confidence level of confidence boxes. . . . .	30
3.1	Simulation results for spherical covariance function. . . . .	42
3.2	Simulation results for exponential covariance function. . . . .	42
3.3	Empirical confidence levels for Gaussian covariance function. . . . .	43
3.4	Estimates of $C_{\alpha/2}^*$ using a stationary, isotropic approximation of $Z'$ . . . .	49

## LIST OF FIGURES

1.1	Creating a contour plot for a three-dimensional surface. . . . .	2
1.2	Probabilistic tornado graphic. . . . .	4
1.3	Illustrations of the alternative definition of a level curve and $T(s_0)$ . . . .	7
1.4	Confidence bands for the true level curve. . . . .	9
1.5	Finding the closest level crossings. . . . .	11
1.6	An image plot of $p(s) = \mathbb{P}(6 < Z(s) < 12)$ . . . . .	13
2.1	Gap between confidence bands. . . . .	17
2.2	Closeup of confidence boxes extending from the linear approximation of the estimated level curve. . . . .	23
2.3	Angle associated with line segment $L_i$ . . . . .	24
2.4	A graphical display of $B_i(\tau_1, \tau_2)$ . . . . .	25
2.5	Process of constructing confidence boxes. . . . .	26
2.6	Covariance functions of simulation procedure. . . . .	29
3.1	A comparison of the true level curve and realized level curves. . . . .	38
3.2	Confidence region for the true level curve. . . . .	44
3.3	Histograms of the proportion of the true level curve contained in the cor- responding confidence region. . . . .	45
3.4	Comparison of confidence regions constructed for varying degrees of de- pendence. . . . .	47
3.5	Relationship between response level of level curve and precision of confi- dence region. . . . .	48

3.6	Estimates of $C_{.10}$ and $C_{.05}$ . . . . .	50
4.1	Covariance function used to generate simulated data. . . . .	56
4.2	Application of methods to simulated data. . . . .	57
4.3	Map of sampled locations for Swiss Jura data. . . . .	58
4.4	Exploratory plots for Jura data. . . . .	59
4.5	Spatial variation of Jura data. . . . .	61
4.6	Confidence regions for $I_{45}$ for the Jura data. . . . .	62
5.1	The relationship between $s_{i,j}$ and a $3 \times 3$ grid of squares in $\mathcal{D}$ . . . . .	74
5.2	A comparison of spatial dependence for various correlation functions. . .	77
5.3	A comparison of estimated excursion probabilities using two methods. .	81
5.4	A visual representation of $M_{(3,1):(2,6)}^5$ . . . . .	85
5.5	A visual approach to Lemma 5.4. . . . .	86
5.6	The number of variables separated by $l_n$ units in $A$ and $B$ . . . . .	88
A.1	Approximation for Gaussian covariance function with $\phi = 0.1$ . . . . .	113
A.2	Approximation for Gaussian covariance function with $\phi = 0.5$ . . . . .	114
A.3	Approximation for Gaussian covariance function with $\phi = 1$ . . . . .	115
A.4	Approximation for Gaussian covariance function with $\phi = 2$ . . . . .	116
A.5	Approximation for spherical covariance function with $\phi = 0.5$ . . . . .	117
A.6	Approximation for spherical covariance function with $\phi = 1$ . . . . .	118
A.7	Approximation for spherical covariance function with $\phi = 2$ . . . . .	119
A.8	Approximation for spherical covariance function with $\phi = 5$ . . . . .	120
A.9	Approximation for spherical covariance function with $\phi = 10$ . . . . .	121
A.10	Approximation for exponential covariance function with $\phi = 0.1$ . . . . .	122
A.11	Approximation for exponential covariance function with $\phi = 0.5$ . . . . .	123
A.12	Approximation for exponential covariance function with $\phi = 1$ . . . . .	124
A.13	Approximation for exponential covariance function with $\phi = 2$ . . . . .	125

## Chapter 1

### INTRODUCTION

#### 1.1 Defining a Contour Plot

The display of three-dimensional data is an important aspect of spatial data analysis. A common display tool used in the representation of spatial data is the *contour plot*. Informally, a contour plot is created by taking a “slice” of a three-dimensional surface at a certain level of the response variable and projecting the slice onto the two-dimensional coordinate-plane. The “slice” at each level is known as a *level curve*. The process of creating a contour plot is shown in Figure 1.1.

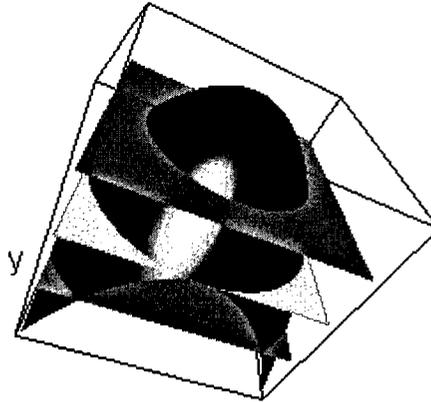
More formally, suppose we have a response surface  $z(x, y)$  and wish to display a contour plot of  $z(x, y)$  corresponding to levels  $u_1, \dots, u_n$  over a region of interest  $\mathcal{D}$ . The contour plot is created by plotting individual level curves  $I_{u_1}, \dots, I_{u_n}$  over the domain  $\mathcal{D}$ , where a level curve  $I_u$  is defined to be

$$I_u = \{(x, y) \in \mathcal{D} : z(x, y) = u\}. \quad (1.1)$$

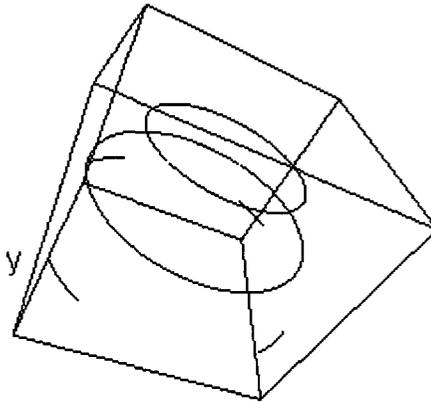
The contour plot is a simple and effective way to understand the behavior of the response variable over the region of interest.

Subsequent discussion will focus on the level curves of a random field  $\{Z(s), s \in \mathcal{D}\}$ , where  $s$  represents location in the two-dimensional region of interest  $\mathcal{D} \subset \mathbb{R}^2$ . Consequently, the level curve  $I_u$  for the response  $Z$  at level  $u$  is defined to be all locations  $s$  such that  $Z(s) = u$ , i.e.,

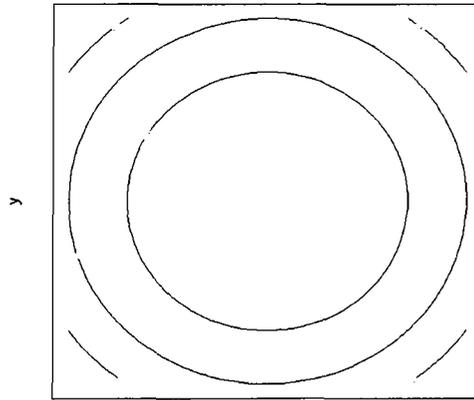
$$I_u = \{s \in \mathcal{D} : Z(s) = u\}. \quad (1.2)$$



x  
(a) Slicing the response surface.



x  
(b) Level curves in three-dimensions.



x  
(c) Level curves projected onto the two-dimensional region of interest.

Figure 1.1: Creating a contour plot for a three-dimensional surface. The response surface is sliced at specified response levels and the slices are projected onto the two-dimensional coordinate plane.

## 1.2 Uncertainty in Contour Plots

There are generally two types of error at work in a contour plot. In a typical spatial setting we observe our response variable  $Z$  at a small number of locations  $s_1, \dots, s_N$ . Because interest lies in the response value at other locations in  $\mathcal{D}$ , the observed responses  $z(s_1), \dots, z(s_N)$  are used to predict  $Z$  throughout the region of interest. Denote the prediction of  $Z$  at location  $s$  by  $\hat{Z}(s)$ . Contour plots are usually constructed from the interpolated response surface  $\{\hat{Z}(s), s \in \mathcal{D}\}$ . The first error type we see in a contour plot is the classical prediction or interpolation error,  $Z(s) - \hat{Z}(s)$ . There will nearly always be a discrepancy between the actual value of  $Z(s)$  and the predicted response value  $\hat{Z}(s) = u$  for  $s \in \hat{I}_u$ . The other related type of error is the discrepancy in the locations of the estimated level curve  $\hat{I}_u = \{s \in \mathcal{D} : \hat{Z}(s) = u\}$  and the true level curve  $I_u = \{s \in \mathcal{D} : Z(s) = u\}$ . This has sometimes been referred to as *spatial* or *horizontal* error (cf. Wameling and Saborowski (2001) and Wameling (2003a)).

In many settings the response value at a specific location is not of primary interest. Instead, researchers may wish to identify regions having a certain characteristic of interest. For example, we may want to find all locations where the response variable is greater than a certain value, or all locations where the response variable is within a certain range of values. Expressed formally, we may wish to find  $\{s \in \mathcal{D} : Z(s) > u\}$  or  $\{s \in \mathcal{D} : u_1 \leq Z(s) \leq u_2\}$ . In both cases, the level curve is the boundary of these regions, so it is of special interest. An example of this is found in the production of weather-related maps in the field of atmospheric science. The Storm Prediction Center in Norman, Oklahoma monitors severe weather across the United States. Short-term forecasts of severe weather and convective weather outlooks are supplied to the public on a regular basis. One of the graphics supplied in these outlooks when appropriate is a “probabilistic tornado graphic” which is simply a contour plot of the probability that a tornado forms within 25 miles of a location. The probabilistic tornado graphic

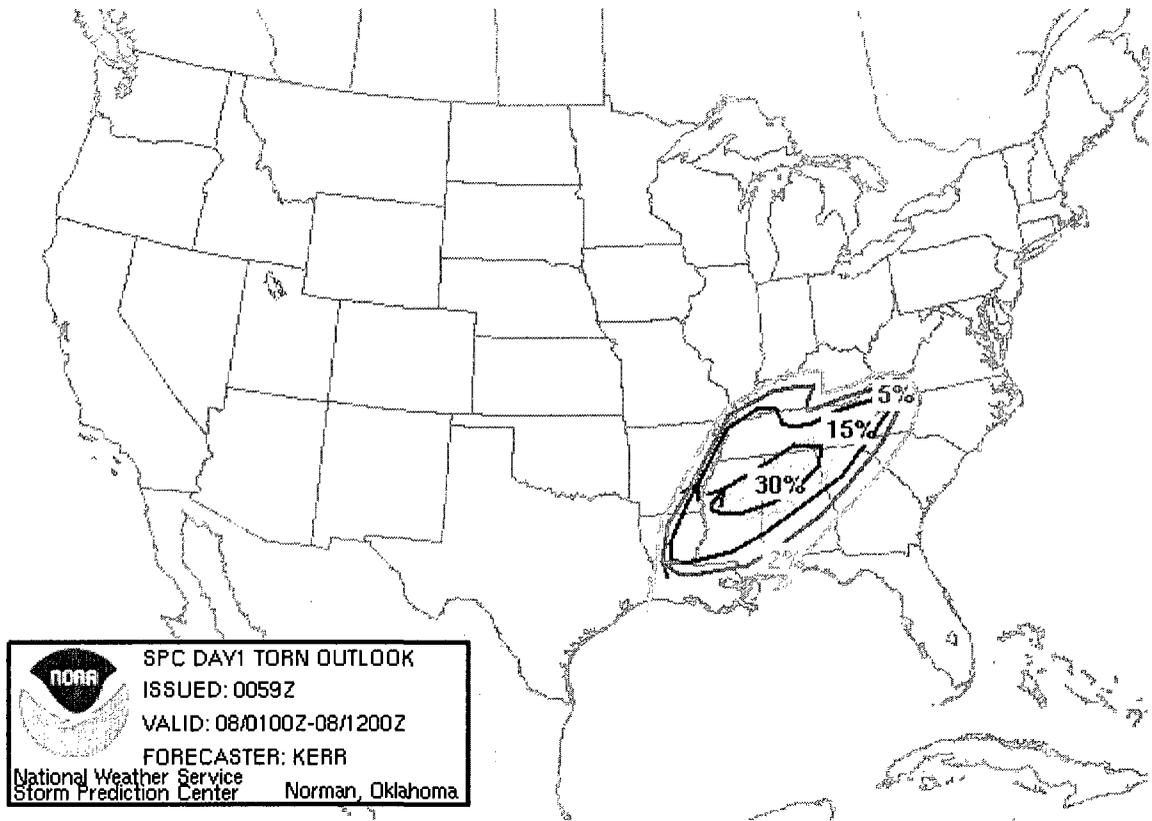


Figure 1.2: Probabilistic tornado graphic. A probabilistic tornado graphic is a contour plot of the probability that a tornado forms within 25 miles of a point. The probabilities have been converted to percentages in the graphic.

for April 8, 2006 (a period of severe weather in the Eastern United States) is shown in Figure 1.2. In this case, researchers are interested in finding regions where the probability is high that a tornado forms within 25 miles of the region. Given the seriousness of having a tornado form in an area unexpectedly, it is of great importance to accurately locate the true level curves of the probability surface in order to provide warnings and alerts to the appropriate agencies. Other disciplines where contour plots are commonly used include meteorology, oceanography, ecology, engineering and many others.

Although contour plots are commonly used to display spatial data, little consideration has been given to the uncertainty of the level curves in contour plots. Lindgren and Rychlik (1995) stated, “The accuracy of level curves seems to have received no attention in the theoretical literature, even if it should be of considerable interest

from a practical point of view.” In the next section we will describe methods currently used to assess the accuracy of contour plots. Our research will supplement the current tools by presenting new approaches to construct confidence regions for the level curves of a contour plot.

### 1.3 Existing Methods for Constructing Confidence Regions for Level Curves

#### 1.3.1 Confidence Bands for Level Curves of Smooth Gaussian Fields

The first theoretical approach for assessing the uncertainty in contour plots appears to have been done by Lindgren and Rychlik (1995). As stated in their paper, their aim was “to present a numerically accurate method to calculate confidence limits for level curves...” To quantify the precision of the estimated contours, they build confidence bands along the estimated contours that individually intersect the unknown true contour with a calculable probability.

Lindgren and Rychlik model the response surface using a Gaussian random field  $\{Z(s), s \in \mathcal{D} \subset \mathbb{R}^2\}$  with  $\mathbb{E}(Z(s)) = 0$  and covariance function

$$C(s, t) = \text{cov}(Z(s), Z(t)), \quad s, t \in \mathbb{R}^2.$$

Conceptually, their method requires no additional assumptions, but to derive specific results, they subsequently assume that  $Z$  is stationary and that the covariance function is isotropic and twice-differentiable. Random measurement error is allowed so that observed values may be realizations of

$$Z_{obs}(s) = Z(s) + \epsilon,$$

where  $\epsilon$  is a Gaussian random variable independent of location with mean zero and variance  $\sigma_\epsilon^2$ . Regardless of whether measurement error is present, confidence limits are constructed for the level curves of the true response process  $Z$ .

Lindgren and Rychlik assume response values  $z(s_1), \dots, z(s_N)$  have been observed at locations  $s_1, \dots, s_N \in \mathcal{D}$ , where  $s_k$  represents the coordinates  $(s_{k,x}, s_{k,y})$ .

Under a squared error loss function, the best linear predictor of  $Z(s)$  given the observed data  $z(s_1), \dots, z(s_N)$  is the simple kriging predictor  $\hat{Z}(s)$ . The covariance between  $\hat{Z}(s)$  and  $\hat{Z}(t)$  is nonstationary and we denote it as  $C_{\hat{Z}}(s, t)$ . The reconstruction error at location  $s$  is given by  $\eta(s) = Z(s) - \hat{Z}(s)$ , where  $\eta(s)$  has mean zero and covariance function  $C_{\hat{Z}}(\cdot, \cdot)$ .

Lindgren and Rychlik make inference on  $\eta(s) + \hat{Z}(s) = Z(s)$  by defining an alternative definition of a level curve using the concept of *level crossings*. We define level crossings and upcrossings following the presentations given by Cramer and Leadbetter (1967) and Lindgren (2004).

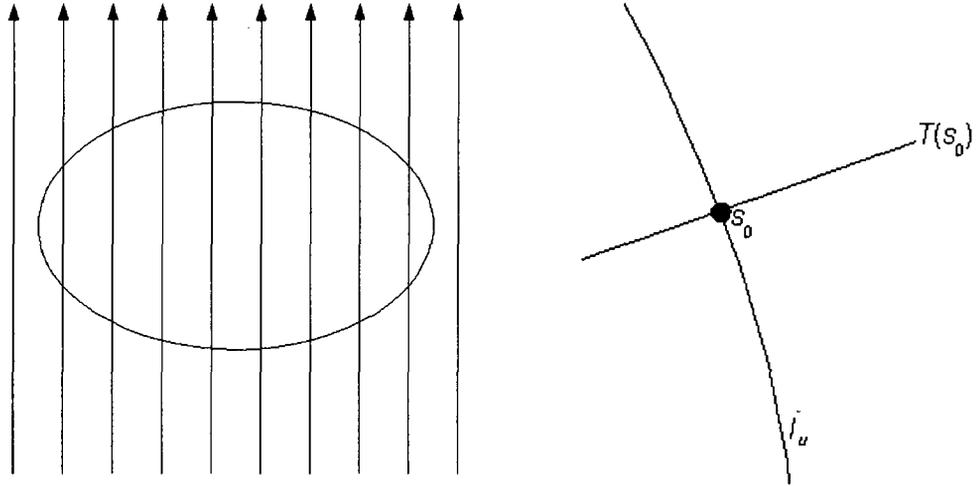
Suppose that  $x(s)$  is a continuous function such that  $x(s)$  is not identically equal to  $u$  in any interval.

**Definition 1.1.**  $x$  is said to have an *upcrossing* of the level  $u$  at  $s_0$  if there exists  $\epsilon > 0$  such that  $x(s) \leq u$  in  $(s_0 - \epsilon, s_0]$ , and  $x(s) \geq u$  in  $[s_0, s_0 + \epsilon)$ .

**Definition 1.2.**  $x$  is said to have a *crossing* of the level  $u$  at  $s_0$  if for every neighborhood of  $s_0$ , there exists  $s_1$  and  $s_2$  such that  $(x(s_1) - u)(x(s_2) - u) < 0$ .

As defined in (1.2), the standard definition of a level curve  $I_u$  is the set of all locations  $s$  such that  $Z(s) = u$ . The alternative definition supplied by Lindgren and Rychlik is that a level curve  $I_u$  is the union of all *crossings* of the level  $u$  by the response surface  $Z$  when traveling along all possible straight lines in the region of interest  $\mathcal{D}$ . A similar definition holds for the level curve  $\hat{I}_u$  of the reconstructed response surface  $\hat{Z}$ . This concept is illustrated in Figure 1.3 (a). The accuracy of the contour plot can be judged by how closely the level crossings along sections of the predicted surface  $\hat{Z}$  represent the level crossings along the sections of the realized surface  $Z$ .

Suppose that  $s_0 \in \hat{I}_u$ . Define  $T(s_0)$  to be a line in  $\mathcal{D}$  going through  $s_0$  at an angle  $\theta_0$  perpendicular to  $\hat{I}_u$  at  $s_0$ . A visual representation of  $T(s_0)$  is shown in Figure 1.3



(a) A level curve can be constructed by taking the union of the level crossings for all possible straight lines in the region of interest. The closed curve is a level curve.

(b)  $T(s_0)$  is a line going through  $s_0 \in \hat{I}_u$  at an angle perpendicular to  $\hat{I}_u$  at  $s_0$ .

Figure 1.3: Illustrations of the alternative definition of a level curve and  $T(s_0)$ .

(b). Any point on the line  $T(s_0)$  can be parameterized by

$$s_0(\tau) = (s_{0,x} + \tau \cos(\theta_0), s_{0,y} + \tau \sin(\theta_0)),$$

where  $\tau \in \mathbb{R}$ . Similarly, the functions  $\hat{Z}$  and  $\eta$  can be respecified in terms of points along  $T(s_0)$  so that

$$\hat{Z}_0(\tau) = \hat{Z}(s_0(\tau))$$

and

$$\eta_0(\tau) = \eta(s_0(\tau)).$$

The goal of Lindgren and Rychlik is to find an interval along  $T(s_0)$  that intersects the true level curve with high probability. This is accomplished using the *first upcrossing intensity*. The first crossing intensity along  $T(s_0)$  in the interval  $[-\tau_1, \tau_2]$  is defined as

$$\tilde{\mu}_0^+(-\tau_1, \tau_2, u) = \mathbb{P}(\eta_0(\tau) + \hat{Z}_0(\tau) \text{ has a } u\text{-upcrossing for } \tau \in [\tau_2, \tau_2 + d\tau_2])$$

$$\text{and } \eta_0(\tau) + \hat{Z}_0(\tau) \neq u \text{ for all } \tau \in [-\tau_1, \tau_2].$$

Thus,  $\tilde{\mu}_0^+(-\tau_1, \tau_2, u)$  is the probability that starting at location  $s_0(-\tau_1)$  and continuing along  $T(s_0)$  in the direction of  $s_0(\tau_2)$ , the first upcrossing of the process  $\eta_0(\tau) + \hat{Z}_0(\tau)$  occurs at  $s_0(\tau_2)$ . Define

$$\tilde{T}(s_0, \alpha) = \{s_0(\tau), -c \leq \tau \leq c : \tilde{\mu}_0^+(-c, c, u) = 1 - \alpha\}.$$

Using the alternate definition of the level curve, we see that  $\mathbb{P}(\tilde{T}(s_0, \alpha) \cap I_u) = 1 - \alpha$ . Numerical methods are used to find the appropriate confidence interval size for each interval individually.

The resulting intervals clarify the amount of error in our estimated level curves  $\hat{I}_u$ . Individually, the probability that each interval intersects the true level curve  $I_u$  is  $1 - \alpha$ . No adjustment is made for multiple comparisons. In general, the widths of our intervals give us insight into the approximate distance between the estimated level curve and the true level curve. Wider intervals indicate that the estimated level curve is more likely to be further from the true level curve. An example of the resulting intervals is shown in Figure 1.4.

### 1.3.2 Confidence Bands for a Broader Range of Processes

Wameling and Saborowski (2001) point out drawbacks of the confidence interval method proposed by Lindgren and Rychlik (1995) for determining the accuracy of contour plots:

- Data are often not normally distributed and positively skewed data sets are more common.
- The confidence intervals produced using this method are limited to symmetric intervals, but it seems likely that the confidence bands should be asymmetric for level curves of extreme values.
- The confidence intervals are constructed individually and do not take into account the multiple comparison problem.

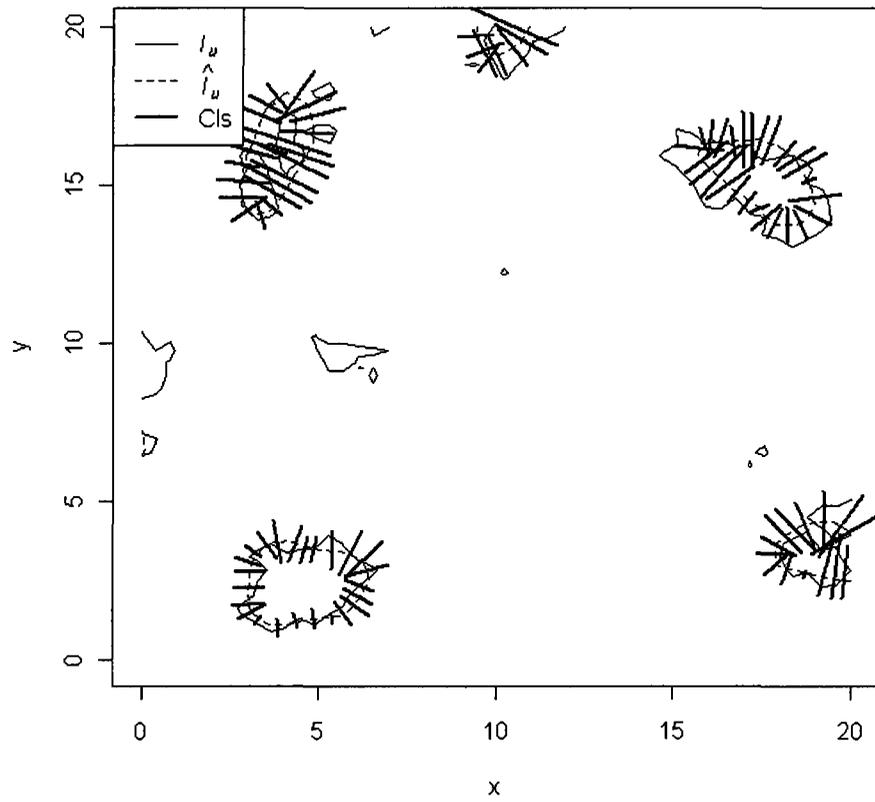


Figure 1.4: Confidence bands for the true level curve. Confidence bands extending perpendicular to  $\hat{I}_u$  should intersect the true level curve  $I_u$  with high probability.

We point out that the first two drawbacks are not necessarily major. Data can often be transformed to approximate normality using a simple transformation. Also, though the ability to construct asymmetric intervals may be useful, it is not clearly shown by Wameling and Saborowski that asymmetric confidence intervals are substantially better than symmetric confidence intervals. The most significant objection is that no adjustment is made for multiple comparisons; however, this is not addressed by Wameling and Saborowski.

In response to the drawbacks mentioned, Wameling and Saborowski propose a method for constructing asymmetric confidence intervals for the true level curve which does not require  $Z$  to be a Gaussian random field through conditional simulation of the process  $\{Z(s) \mid z(s_1), \dots, z(s_N)\}$ . The only explicit assumption used in this method is that an appropriate simulation algorithm exists for approximating realizations of  $\{Z(s) \mid z(s_1), \dots, z(s_N)\}$ .

As in Section 1.3.1, for  $s_0 \in \hat{I}_u$  we define  $T(s_0)$  to be a line in  $\mathcal{D}$  going through  $s_0$  at an angle  $\theta_0$  perpendicular to  $\hat{I}_u$ . Any point on the line  $T(s_0)$  can be parameterized as

$$s_0(\tau) = (s_{0,x} + \tau \cos(\theta_0), s_{0,y} + \tau \sin(\theta_0)).$$

Wameling and Saborowski make use of the fact that as  $T(s_0)$  extends away from  $s_0$  it will intersect the true level curve when it crosses the level  $u$ . The following steps are proposed for constructing a confidence interval for  $I_u$  along  $T(s_0)$ . Simulate  $M$  realizations of  $\tilde{Z}(s) = Z(s) \mid z(s_1), \dots, z(s_N)$  along  $T(s_0)$  and record the crossing closest to  $s_0$ . Denote the location of the closest crossing for the  $i$ th realization as  $t_{\min}^{(i)}$ . This location will correspond to  $\tau_{\min}^{(i)}$  such that

$$t_{\min}^{(i)} = s_0 + \tau_{\min}^{(i)}(\cos \theta, \sin \theta).$$

These steps are illustrated in Figure 1.5. In Figure 1.5 the estimated response  $\hat{Z}(s)$  along  $T(s_0)$  is shown by a dashed line. The estimated response  $\hat{Z}(s)$  crosses the level

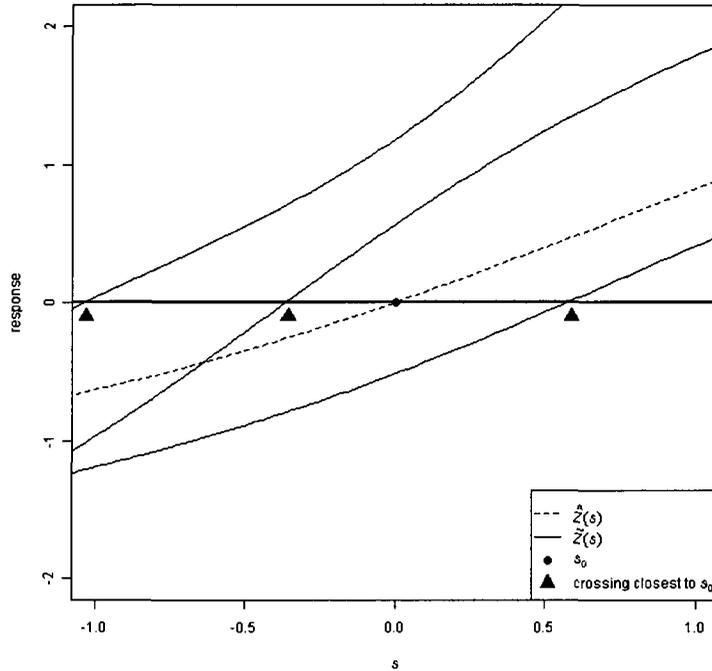


Figure 1.5: Finding the closest level crossings. To construct confidence intervals of the appropriate size we must simulate realizations of  $\tilde{Z}(s) = Z(s) \mid z(s_1), \dots, z(s_N)$  (shown as a solid black line) and find the level crossing closest to  $s_0$ .

zero at  $s_0$ . Three realizations of  $\tilde{Z}(s)$  along  $T(s_0)$  are shown as solid lines. The 0-crossing closest to  $s_0$  for each realization is marked by a triangle. After finding the nearest level crossing  $t_{\min}^{(i)}$  for each realization, compute  $\tau_{\alpha/2}$  and  $\tau_{1-\alpha/2}$ , the  $\alpha/2$  and  $1 - \alpha/2$  quantiles of  $\{\tau_{\min}^{(i)}, i = 1, \dots, M\}$ . The line segment stretching from  $s_0(\tau_{\alpha/2})$  to  $s_0(\tau_{1-\alpha/2})$  should intersect the true level curve with approximate confidence level  $1 - \alpha$ . The resulting intervals will look similar to the ones produced by Lindgren and Rychlik (1995).

### 1.3.3 Measuring accuracy via individual confidence levels

Polfeldt (1999) assesses the accuracy of a contour plot by calculating individual confidence levels throughout the region of interest. Polfeldt assumes that the response process  $Z$  has been observed at locations  $\{s_1, \dots, s_N\} \in \mathcal{D}$ , so that for any location  $s \in \mathcal{D}$ , we are able to predict responses  $\hat{Z}(s)$  and calculate the associated prediction error  $\hat{\sigma}^2(s)$ .

From the prediction surface  $\hat{Z}$  we construct a contour plot for levels  $u_1 < \dots < u_p$ . Ignoring the problems at locations along the boundary of  $\mathcal{D}$ , any point  $s$  will lie between two contours for levels  $u_i$  and  $u_j$  (assume  $u_i < u_j$ ). Implicitly, our belief is that the true response  $Z(s)$  is between the values  $u_i$  and  $u_j$ . Polfeldt measures how well the contour plot represents the distribution of the response process  $Z$  over the domain of interest by calculating the probability

$$p(s) = \mathbb{P}(u_i < Z(s) < u_j).$$

Under the assumption that  $Z$  is Gaussian, this probability can be easily calculated.

Contour plots in which  $p(s)$  tends to be high represent the distribution of the process  $Z$  better than plots in which  $p(s)$  is low. When  $p(s)$  is low, it indicates that it is uncertain whether the level curves  $u_i$  and  $u_j$  are in the “correct” location. i.e., it is not clear whether the level curves  $u_i$  and  $u_j$  correctly separate the locations where  $Z(s) < u_i$  from the locations where  $Z(s) > u_j$ . When  $p(s)$  is high, we can have greater confidence that the estimated level curves accurately represent the behavior of  $Z$  at that location and that  $Z(s)$  takes a value between  $u_i$  and  $u_j$  at  $s$ . An image plot of  $p(s)$  is shown in Figure 1.6; the predictions  $\hat{Z}(s)$  and the prediction error  $\hat{\sigma}^2(s)$  were calculated using simple kriging and the levels under consideration were  $u_1 = 6$  and  $u_2 = 12$ .

### 1.3.4 Credible Bands for Level Curves

Bayarri et al. (2008) use credible bands to assess the accuracy of a contour plot in their paper dealing with risk assessment of natural hazards. Bayarri et al. seek to model volcanic behavior as a function of two input variables. There is special interest in the critical event contour of the input space, which is a level curve separating catastrophic events from benign events.

The response surface is a deterministic function of the input variables calculated by a complex computer model. Because this computation is prohibitively expensive,

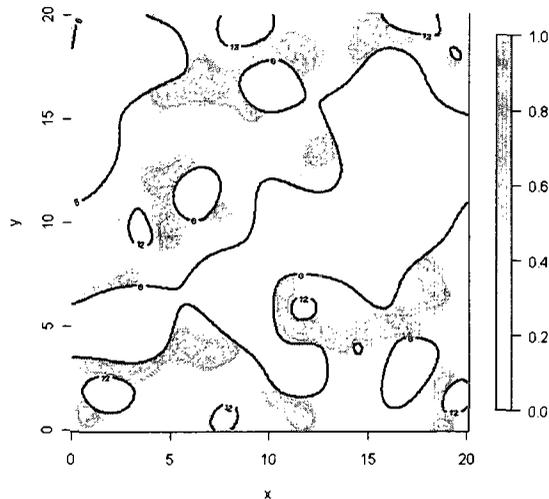


Figure 1.6: An image plot of  $p(s) = \mathbb{P}(6 < Z(s) < 12)$ . The accuracy of a contour plot is better when  $p(s)$  tends to be high over the region of interest.

a Bayesian model is used to approximate the computer model. The true critical event contour is approximated by the critical event contour of the median of the posterior predictive distribution. To reflect the uncertainty in the approximation of the critical event contour, 90% credible bands are provided by finding the critical event contours for the 0.05 and 0.95 quantiles of the posterior predictive distribution. However, Bayarri et al. do not clarify the sense in which these bands are “credible” bands. A typical credible region contains an object of interest with known probability. The credible bands provided by Bayarri et al. reflect the uncertainty of the estimated critical event contour, but it is not clear that these bands contain the true contour with probability 0.90. The credible bands provided in one of the figures actually cross each other; this seems unreasonable if the bands actually contain the true critical event contour with high probability.

### 1.3.5 Related Work

Several papers apply the results previously discussed. Lindgren and Rychlik (1996) use their method discussed in Lindgren and Rychlik (1995) to construct confidence bands for level curves of sparsely observed images. Wameling (2003a) uses

the method presented in Wameling and Saborowski (2001) to find confidence bands for the level curves of yearly precipitation in the lower part of Saxony, Germany. Cullman and Saborowski (2005) use indicator kriging and the method presented in Polfeldt (1999) to approximate the conditional probability given the observed data that a random process fails to exceed a response level of interest at specified locations.

#### **1.4 Outline of Dissertation**

Chapters 2 and 3 will present two methods for constructing confidence regions for level curves. The method presented in Chapter 2 is an extension of Lindgren and Rychlik (1995) and Wameling and Saborowski (2001). Instead of constructing confidence bands along the estimated level curve, an approach using confidence boxes (rectangular confidence regions) is proposed. Chapter 2 begins with a brief overview of spatial interpolation, continues with a description of the proposed method, and then concludes with empirical results of a small simulation study. Chapter 3 presents an approach to construct a confidence region for the true level curve which addresses the multiple comparison problem. This method models the response surface as a Gaussian random field and constructs the confidence region using hypothesis testing. The critical value is adjusted to control the simultaneous Type I error of the hypothesis tests. The chapter begins with an explanation of how hypothesis testing can be used to construct this confidence region, describes a simulation method to estimate the critical value needed to control the simultaneous error rate, and then confirms the validity of this approach using a simulation study. After this, discussion will be given to additional error criteria which might be considered in combination with this approach.

In Chapter 4 these two methods will be applied to three different data sets; the first two are simulated data sets designed to highlight the differences in the methods, while the last data set is related to heavy metal contamination in the Swiss Jura.

Chapter 5 introduces a limit theorem for the distribution of the maximum of a triangular sequence of stationary Gaussian random fields. The result is an extension of the work presented by Hsing et al. (1996) to two dimensions. The result was motivated by the desire to control the simultaneous Type I error rate of hypothesis tests at locations on an  $n \times n$  lattice where the test statistics are Gaussian and correlated. Consideration was given to using this result to approximate the critical value of the hypothesis tests discussed in Chapter 3, but the approximation was typically conservative and application of this result needs further exploration. Under certain dependence and limiting conditions we show that the maximum of the random fields exhibits extremal clustering in the limit. Consideration is then given to the use of this result in approximating  $\mathbb{P}(\max_{1 \leq i, j \leq n} Z_{i, j} \leq u)$ , where  $\{Z_{i, j}\}$  is a stationary Gaussian random field on a square  $n \times n$  lattice of equally spaced locations.

Concluding remarks and discussion of future work will be given in Chapter 6.

## Chapter 2

### CONFIDENCE BOXES

#### 2.1 Introduction

In this chapter we extend the work of Lindgren and Rychlik (1995) and Wameling and Saborowski (2001) to construct rectangular confidence regions along the estimated level curve which individually intersect the true level curve with high probability. We call these rectangular regions “confidence boxes”. The confidence boxes extend perpendicular to the estimated level curve and each box is constructed so that the outside edges of each box will touch the edges of adjacent boxes and there are no gaps between boxes. In general, the only conditions necessary to use the method proposed in this chapter are that the random field is continuous almost surely and can be simulated given the observed data. However, for practical purposes we will also assume that the random process is Gaussian and stationary. Additional processes may be considered, but simulation becomes more difficult outside of the Gaussian case.

There are some advantages of our approach over its predecessors. The first is that the assumptions are weaker than those assumed by Lindgren and Rychlik (1995) (since we do not assume differentiability of the random field) and are consistent with the assumptions of Wameling and Saborowski (2001). The method presented in this chapter is valid for any continuous process which can be simulated given the observed data. The second advantage is that the approach presented in this chapter helps to fill in the “gaps” between confidence bands. As seen in Figure 2.1,

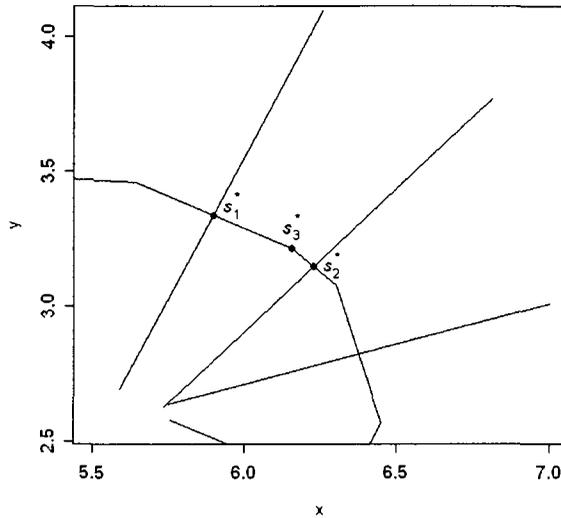


Figure 2.1: Gap between confidence bands. While confidence statements can be made about locations perpendicular to the estimated level curve at  $s_1^*$  and  $s_2^*$ , no such statement can be made about  $s_3^*$ .

confidence intervals built along  $\hat{I}_u$  leave gaps along  $\hat{I}_u$  where no confidence statement is made. Suppose that confidence intervals have been constructed perpendicular to  $\hat{I}_u$  at two points  $s_1^*$  and  $s_2^* \in \hat{I}_u$ . Consider making a confidence statement about the location of the true level curve in directions perpendicular to  $\hat{I}_u$  for a point  $s_3^*$  between  $s_1^*$  and  $s_2^*$ . It seems reasonable that the confidence interval at this point would be similar to the ones at  $s_1^*$  and  $s_2^*$ , but no formal statement can be made using the previous methods without constructing an additional confidence interval. The proposed method overcomes this problem by constructing the boxes so that the outside edges of each box touch the edges of the adjacent boxes. Every point on the estimated level curve will be associated with a confidence box, so every point on the estimated level curve will have an associated confidence statement. In this case, the confidence statement is not about the particular point, but about the confidence box associated with the point. Specifically, individually for each confidence box the probability that the true level curve passes through the box should be  $1 - \alpha$ .

We continue this chapter by going over some basic ideas and concepts related to the interpolation of spatial processes. We will then describe the procedure for constructing the confidence boxes. This chapter will conclude with a small simulation study which confirms that the empirical confidence levels of our confidence regions are similar to the nominal levels.

## 2.2 Interpolation of Spatial Processes

### 2.2.1 Characterizing Spatial Variation

Consider a random field  $\{Z(s)\}$ , where  $s$  is the location of  $Z$  in  $\mathcal{D} \subset \mathbb{R}^2$ . Typically,  $Z(s)$  is decomposed into the sum

$$Z(s) = m(s) + Y(s),$$

where  $m(s)$  is a smoothly varying trend function that captures the large scale variability of  $Z$  and  $Y(s)$  is a mean zero random function capturing the small scale behavior of  $Z$ .

The process  $Z$  is said to be *second-order stationary* if it satisfies the following properties:

1. The mean of  $Z$  is constant so that  $\mathbb{E}(Z(s)) = \mu$  for all  $s$ .
2. The variance of  $Z$  is finite, and for any  $s \in \mathbb{R}^2$  and displacement vector  $\mathbf{h} \in \mathbb{R}^2$ , we have  $\text{Cov}(Z(s), Z(s + \mathbf{h})) = C(\mathbf{h})$ .

If the covariance does not depend on direction, the spatial variation is said to be *isotropic*. In this case, the displacement vector  $\mathbf{h}$  may be replaced by the distance  $|\mathbf{h}|$ . Hereafter, we use the term “stationary” to refer to a second-order stationary process.

The covariance function of a process gives us insight into its continuity and differentiability. Since we will be concerning ourselves with Gaussian processes, we

present results related to the continuity and differentiability of Gaussian processes. A stationary Gaussian process is continuous almost surely if its covariance function

$$C(\mathbf{h}) = C(\mathbf{0}) - K|\mathbf{h}|^\beta + o(|\mathbf{h}|^\beta), \quad \text{as } \mathbf{h} \rightarrow \mathbf{0},$$

where  $K > 0$  and  $0 < \beta \leq 2$ . Similarly, a stationary Gaussian process is continuously differentiable if

$$-C''(\mathbf{h}) = -C''(\mathbf{0}) - K|\mathbf{h}|^\alpha + o(|\mathbf{h}|^\alpha),$$

where  $K > 0$  and  $0 < \alpha \leq 2$ . The conditions for continuity almost surely are satisfied for almost all covariance functions encountered in applied probability (cf. Lindgren (2004, p. 25)). However, it is common for continuous processes to not be continuously differentiable.

A necessary (but not sufficient) condition for a process  $Z$  to be continuous almost surely is that  $C(\mathbf{h}) \rightarrow C(\mathbf{0})$  as  $\mathbf{h} \rightarrow \mathbf{0}$ . In practice, our observed responses may be measured with measurement error. In this case, our observed responses are realizations of

$$Z_{obs}(s) = Z(s) + \epsilon,$$

where  $\epsilon$  is a Gaussian random variable independent of location with mean zero and variance  $\tau^2$ . In these cases, the empirical covariance function will exhibit a nugget effect in which  $C(\mathbf{h})$  is discontinuous at zero. We emphasize the fact that the underlying random field is always assumed to be continuous almost surely. Even when measurement error is present in the observed responses, interest lies in the behavior of the underlying continuous random field  $Z$ .

Some common covariance functions in geostatistics include the Gaussian, spherical, exponential, and Matérn covariance functions. These covariance functions are parameterized in terms of several parameters: the partial sill  $\sigma^2$ , the range parameter  $\phi$ , the nugget  $\tau^2$ , and in the case of the Matérn covariance function, an additional smoothness parameter  $\nu$ . The partial sill  $\sigma^2$  represents the variance of  $Z(s)$ , i.e.,

$\sigma^2 = \text{Var}(Z(s))$ . The nugget is the variance of the random measurement error. The variance of the observed responses is  $\sigma^2 + \tau^2$  when measurement error is present. The parameter  $\phi$  models the dependence of the random field. When  $\phi$  is larger, the dependence between neighboring locations is greater (in our parameterization). The smoothness parameter  $\nu$  models the smoothness of the underlying random field. The process becomes smoother as  $\nu$  increases.

We now specify the forms of the Gaussian, spherical, exponential, and Matérn covariance functions. For simplicity, we assume isotropic covariance functions. We include the variance  $\tau^2$  of the measurement error in the parameterizations with the understanding that this is zero when no measurement error is present. The Gaussian covariance function is defined by

$$C(h) = \begin{cases} \sigma^2 + \tau^2 & \text{if } h = 0 \\ \sigma^2 \exp\left(-\frac{h^2}{\phi^2}\right) & \text{if } h > 0. \end{cases}$$

Processes having Gaussian covariance functions will have infinitely differentiable sample paths. The spherical covariance function is given by

$$C(h) = \begin{cases} \sigma^2 + \tau^2 & \text{if } h = 0 \\ \sigma^2 \left(1 - \frac{3h}{2\phi} + \frac{1}{2} \left(\frac{h}{\phi}\right)^3\right) & \text{if } 0 < h < \phi \\ 0 & \text{if } h \geq \phi. \end{cases}$$

The exponential covariance function is given by

$$C(h) = \begin{cases} \sigma^2 + \tau^2 & \text{if } h = 0 \\ \sigma^2 \exp\left(-\frac{h}{\phi}\right) & \text{if } h > 0. \end{cases}$$

Processes having spherical or exponential covariance functions are not smooth enough to have differentiable sample paths. The last covariance function we mention is the Matérn covariance function. The Matérn covariance function can be parameterized in several ways, but we follow the parameterization given in Stein (1999). The Matérn covariance function is parameterized by

$$C(h) = \begin{cases} \sigma^2 + \tau^2 & \text{if } h = 0 \\ \sigma^2 \left(\frac{h}{\phi}\right)^\nu \mathcal{K}_\nu\left(-\frac{h}{\phi}\right) & \text{if } h > 0, \end{cases}$$

where  $\mathcal{K}_\nu$  is modified Bessel function of the second kind. Random fields having the Matérn covariance function are  $\lceil \frac{\nu-2}{2} \rceil$  (the ceiling function of  $(\nu - 2)/2$ ) times differentiable (cf. Paciorek (2003, p. 44)).

### 2.2.2 Kriging

Kriging predictors are commonly used to interpolate spatial data. Kriging predictors make use of the way a process varies through space, usually modeled through either a covariance or semivariance model, and are optimal in the sense that they are unbiased and minimize the mean squared error of prediction.

As with most forms of classical kriging, the covariance function of  $Z(s)$  is assumed to be known. In practice, the parameters are estimated empirically and then used as if they were the true values. This estimation process is described in a number of excellent resources such as Chiles and Delfiner (1999), Schabenberger and Gotway (2005), Cressie (1993), Goovaerts (1997), Banerjee et al. (2004), Stein (1999), and Webster and Oliver (2007), and will not be discussed here.

Assume that the random field has the form

$$Z(s) = \mathbf{x}(s)' \beta + Y(s),$$

where  $\mathbf{x}(s)$  is a vector of explanatory variables associated with  $s$ ,  $\beta$  is a vector of fixed constants, and  $Y(s)$  is a mean zero random function with covariance function  $C(\cdot)$ .

The kriging prediction at location  $s$  is given by

$$\hat{Z}(s) = \lambda' \mathbf{Z},$$

where  $\mathbf{Z}' = [Z(s_1), \dots, Z(s_N)]$  and  $\lambda'$  is a vector of weights chosen to minimize  $\mathbb{E}(Z(s) - \hat{Z}(s))^2$  and ensure that  $\mathbb{E}(Z(s)) = \mathbb{E}(\hat{Z}(s))$ . Special cases of kriging include:

1. Simple kriging when  $\beta = 0$  for all  $s$ .

2. Ordinary kriging when  $\mathbf{x}(s)$  is a vector of ones for all  $s$  and  $\beta$  is unknown.
3. Universal kriging for more general  $\mathbf{x}(s)$  and  $\beta$ .

In general, one can assume that  $\beta = 0$  without added difficulty. Ordinary and universal kriging are computationally identical to simple kriging after the generalized least squares estimate of the mean is subtracted from the observed responses.

### 2.3 Details of Constructing Confidence Boxes

The goal of the method proposed in this chapter is to construct a set of rectangular regions such that the probability that the true level curve intersects each region individually is quite high. Denoting the  $i$ th confidence box by  $S_i$ , our desire is to construct  $S_i$  such that

$$\mathbb{P}(Z(s) = u \text{ for some } s \in S_i) \geq 1 - \alpha.$$

We will do this using the concept of *level crossings* and modifying the ideas of Lindgren and Rychlik (1995) and Wameling and Saborowski (2001).

The approach presented in Wameling and Saborowski (2001) is more flexible and straightforward than that of Lindgren and Rychlik (1995) and will be the starting point of our approach. As summarized in Section 1.3.2, the basic approach of Wameling and Saborowski (2001) is:

1. Use kriging to construct estimated level curves  $\hat{I}_u$ .
2. Along transects perpendicular to  $\hat{I}_u$ , generate a large number of realizations of the process  $Z(s)$  given the observed data  $z(s_1), \dots, z(s_N)$  (an example of this was shown in Figure 1.5).
3. Find the level crossing nearest to  $\hat{I}_u$  for each realization.
4. Use this information to construct confidence bands of the appropriate size.

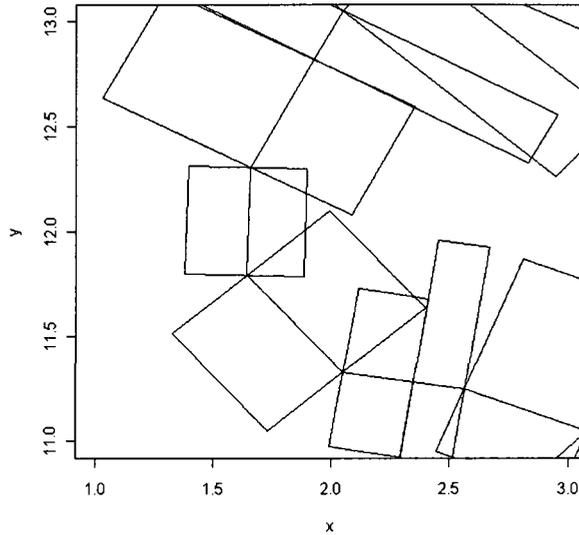


Figure 2.2: Closeup of confidence boxes extending from the linear approximation of the estimated level curve.

Our approach is similar, except that in step two, we generate realizations of  $Z(s)$  given the observed data in rectangular regions stretching perpendicular to the estimated level curve. The procedure we follow will be illustrated through several figures and is described in detail below.

First, kriging is used to construct the estimated level curves  $\hat{I}_u$  in our domain of interest  $\mathcal{D}$ . Next, we wish to simulate realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$  in rectangular regions perpendicular to  $\hat{I}_u$ . The width of each box is chosen so that the edge of each box touches the neighboring box and there are no gaps between the boxes along the estimated level curve. In practice,  $\hat{I}_u$  will be a linear approximation of the set  $\{s : \hat{Z}(s) = u\}$ . Thus, the approximation will be a set of line segments and we denote the set of line segments comprising this linear approximation by  $\{L_i\}$ , where  $L_i$  represents the  $i$ th transect. A convenient way to choose the widths of the rectangular regions is to associate a rectangular region with each line segment of the linear approximation. The width of the  $i$ th rectangular region then corresponds to the length of the associated line segment  $L_i$ . This relationship is shown in Figure 2.2. Care must be taken in choosing the proper length of the transects making up the linear approximation of  $\hat{I}_u$ . Choosing the length of the line transects to be a fixed distance

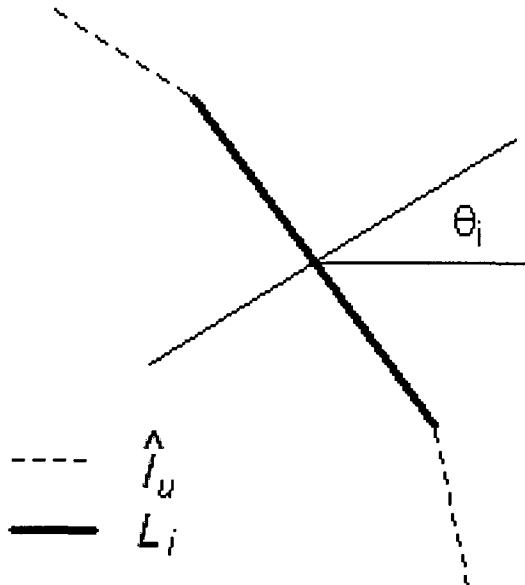


Figure 2.3: Angle associated with line segment  $L_i$ . Angle  $\theta_i$  is the angle perpendicular to line segment  $L_i$ .

can create problems. If the lengths of the transects are too small, then the confidence boxes will be narrow and look similar to confidence bands. If the lengths are too long, then the linear approximation may not be an adequate approximation of  $\hat{I}_u$ . Additionally, since the true level curve may intersect a confidence box at any location in the box, a wide confidence box may yield little insight into the location of the true level curve. For this reason, we prefer a linear approximation which allows the lengths of the transects to vary. For the analysis done in this dissertation, the `contourLines()` function in the R software package provided a suitable linear approximation of the estimated level curve.

For each line segment  $L_i$ , we will generate  $M$  realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$  in a rectangular region  $B_i(\tau_1, \tau_2)$  perpendicular to  $L_i$  (this will be done independently of the realizations for other regions). Let  $\theta_i$  represent the angle in relation to the  $x$ -axis which is perpendicular to segment  $L_i$ , as shown in Figure 2.3. Define  $t_\theta(s, \tau)$  to be the location a distance of  $\tau$  away from  $s$  at angle  $\theta$ . More formally,

$$t_\theta(s, \tau) = (s_x + \tau \cos(\theta), s_y + \tau \sin(\theta)),$$

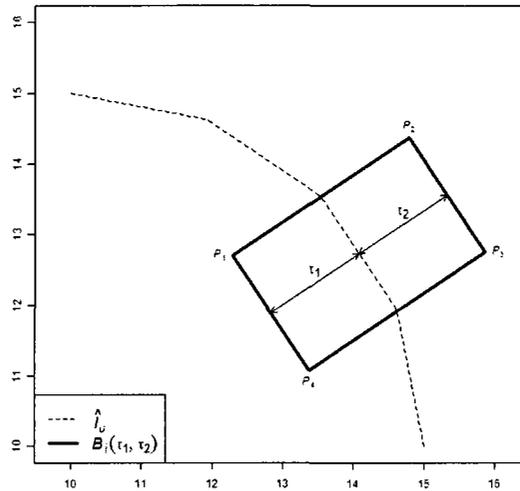


Figure 2.4: A graphical display of  $B_i(\tau_1, \tau_2)$ .  $B_i(\tau_1, \tau_2)$  is a rectangular region stretching a distance of  $\tau_1$  and  $\tau_2$  from line segment  $L_i$ .

where  $s_x$  and  $s_y$  represent the  $x$ - and  $y$ -coordinates of  $s$ . Let  $s_{i,1}$  and  $s_{i,2}$  represent the two endpoints of line transect  $L_i$  and let

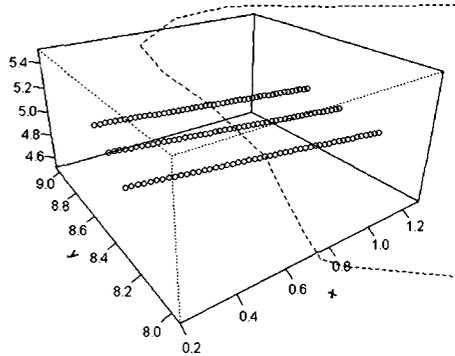
$$P_{i,1} = t_{\theta_i}(s_{i,1}, \tau_1),$$

$$P_{i,2} = t_{\theta_i}(s_{i,1}, \tau_2),$$

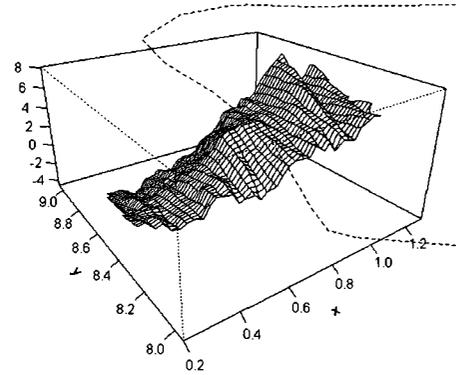
$$P_{i,3} = t_{\theta_i}(s_{i,2}, \tau_2),$$

$$P_{i,4} = t_{\theta_i}(s_{i,2}, \tau_1),$$

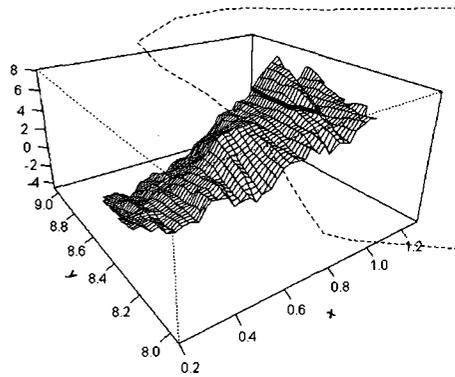
denote points at distances  $\tau_1$  and  $\tau_2$  and angle  $\theta_i$  away from the endpoints of  $L_i$ .  $B_i(\tau_1, \tau_2)$  is the rectangular region extending distances of  $\tau_1$  and  $\tau_2$  from line segment  $L_i$ . Thus,  $B_i(\tau_1, \tau_2)$  is the rectangular region bounded by line segments  $\overline{P_{i,1}P_{i,2}}$ ,  $\overline{P_{i,2}P_{i,3}}$ ,  $\overline{P_{i,3}P_{i,4}}$ , and  $\overline{P_{i,4}P_{i,1}}$ .  $B_i(\tau_1, \tau_2)$  is illustrated in Figure 2.4. Note that typically,  $\tau_1$  is a negative value while  $\tau_2$  is a positive value. Because we cannot generate observations at every location in  $B_i(\tau_1, \tau_2)$ , we choose to generate realizations at locations along three transects in each rectangular region. Two of the transects run along the edges of each rectangle, while the other transect runs along the center of the rectangle, as shown in Figure 2.5 (a).



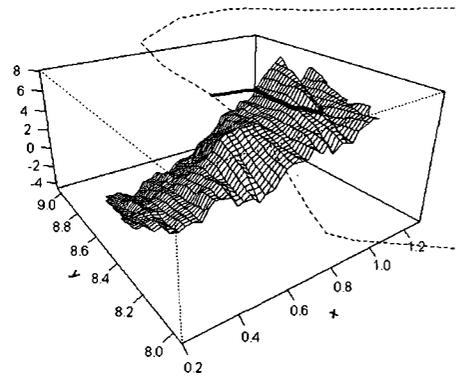
(a) Realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$  are simulated at locations falling on three transects running perpendicular to the estimated level curve. The estimated level curve is represented by the dashed line.



(b) The realized responses are interpolated in the rectangular region to produce a surface. The dashed line represents the estimated level curve.



(c) After interpolating the realized surface over the grid of locations, we construct the level curve of the realized process. The realized level curve is shown as a thick black line.



(d) After interpolating the realized surface and constructing the realized level curve, we locate the nearest level crossing on the realized level curve and measure its distance to the estimated level curve.

Figure 2.5: Process of constructing confidence boxes.

After generating a realization of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$  at the points along the three transects, we interpolate the realized responses over the rectangular region to get a response surface looking something like the one shown in Figure 2.5 (b). Having constructed the realized response surface, our next step is to find the location nearest to the estimated level curve where the response surface crosses the response level  $u$ . To do this, we first construct the level curve of the realized process as shown in Figure 2.5 (c). After constructing the realized level curve, we locate the point on the realized level curve which is closest to  $\hat{I}_u$  (which corresponds to the level crossing nearest to the estimated level curve) and measure the distance between this point and the estimated level curve. This is depicted in Figure 2.5 (d). Note that this distance will always be measured at an angle perpendicular to the estimated level curve. Denote the minimum distance for the  $j$ th realization in rectangular region  $i$  as  $\tau_{i,j}$ , and denote the set of nearest level crossing distances for the  $i$ th rectangular region as  $T_{\min}^i = \{\tau_{i,j}\}_{j=1}^M$ .

An empirical confidence box of the appropriate size for rectangular region  $i$  is given by  $B_i(\tau_{\alpha/2}^i, \tau_{1-\alpha/2}^i)$ , where  $\tau_{\alpha/2}^i$  and  $\tau_{1-\alpha/2}^i$  denote the  $\alpha/2$  and  $1-\alpha/2$  quantiles of  $T_{\min}^i$ . Since a  $u$ -level crossing occurred in the region  $B_i(\tau_{\alpha/2}^i, \tau_{1-\alpha/2}^i)$  in  $(1-\alpha) \times 100\%$  of the realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$ , we expect that the true level curve  $I_u$  will intersect the region  $B_i(\tau_{\alpha/2}^i, \tau_{1-\alpha/2}^i)$  with approximate confidence level  $1-\alpha$ .

## 2.4 Simulation Study

The goal of the method presented in this chapter is to construct rectangular confidence regions  $S_i$  along the estimated level curve such that each individual confidence region intersects the true level curve with high probability (no confidence statement is made regarding the simultaneous probability). We will check whether the empirical confidence levels of our confidence regions are similar to the desired confidence levels through a small simulation study.

### 2.4.1 General Description of Simulation Procedure

We first generate a realization of the response variable  $Z$  over a fine grid of locations in the domain of interest  $\mathcal{D}$ . The realized response surface is interpolated over  $\mathcal{D}$  and the level curve for this realization is constructed. The level curve of the realized surface will be used as the true level curve  $I_u$  for comparison purposes.

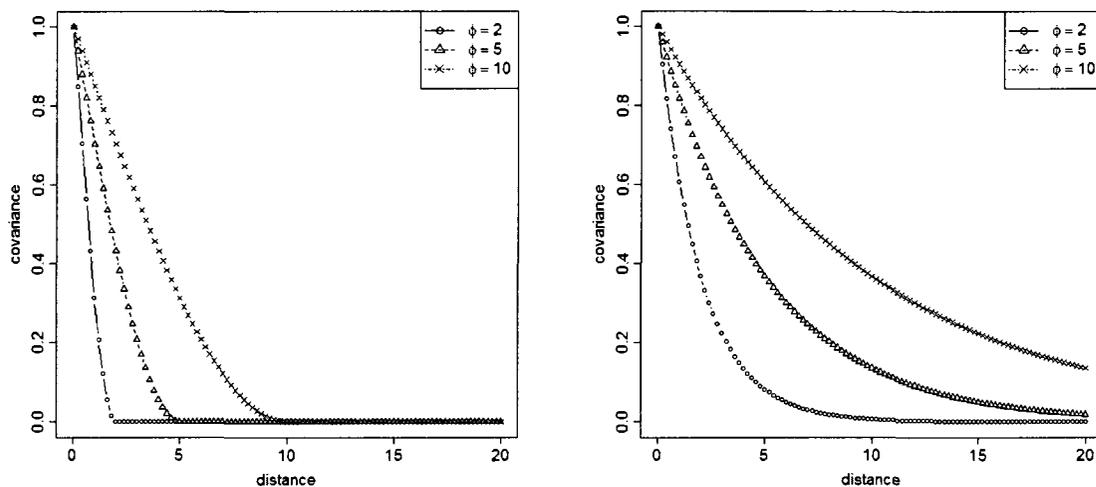
From this realized true surface, we randomly sample  $N$  locations and use the corresponding responses as our observed responses  $z(s_1), \dots, z(s_N)$ . The observed responses are used to make kriging predictions over a grid of densely spaced locations in the domain of interest. These predictions are interpolated over  $\mathcal{D}$ , and from this we produce a linear approximation of the estimated level curve  $\hat{I}_u$  using the `contourLines()` function in R.

For each line segment  $L_i$  which makes up the linear approximation of  $\hat{I}_u$ , we simulate  $M$  realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$  at locations along three transects in the region  $B_i(\tau_1, \tau_2)$ . The nearest level crossing is determined for each realization, and the set of nearest crossing distances is used to determine the appropriate size of the  $i$ th confidence region  $S_i$ . After constructing each confidence region  $S_i$ , we check whether the true level curve intersects the  $i$ th confidence region.

This procedure will be performed for numerous realizations of the “true” level curve. As an overall measure of performance, we calculate the proportion of confidence boxes intersecting the true level curve across all realizations of  $Z$ .

### 2.4.2 Specifics of Simulation Procedure

We assume that  $Z$  has mean zero, variance  $\sigma^2 = 1$ , and nugget  $\tau^2 = 0$ . Simulations were performed using both the spherical and exponential covariance functions. For each covariance function we ran simulations using range parameters  $\phi = 2, 5$ , and 10. The procedure can be performed using smaller values of  $\phi$ , but the weaker dependence typically results in longer simulation times. The covariance functions



(a) Spherical covariance functions used in the simulation procedure.

(b) Exponential covariance functions used in the simulation procedure.

Figure 2.6: Covariance functions of simulation procedure.

used in the simulations, relative to the size of the domain, are shown in Figures 2.6 (a) and (b).

Following the example of Wameling (2003b), we assumed the region of interest  $\mathcal{D}$  to be a square region of size  $[0, 20] \times [0, 20]$ . The initial unconditional realization of  $Z$  was generated at locations on a grid of size  $200 \times 200$ , and the response surface was interpolated over these locations. The true level curve was constructed for the level  $u = 1$ .

From the initial set of realized locations, we randomly selected  $N = 200$  locations to use as our observed responses  $z(s_1), \dots, z(s_{200})$ . For each realization of  $Z$ , we constructed the estimated level curve  $\hat{I}_1$  using both simple and ordinary kriging.

To determine the appropriate size of the  $i$ th confidence box  $S_i$ , 100 realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_{200})$  were produced in  $B_i(\tau_1, \tau_2)$ .  $\tau_{\alpha/2}$  and  $\tau_{1-\alpha/2}$  were then chosen according to confidence level 0.90.

A confidence box is “successful” if the true level curve passes through the confidence box. For a confidence level of 0.90, approximately 90 percent of the confidence

Table 2.1: Empirical confidence level of confidence boxes.

covariance function	$\phi$	sk	ok
spherical	2	.91	.91
spherical	5	.89	.90
spherical	10	.91	.90
exponential	2	.90	.90
exponential	5	.90	.90
exponential	10	.89	.88

boxes should intersect the true level curve. For each combination of covariance function and range parameter, we generated 20 realizations of the true level curve  $I_1$ . In each realization, there were typically 100 or more confidence boxes. The empirical confidence level of our method was determined by calculating the proportion of confidence boxes intersecting the true level curve across all realizations.

### 2.4.3 Results

This empirical confidence levels for each covariance function are provided in Table 2.1. The type of kriging used when constructing the confidence boxes is specified by “sk” for simple kriging and “ok” for ordinary kriging. The empirical confidence levels were close to the intended level of 0.90 in all cases. In addition to the empirical confidence level calculated across all realizations, we also looked at the success rate of the confidence boxes in each realization. More specifically, we looked at the proportion of boxes that “successfully” intersected the true level curve in each realization. We noted that though the confidence boxes are constructed independently of one another, there is still correlation between the boxes of the same realization. When the estimated level curve for a particular realization is a poor approximation of the true level curve, it will likely lead to a lower success rate for the confidence boxes of that particular realization. Conversely, when the estimated level curve of a realization approximates the true level curve well, it is likely that the success rate of the confidence boxes for that particular realization will be higher than the nominal

confidence level. Typically, the success rate of the confidence boxes for individual realizations of the true level curve was within 0.05 of the confidence level 0.90.

## Chapter 3

### HYPOTHESIS TESTING

#### 3.1 Introduction

In this chapter we present an approach for the construction of a confidence region for the level curve  $I_u$  of a Gaussian random field. The confidence region will be constructed through hypothesis testing, adjusting the critical value of the test to control the familywise error rate (FWER) through empirical simulation of the test statistic.

We would like to construct a confidence region  $S$  for the level curve  $I_u$  such that  $\mathbb{P}(I_u \subseteq S) \geq 1 - \alpha$ . Instead of finding  $S$  directly, we will try to find a set  $R$  which does not intersect  $I_u$  with high probability, i.e.,  $R$  such that  $\mathbb{P}(\{I_u \cap R\} = \emptyset) \geq 1 - \alpha$ . Consequently, the set  $S = R^c$  will satisfy our goal since  $\mathbb{P}(I_u \subseteq R^c) \geq 1 - \alpha$ .

Suppose we observe responses  $z(s_1), \dots, z(s_N)$  of a single realization of a Gaussian random field  $Z$ , where  $s_1, \dots, s_N$  are locations in the two-dimensional region of interest  $\mathcal{D} \subset \mathbb{R}^2$ . We assume the random field with known mean  $\mu$  (without loss of generality we take  $\mu = 0$ ) is stationary and has continuous sample paths almost surely. The region  $R$  is formally constructed by testing for each  $s$ ,  $H_0 : Z(s) = u$  versus  $H_a : Z(s) \neq u$ . The region  $R$  is then the union of all  $s$  for which we conclude that  $Z(s) \neq u$ .

To construct a test statistic, we define

$$\hat{Z}(s) = \mathbb{E}(Z(s) \mid Z(s_1), \dots, Z(s_N)),$$

and

$$\hat{\sigma}^2(s) = \mathbb{E}(Z(s) - \hat{Z}(s))^2.$$

Thus,

$$Z'(s) = \frac{Z(s) - \hat{Z}(s)}{\hat{\sigma}(s)} \sim N(0, 1),$$

and is a convenient choice as test statistic. Since  $Z$  is Gaussian with mean zero,  $\hat{Z}$  corresponds to the simple kriging predictor.

Our confidence region will fail to contain the true level curve when any Type I error is made in our hypothesis tests. In order to control the confidence level of our confidence region, we will need to control the familywise error rate (FWER) of our hypothesis tests. A Type I error can only be made when  $Z(s) = u$ , and thus, only at locations  $s \in I_u$ . To control the FWER at significance level  $\alpha$ , we should only reject  $H_0$  and conclude that  $Z(s) \neq u$  when  $|Z'(s)| \geq C_\alpha$ , where  $C_\alpha$  is a critical value chosen so that

$$\mathbb{P}\left(\sup_{s \in I_u} |Z'(s)| \geq C_\alpha\right) = \alpha. \quad (3.1)$$

There are two main difficulties in calculating  $C_\alpha$ . First, the level curve  $I_u$  is random and unknown. Second,  $\{Z'(s)\}$  is a nonstationary Gaussian random field; problems involving the extremes of Gaussian random fields are typically difficult to solve and are more difficult when the random field is nonstationary.

A natural approach to dealing with the first difficulty is to find  $C_{\alpha/2}^*$  such that

$$\mathbb{P}\left(\sup_{s \in \mathcal{D}} Z'(s) \geq C_{\alpha/2}^*\right) = \alpha/2. \quad (3.2)$$

$C_{\alpha/2}^*$  will be a conservative approximation of  $C_\alpha$ , but is easier to approximate since we no longer have to consider the behavior of the level curve  $I_u$ . There are two reasons that  $C_{\alpha/2}^*$  will be a conservative approximation. The first reason is that we are considering the distribution of  $Z'(s)$  for all  $s \in \mathcal{D}$  instead of  $s \in I_u$ . The second reason is because a realization of  $\{Z'(s), s \in \mathcal{D}\}$  may have extreme positive and

negative values occur simultaneously. Suppose that  $\mathbb{P}(\sup_{s \in \mathcal{D}} Z'(s) \geq C_{\alpha/2}^*) = \alpha/2$ .

Then

$$\begin{aligned}
& \mathbb{P}\left(\sup_{s \in \mathcal{D}} |Z'(s)| \geq C_{\alpha/2}^*\right) \\
&= \mathbb{P}\left(\left\{\inf_{s \in \mathcal{D}} Z'(s) \leq -C_{\alpha/2}^*\right\} \cup \left\{\sup_{s \in \mathcal{D}} Z'(s) \geq C_{\alpha/2}^*\right\}\right) \\
&= \mathbb{P}\left(\inf_{s \in \mathcal{D}} Z'(s) \leq -C_{\alpha/2}^*\right) + \mathbb{P}\left(\sup_{s \in \mathcal{D}} Z'(s) \geq C_{\alpha/2}^*\right) \\
&\quad - \mathbb{P}\left(\left\{\inf_{s \in \mathcal{D}} Z'(s) \leq -C_{\alpha/2}^*\right\} \cap \left\{\sup_{s \in \mathcal{D}} Z'(s) \geq C_{\alpha/2}^*\right\}\right) \\
&\leq \alpha.
\end{aligned}$$

Approximating (3.2) may still be difficult in principle since  $\{Z'(s)\}$  is nonstationary. To overcome this difficulty, instead of considering the behavior of the nonstationary random field  $\{Z'(s)\}$ , one might consider the behavior of a stationary Gaussian random field as an approximation.

The behavior of the extremes of Gaussian random variables has long been of interest. Theory related to the excursion probability

$$\mathbb{P}\left(\sup_{t \in T} Z(t) \geq u\right), \quad (3.3)$$

has a rich literature history, where  $Z$  is a real-valued Gaussian random field over some domain of interest  $T$  (which may be continuous or discrete and multidimensional). For a helpful overview of results related to the extremal behavior of Gaussian random fields, see Adler (2000) and Chapters 2 and 4 of Adler and Taylor (2007). There have been many approximations of (3.3) given in the literature. The theoretical papers have generally treated only stationary processes over smooth Euclidean domains and have never been able to identify the precise level of accuracy for the approximations (cf. Adler and Taylor (2007, p. 350)). Typically, computations for nonstationary and non-differentiable processes become “prohibitively complicated, and, if doable, then only in very special cases,” (cf. Adler (2000, p. 41)). Since  $\{Z'(s)\}$  is nonstationary

and not necessarily differentiable, there does not appear to be any method that easily allows us to estimate  $C_{\alpha/2}^*$ . Recently, Adler (2008) presented an approximation of (3.3) which does not require stationarity or isotropy a priori. An explicit, relatively simple result was provided for stationary, isotropic, twice-differentiable Gaussian random fields when  $T \subset \mathbb{R}^d$  for  $d \geq 1$  and certain regularity assumptions are met. Though the process  $Z'(s)$  does not meet the assumptions needed to use this explicit result, we did explore using the result to estimate  $C_{\alpha/2}^*$  (and thus gain an approximation for  $C_\alpha$ ). As discussed in Section 3.4, this estimate of  $C_{\alpha/2}^*$  is often not an accurate approximation of  $C_\alpha$  and is typically quite conservative. It is possible that the approximation could be improved if the nonstationarity of  $Z'$  were taken into account. However, an explicit result for the nonstationary case was not provided by Adler and it is not clear whether a calculable expression is obtainable. Another approach we considered was to approximate  $C_\alpha$  using the main result in Chapter 5 regarding the behavior of the maxima of a triangular sequence of stationary Gaussian random fields. The conditions needed for this result are relatively mild, and Gaussian processes having common covariance functions such as the Gaussian, spherical, exponential, and Matérn all satisfy the conditions. However, this approximation was also typically conservative and further exploration needs to be done regarding the use of this result in approximating  $C_\alpha$ .

Given the difficulties in estimating  $C_\alpha$ , an expedient way to find the proper critical value is through simulation. Because of the difficulties outlined above, care must be taken in the simulation process and the details of the simulation are described below. We will continue this chapter by first describing a simulation procedure which can be used to estimate  $C_\alpha$ . This is followed by results of a simulation study which validates the method proposed in this chapter by checking the empirical confidence level of confidence regions constructed using this method. Next, we briefly compare the estimates of  $C_\alpha$  obtained using the excursion probability approximation in Adler

(2008) (for the stationary, isotropic, twice-differentiable case) with the estimates obtained through simulation. Lastly, since it may be unnecessarily strict to control the simultaneous Type I error rate of the hypothesis tests, we finish by discussing alternative error criteria which may be used.

### 3.2 Proposed Method of Simulation for Estimating $C_\alpha$

It is clear from (3.3) that in order to properly estimate  $C_\alpha$  we must consider the distribution of  $\{Z'(s), s \in I_u\}$ . Thus, there are two main elements to consider in correctly estimating  $C_\alpha$  through simulation:

1. Proper simulation of  $Z'$ .
2. Properly modeling the randomness of the true level curve  $I_u$ .

As previously mentioned, we assume that  $Z(s)$  is a stationary Gaussian random field which is continuous almost surely. Without loss of generality, we assume that the mean of  $Z$  is zero. Under these assumptions, the predictor  $\hat{Z}(s) = \mathbb{E}(Z(s) | Z(s_1), \dots, Z(s_N))$  corresponds to the simple kriging predictor

$$\hat{Z}(s) = \mathbf{Z}'\Sigma^{-1}\mathbf{c}_s,$$

where  $\mathbf{Z}' = [Z(s_1), \dots, Z(s_N)]$ ,  $\Sigma$  is the  $N \times N$  covariance matrix of  $\mathbf{Z}$ , and  $\mathbf{c}'_s = [\text{cov}(Z(s), Z(s_1)), \dots, \text{cov}(Z(s), Z(s_N))]$ . Correspondingly, the kriging variance  $\hat{\sigma}^2(s)$  is given by

$$\hat{\sigma}^2(s) = \mathbb{E}(Z(s))^2 - \mathbf{c}'_s\Sigma^{-1}\mathbf{c}_s.$$

We wish to model the behavior of  $Z'(s)$  given the observed responses  $z(s_1), \dots, z(s_N)$ . Because  $\hat{Z}(s)$  and  $\hat{\sigma}(s)$  are deterministic functions, the random behavior of  $Z'(s)$  comes from the random response value  $Z(s)$ . The observed responses give us information about likely response values of  $Z(s)$ . Thus, instead of considering unconditional realizations of  $Z(s)$  to model the variability in the potential responses of  $Z(s)$ ,

we should only consider the ensemble of conditional realizations given the observed data. Denoting the conditional random variable  $Z(s)|z(s_1), \dots, z(s_N)$  by  $\tilde{Z}(s)$ , we can explore the behavior of  $Z'(s)$  by generating realizations of

$$\frac{\tilde{Z}(s) - \hat{Z}(s)}{\hat{\sigma}(s)}.$$

Next, we deal with the challenge that  $I_u$  is unknown and random. As previously stated, observed responses  $z(s_1), \dots, z(s_N)$  give us insight into likely responses of  $Z(s)$ . Similarly, realizations of  $\tilde{Z}(s) = Z(s) | z(s_1), \dots, z(s_N)$  give us insight into plausible locations of  $I_u$ . Define  $\tilde{I}_u = \{s : \tilde{Z}(s) = u\}$ . Figure 3.1 compares the true level curve of a realization of  $Z$  with the level curves  $\tilde{I}_u$  resulting from five realizations of the conditional random field  $\tilde{Z}$ .  $\tilde{I}_u$  tends to be located fairly close to  $I_u$  and it seems reasonable to use realizations of  $\tilde{I}_u$  to approximate the behavior of  $I_u$ .

The last thing we need to consider in our estimation of  $C_\alpha$  is how we use the distribution of  $|Z'(s)|$  to find the appropriate critical value. Recall that we desire  $C_\alpha$  such that  $\mathbb{P}(\sup_{s \in I_u} |Z'(s)| \geq C_\alpha) = \alpha$ . Hence,  $C_\alpha$  should be greater than the  $\sup\{|Z'(s)|, s \in I_u\}$  with probability  $1 - \alpha$ . Replacing  $I_u$  by  $\tilde{I}_u$ , our critical value should be estimated by considering the maximum of  $\{|Z'(s)|, s \in \tilde{I}_u\}$  for each realization, and then taking the  $(1 - \alpha)$  quantile of these values.

Combining all of the previous information, we propose the following method for estimating the critical value  $C_\alpha$ :

- Generate realizations of the conditional field  $\tilde{Z}(s)$  throughout the domain  $\mathcal{D}$ .
- For each realization, construct  $\tilde{I}_u = \{s : \tilde{Z}(s) = u\}$ .
- In each realization, calculate the standardized prediction error

$$\{|Z'(s)|, s \in \tilde{I}_u\} = \left\{ \left| \frac{u - \hat{Z}(s)}{\hat{\sigma}(s)} \right|, s \in \tilde{I}_u \right\}.$$

- Find the maximum of  $\{|Z'(s)|, s \in \tilde{I}_u\}$  for each realization.
- Calculate the  $(1 - \alpha)$  quantile of the set of maxima from the previous step and use it as the estimate of  $C_\alpha$ .

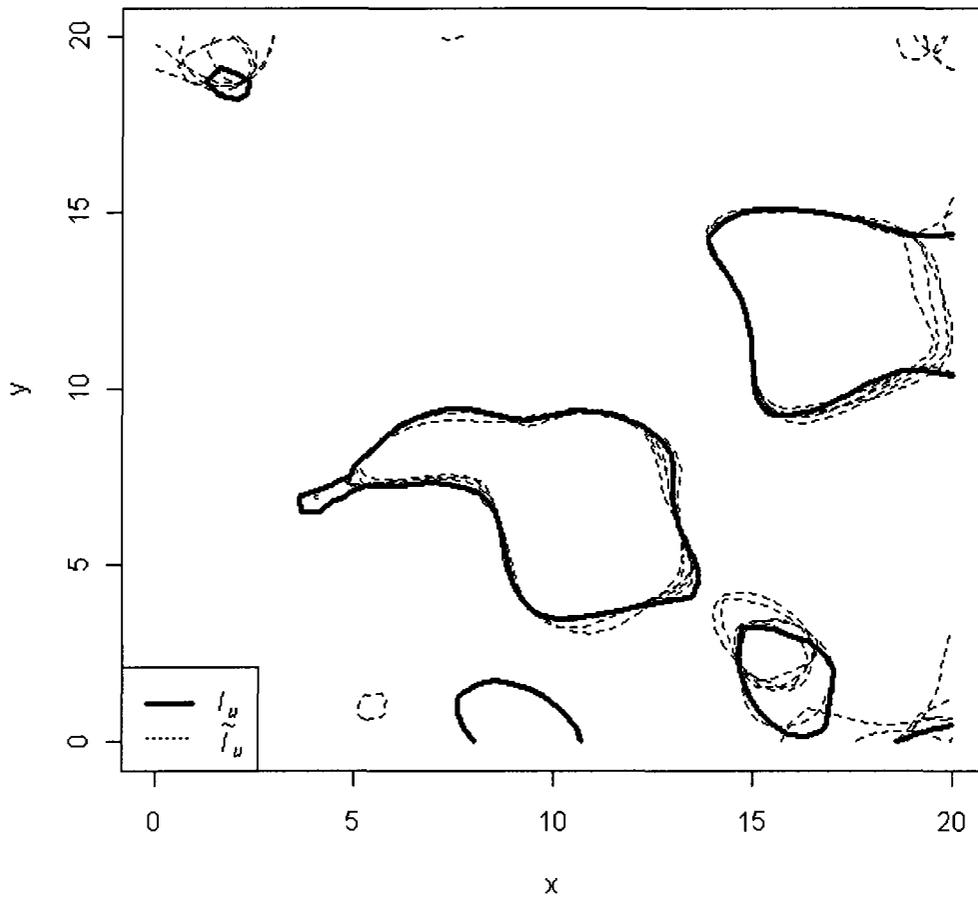


Figure 3.1: A comparison of the true level curve and realized level curves. The true level curve  $I_u$  is shown along with the level curves  $\tilde{I}_u$  from five realizations of  $Z(s)$  given  $z(s_1), \dots, z(s_N)$ .

### 3.3 Validation of Proposed Method

Having outlined a method for constructing our confidence region  $S$  and a simulation procedure to estimate  $C_\alpha$ , we proceed by verifying that our method produces confidence regions of the appropriate confidence level.

#### 3.3.1 General Description of Validation Procedure

In order to test the effectiveness of our method, we first generate a realization of a Gaussian random field for a fixed covariance function over a very fine grid in our domain  $\mathcal{D}$ . The realized response surface is interpolated over  $\mathcal{D}$  and the level curve for this realization is constructed. For comparison purposes the level curve of this interpolated surface is used as the “true” level curve  $I_u$ .

From this realized “true” surface we randomly sample  $N$  locations and note the corresponding response. These sampled responses will be used as our observed responses  $z(s_1), \dots, z(s_N)$ . The observed responses are used to calculate kriging predictions over a fine grid of locations in the domain of interest, as well as to calculate the associated kriging error. For each of these locations we construct the test statistic

$$|Z'(s)| = \left| \frac{\hat{Z}(s) - u}{\hat{\sigma}(s)} \right|.$$

Using the procedure described in Section 3.2, we approximate the distribution of  $\{|Z'(s)|, s \in I_u\}$  through simulation and estimate our critical value  $C_\alpha$ . Using our estimated critical value  $\hat{C}_\alpha$  and the previously calculated test statistics, we construct a rejection region  $R = \{s : |Z'(s)| > \hat{C}_\alpha\}$ . Our confidence region then becomes  $S = R^c$ . If the true level curve  $I_u$  is contained in our confidence region  $S$ , then the simulation is considered a success. Our procedure should produce a confidence region that contains the true level curve in approximately  $(1 - \alpha) \times 100\%$  of the realizations of the “true” response surface.

### 3.3.2 Specifics of Validation Procedure

In order to test the effectiveness of our method we looked at validation results for Gaussian, spherical, and exponential covariance functions. Since  $Z'$  will have a standard normal distribution regardless of the covariance function of  $Z$ , we fixed the mean of  $Z$  to be zero, the partial sill  $\sigma^2 = 1$ , and the nugget  $\tau^2 = 0$  across all simulations. For each covariance function we ran simulations using six different values of the range parameter  $\phi$ : 0.1, 0.5, 1, 2, 5, and 10. To check whether our method worked for level curves having different response values, results were determined for levels curves when  $u$  was equal to 0, 1, and 2.

In order to maintain consistency with the validation simulations performed for the confidence boxes method proposed in Chapter 2, we once again let the domain of interest be a square region of size  $[0, 20] \times [0, 20]$ . The square region was first divided into a regular grid of  $255 \times 255$  squares. The initial unconditional realization of  $Z$  was generated at the vertices and center points of the squares, and the response surface was interpolated over these locations.

From the original set of locations we randomly selected 200 locations to use as our observed responses  $z(s_1), \dots, z(s_{200})$ . Using the procedure described in Section 3.2, we empirically estimated the critical value  $C_\alpha$  using 100 realizations of  $\{Z'(s)\}$  for  $\alpha$  equal to both 0.05 and 0.10.

To make construction of the confidence region more manageable,  $\mathcal{D}$  was divided into small, equal sized pixels. The center of each pixel is taken to be a good representative of the process over the pixel and the hypothesis test  $H_0 : Z(s) = u$  versus  $H_a : Z(s) \neq u$  was performed at the center of these pixels. The confidence region  $S$  was taken to be the union of all pixels for which the null hypothesis was not rejected at the center point. This grid of pixels was typically of size  $256 \times 256$ , but increased up to size  $600 \times 600$  for the Gaussian covariance function. After constructing the confidence region for each realization of  $Z$ , it was determined whether the confidence region contained the “true” level curve.

For each combination of covariance function and range parameter we generated 100 realizations of the true level curve  $I_u$ . We assessed the performance of our method by calculating the proportion of realizations in which the confidence region contained the true level curve.

### 3.3.3 Simulation Results for Hypothesis Testing

The results of our simulations are presented for each covariance function individually in the form of a table. The measurement of interest is the proportion of realizations in which the confidence region contained the true level curve. This measurement is provided for each possible combination of the range parameter  $\phi$ , the confidence level of the region, and the value  $u$  of the level curve under consideration. Each row of the table indicates the results for a different value of  $\phi$ . The first three columns of results are for the levels  $u = 0, 1, \text{ and } 2$  respectively at the 0.90 confidence level. Similarly, the last three columns of results are for the 0.95 confidence level. The empirical confidence levels should be close to 0.90 for the first three columns of results and close to 0.95 in the remaining columns.

#### Results for Spherical Covariance Function

The confidence regions for the spherical covariance function consistently attained empirical confidence levels close to the desired confidence levels. As shown in Table 3.1, across all values of  $\phi$  and  $u$  the empirical and desired confidence levels of our confidence regions were fairly close. Some of the individual results were slightly high or low, but there is no overall pattern of liberality or conservativeness among the results.

#### Results for Exponential Covariance Function

The validation results for the exponential covariance function were similar to the results for the spherical covariance function. The results are shown in Table 3.2. Once

Table 3.1: Simulation results for spherical covariance function.

Spherical						
Conf. Level	.90			.95		
$\phi \backslash u$	0	1	2	0	1	2
0.1	.94	.89	.92	.96	.92	.93
0.5	.93	.89	.90	.99	.96	.95
1	.93	.88	.88	.97	.94	.94
2	.88	.89	.87	.94	.93	.93
5	.87	.92	.97	.93	.95	.99
10	.92	.91	.90	.96	.96	.93

Table 3.2: Simulation results for exponential covariance function.

Exponential						
Conf. Level	.90			.95		
$\phi \backslash u$	0	1	2	0	1	2
0.1	.93	.92	.95	.98	.98	.97
0.5	.86	.94	.90	.93	.95	.96
1	.92	.91	.92	.95	.94	.96
2	.89	.91	.90	.94	.96	.96
5	.93	.91	.87	.95	.95	.92
10	.91	.87	.88	.96	.96	.91

Table 3.3: Empirical confidence levels for Gaussian covariance function.

Gaussian						
Conf. Level	.90			.95		
$\phi \backslash u$	0	1	2	0	1	2
0.1	.90	.85	.91	.98	.92	.93
0.5	.86	.90	.90	.89	.92	.93
1	.63	.76	.91	.69	.86	.92
2	.73	.75	.81	.86	.85	.90
5	.75	.89	.90	.84	.95	.94
10	.84	.95	.96	.89	.96	.99

again, the empirical and desired confidence levels are fairly close across all values of  $\phi$  and  $u$ .

### Results for Gaussian covariance function

The confidence regions for the Gaussian covariance function were not as reliable as the confidence regions for the other two covariance functions. The overall results are shown in Table 3.3. The empirical confidence level is sometimes much lower than the intended confidence level. However, the method actually performs better than Table 3.3 indicates. Visual results indicate that the confidence regions for the true level curve are too small by a small margin. Figure 3.2 shows a 90 percent confidence region for a level curve generated from a Gaussian random field having a Gaussian covariance function with  $\phi = 5$ . The confidence region  $S$  is indicated in grey while the solid line indicates the true level curve for this particular realization. The figure clearly shows that the confidence region does an excellent job at finding the true level curve; however, the confidence region is slightly too small and the true level curve strays slightly outside the confidence region. One of the locations where the level curve strays outside the confidence region is indicated by an arrow. An alternative measure of the effectiveness of this method is the proportion of the true level curve contained in the confidence region. This proportion is calculated by finding the length of the level curve contained inside the confidence region and dividing it by the total

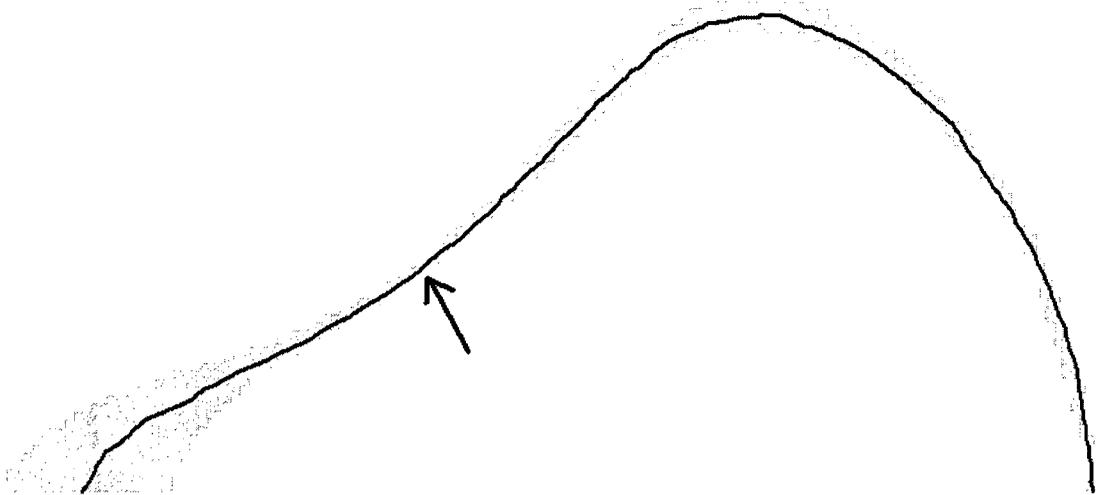
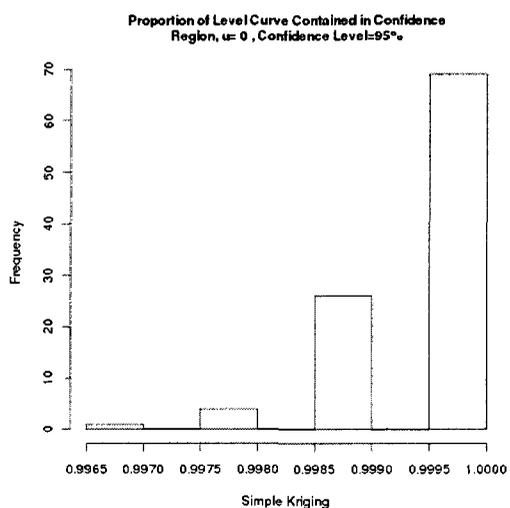


Figure 3.2: Confidence region for the true level curve. A 90 percent confidence region is shown in grey while the true level curve is given by the solid line. The true level curve strays slightly outside the confidence region. The arrow indicates a region where the level curve strays outside the confidence region.

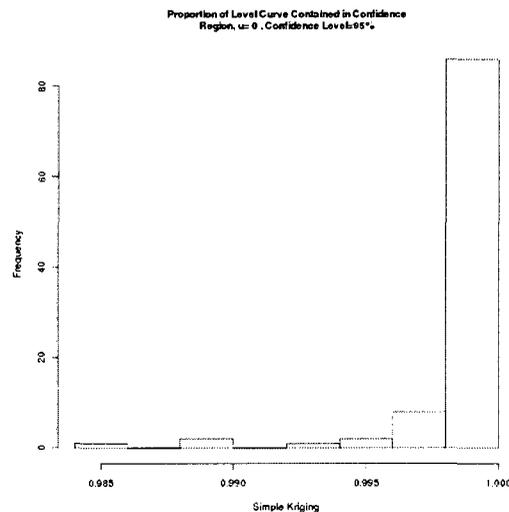
length of the level curve. If this proportion is high, then for practical purposes, the confidence region is successful in containing the true level curve. Histograms of this proportion for 95 percent confidence regions are shown in Figure 3.3 for Gaussian covariance functions with range parameter  $\phi = 1, 2, 5,$  and  $10$  for the level  $u = 0$ . It is seen from these histograms that the proportion of the true level curve contained by the confidence region is extremely high for nearly every realization of the level curve, and fully contained in a majority of the realizations. Thus, though the hypothesis test method does not consistently attain the desired confidence level for the Gaussian covariance function, the confidence regions produced by the method contain nearly the entire level curve with high probability.

### General Discussion of Results

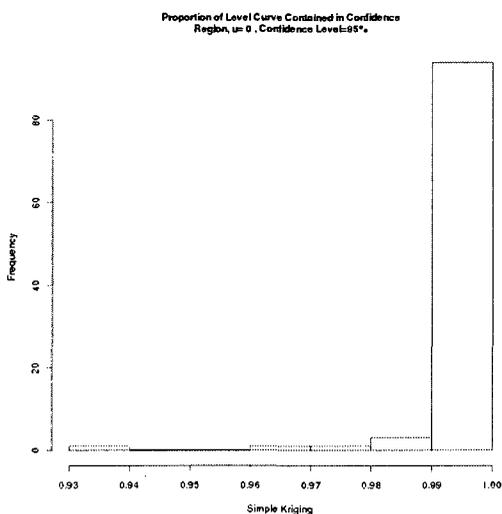
The hypothesis test method appears to be reasonably successful in constructing a confidence region containing the true level curve at the appropriate confidence level. This method is very stable for the spherical and exponential covariance functions, but more sensitive for the Gaussian covariance function. The empirical confidence level



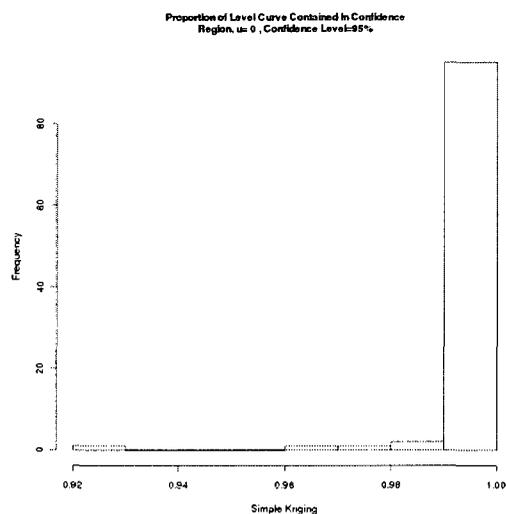
(a) Histogram for Gaussian covariance function with  $\phi = 1$ .



(b) Histogram for Gaussian covariance function with  $\phi = 2$ .



(c) Histogram for Gaussian covariance function with  $\phi = 5$ .



(d) Histogram for Gaussian covariance function with  $\phi = 10$ .

Figure 3.3: Histograms of the proportion of the true level curve contained in the corresponding confidence region. (a)–(d) are histograms of the proportion of the true level curve contained by the corresponding confidence region for one hundred realizations of the true level curve.

was typically close to the desired confidence level regardless of the range parameter for the spherical and exponential covariance functions. Confidence regions produced using the Gaussian covariance function did not always attain the nominal confidence level, but did contain a large proportion of the true level curve with high probability. Results seem to indicate that a reasonable estimate of  $C_\alpha$  can be obtained using only 100 realizations of  $\{Z'(s)\}$ .

### **Size of Confidence Region in Relation to the Covariance Function, Dependence, and Level**

There is a relationship between the precision (or geographical area) of the confidence region and the covariance function, dependence, and response level considered. The confidence regions are typically more precise for smoother covariance functions. In this case, the confidence regions are generally more precise for the Gaussian covariance function and less precise for the exponential and spherical covariance functions. The confidence regions also tend to be more precise when the range parameter  $\phi$  is larger. This pattern was consistent across simulations for the Gaussian, spherical, and exponential covariance functions. In Figure 3.4 we see 90 percent confidence regions constructed for the spherical covariance function using different values of the range parameter  $\phi$ . The rejection region is shown in black and the confidence region in white, while the solid lines represents the true level curve for  $u = 1$  for this particular realization of the random field. Notice that for  $\phi = 0.1$ , the confidence region makes up almost the entire region of interest (the rejection region is difficult to see because of the scale of the figure). However, as the range parameter  $\phi$  increases, we clearly see that the confidence regions become smaller and tighter around the true level curve. Lastly, the confidence regions tend to become more precise as the level  $u$  increases. To display this relationship, we constructed confidence regions for level curves at levels  $u = 0, 1, \text{ and } 2$  for random fields having a spherical covariance function with  $\sigma^2 = 1, \phi = 5, \text{ and } \tau^2 = 1$ . We compare the relationship between the precision of

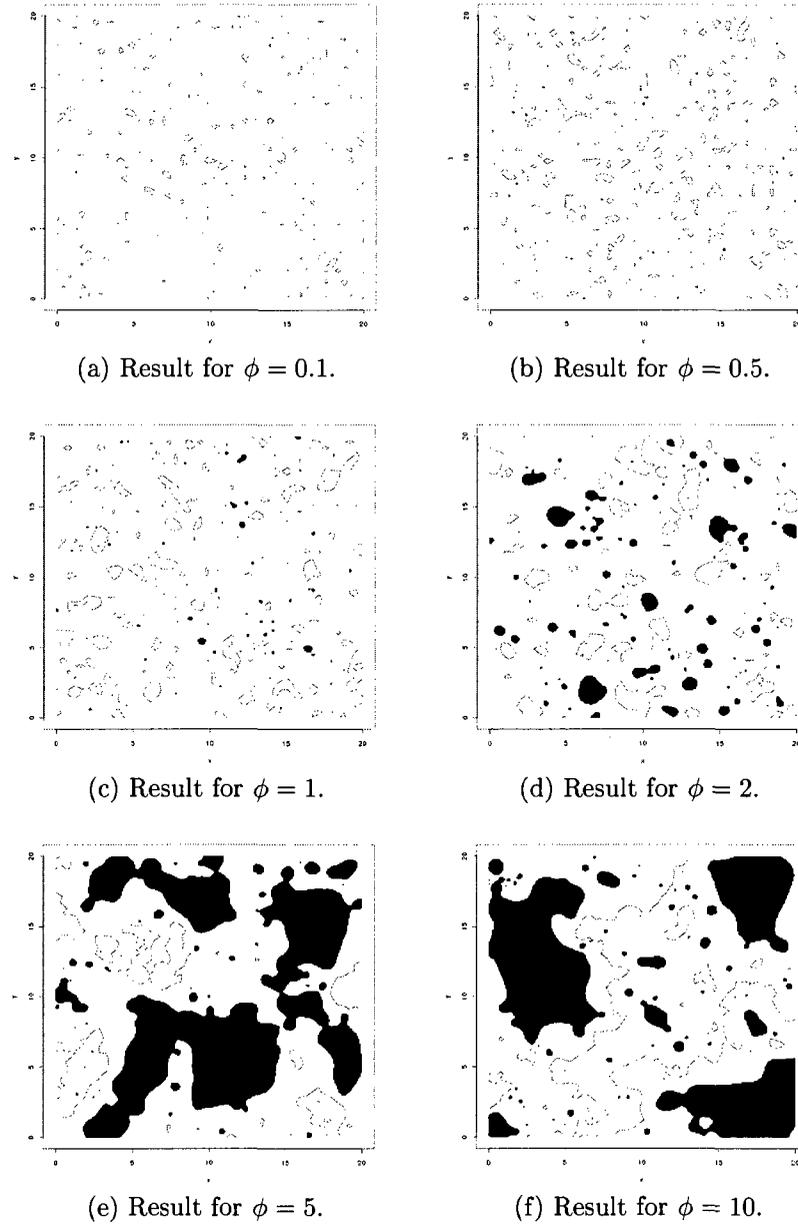


Figure 3.4: Comparison of confidence regions constructed for varying degrees of dependence. 90 percent confidence regions for the true level curve are shown in white and the rejection regions are shown in black. The true level curve under consideration is represented a solid line.

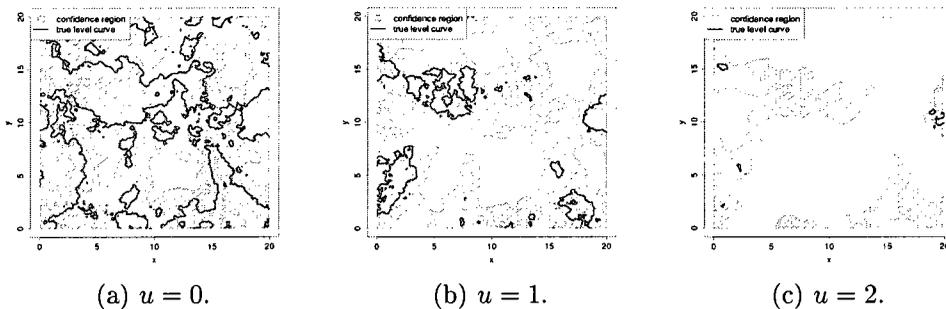


Figure 3.5: Relationship between response level of level curve and precision of confidence region. The confidence regions were constructed for the spherical covariance function with  $\phi = 5$ . The level  $u$  of the true level curve is specified in each figure.

a confidence region and the response level  $u$  of the level curve in Figure 3.5. When the response value of a level curve is more extreme, we can expect that there will be fewer locations having response values close to this level and it will be easier to exclude locations from our confidence region. This pattern is clearly demonstrated in Figure 3.5 as the level  $u$  increases.

### 3.4 Use of Adler (2008) to Estimate $C_\alpha$

As mentioned in Section 3.1, consideration was given to using the excursion probability approximation for isotropic Gaussian random fields presented in Adler (2008) to approximate  $C_\alpha$ . Adler provides a simple, closed form expression for the excursion probability when the covariance function of the random field is isotropic, the domain of interest is a rectangle, and a few additional assumptions are met. We seek to find  $C_\alpha$  such that  $\mathbb{P}(\sup_{s \in I_u} |Z'(s)| \geq C_\alpha) = 1 - \alpha$ . We will explore approximating  $C_\alpha$  by  $C_{\alpha/2}^*$  such that  $\mathbb{P}(\sup_{s \in \mathcal{D}} Z(s) \geq C_{\alpha/2}^*) = 1 - \alpha/2$ , where  $Z(s)$  is a stationary Gaussian random field on a rectangular domain of interest with an isotropic covariance function. The additional conditions necessary to use Adler's approximation are all generally met in our current setting, except that  $Z(s)$  must have twice-differentiable sample paths almost surely. Because of this, we restrict our consideration to the case where  $Z$  has a Gaussian covariance function;  $Z$  is not differentiable for the spherical

Table 3.4: Estimates of  $C_{\alpha/2}^*$  using a stationary, isotropic approximation of  $Z'$ .

$\phi$	$C_{.05}^*$	$C_{.025}^*$
0.1	5.13	5.27
0.5	4.43	4.59
1	4.02	4.27
2	3.72	3.92
5	3.26	3.42
10	2.79	3.04

and exponential covariance functions. Under the assumptions that the domain of interest is a  $[0, 20] \times [0, 20]$  region and the covariance function of  $Z$  is isotropic Gaussian with  $\sigma^2 = 1$  and  $\tau^2 = 0$ , the excursion probability

$$\mathbb{P}\left(\sup_{s \in \mathcal{D}} Z(s) \geq c\right) \approx \left(\frac{20^2 \lambda_2}{(2\pi)^{3/2}} c + \frac{2(20)\sqrt{\lambda_2}}{2\pi}\right) e^{-\frac{c^2}{2}} + 1 - \Phi(c), \quad (3.4)$$

where

$$\lambda_2 = -\left.\frac{\partial^2 C(h)}{\partial x_i^2}\right|_{h=0} = \frac{2}{\phi^2},$$

and  $\Phi(\cdot)$  is the cumulative distribution function of a standard normal random variable. Our estimate of  $C_{\alpha/2}^*$  will be  $c$  such that  $\mathbb{P}(\sup_{s \in \mathcal{D}} Z(s) \geq c) = \alpha/2$ . An estimate of  $C_{\alpha/2}^*$  was obtained for  $\phi = 0.1, 0.5, 1, 2, 5$ , and  $10$  and  $\alpha = 0.05$  and  $0.10$ . Results are shown in Table 3.4.

To assess the accuracy of this approximation in estimating  $C_\alpha$ , we compared the estimates in Table 3.4 to estimates of  $C_\alpha$  obtained via simulation. An important consideration is that the value of  $C_\alpha$  is dependent on the distribution of  $Z'$ , which is in turn dependent on the locations of the observed responses  $z(s_1), \dots, z(s_N)$ . In order to take this into account, estimates of  $C_\alpha$  were obtained for 20 independent realizations of a Gaussian random field  $Z$ . The level curve under consideration for each realization was for the level  $u = 2$ . For each of these realizations, 200 responses were randomly selected and used to estimate  $C_{.10}$  and  $C_{.05}$  through the simulation process outlined in Section 3.2. Boxplots of the resulting estimates are shown for  $\phi = 0.1, 1$  and  $2$  in Figure 3.6. Individual estimates of  $C_\alpha$  are shown as dashed

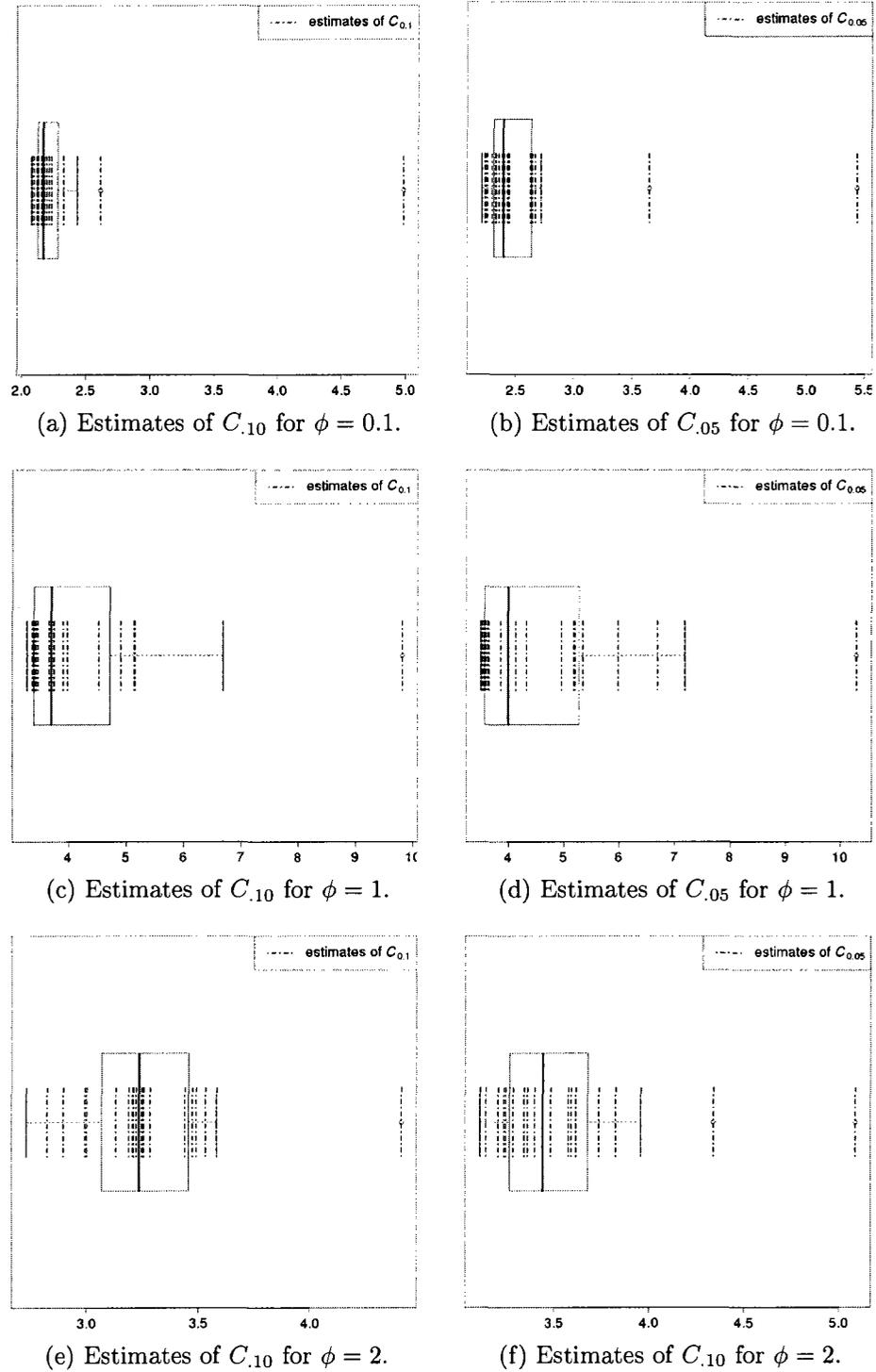


Figure 3.6: Estimates of  $C_{.10}$  and  $C_{.05}$ . The estimates were obtained via simulation for processes having Gaussian covariance functions with  $\sigma^2 = 1$  and  $\tau^2 = 0$ . The range parameter  $\phi$  is specified in each figure.

vertical lines. Comparing the individual estimates computed via simulation to the estimates using the approximation of Adler, we see that the individual estimates are often quite different than the Adler approximation. When  $\phi = 0.1$ , the estimates of  $C_{.10}$  and  $C_{.05}$  obtained using the approximation of Adler were nearly always extremely conservative compared to the simulated estimates. With the exception of one large outlier, the simulation estimates of  $C_{.10}$  ranged from 2.08 to 2.61, which is well below the estimate of 5.13 calculated using Adler's approximation. Results are similar for  $C_{.05}$ . For  $\phi = 1$ , the simulation estimates were often considerably different than the estimates obtained using the approximation of Adler. While the estimates of  $C_{.10}$  and  $C_{.05}$  for the Adler approximation were 4.02 and 4.27 respectively, the simulation estimates ranged between 3.39 to 9.81 for  $C_{.10}$  and 3.51 to 10.29 for  $C_{.05}$ . For  $\phi = 2$ , the simulation estimates were typically smaller than the Adler estimates for both  $C_{.10}$  and  $C_{.05}$ , though there were a few simulation estimates which were noticeably higher than the estimates obtained using the result of Adler. In general, it does not appear that an accurate approximation of  $C_\alpha$  can be obtained using the current version of the approximated excursion probability. However, as previously noted, the result provided in Adler (2008) does not require stationarity a priori; the isotropic approximation was used because it is very easy to calculate. If a similar result can be derived for the nonstationary case, then that approximation may produce more accurate estimates of  $C_\alpha$ .

### 3.5 Future Alternative Error Criteria for Consideration

The true location of a level curve can be of great importance. A level curve may be the boundary separating a region with a high risk of severe weather from a region with moderate risk. A level curve may separate locations where air pollution levels are acceptable versus dangerous in a particular county. It may separate contaminated locations from uncontaminated locations in hazardous waste cleanup. With these settings in mind, it is important that the confidence regions are precise (i.e., have

small geographical area). An accurate confidence region which is not very precise may not be helpful in practice.

Consider the hazardous waste example previously mentioned. The level curve separates dangerous locations from safe locations. It is important that cleanup occur at all contaminated locations for the safety of the community. However, it is also important that cleanup not be performed unnecessarily at safe locations to conserve time, money, and resources. If the geographical area of the confidence region for the level curve is large, the cleanup may prove too costly to attempt or there may not be enough resources to accomplish the cleanup.

The error rate controlled by the hypothesis test method is the well known familywise error rate. Suppose that we are performing multiple hypothesis tests. Define  $V_n$  to be the number of tests that reject  $H_0$  when  $H_0$  is in fact true. The familywise error rate (FWER) is  $\mathbb{P}(V_n > 0)$  and typically controlled so that  $\mathbb{P}(V_n > 0) \leq \alpha$ . In our specific situation, we inverted hypothesis tests to construct a confidence region for the true level curve. By controlling the FWER at  $\alpha$  for our tests, we were able to construct confidence regions that contained the true level curve with approximate probability  $1 - \alpha$ . Benjamini and Hochberg (1995) point out that controlling the familywise error rate can dramatically reduce the power of a test in comparison to the per comparison error rate (controlling the Type I error rate for each test individually). In our setting, controlling the FWER may result in including a large proportion of non-null locations in our confidence region. Benjamini and Hochberg also point out that in many settings, control of the FWER is not really needed.

Probably the best known alternative to the FWER is the false discovery rate (FDR) proposed by Benjamini and Hochberg (1995). Define  $R_n$  to be the number of rejected hypotheses among  $M$  hypothesis tests. FDR controls  $\mathbb{E}(V_n/R_n)$ , the expected proportion of false discoveries among the  $M$  hypothesis tests. This error criterion is quite popular because it can offer substantial gains in power compared to situations where FWER is controlled. Unfortunately, this criterion does not appear

to be useful for controlling the error rate of the hypothesis tests in our current setting. In practice, our confidence region is constructed by pixelating the domain of interest and performing a hypothesis test at the center point of each pixel. When a pixel is included in our rejection region, our belief is that the true level curve does not pass through that pixel. Thus, we may define a “false discovery” to be the inclusion of a pixel in our rejection region when in fact the true level curve passes through that pixel. In general, the number of pixels through which the level curve passes will be small relative to the total number pixels. In other words, the proportion of pixels intersecting the level curve is generally small. If the proportion of pixels intersecting the true level curve relative to the total number of pixels is expected to be less than  $\alpha$ , then one could include all pixels in the rejection region and still control the FDR (since the proportion of false discoveries would be less than  $\alpha$ ). Thus, the usefulness of this error criterion is questionable when seeking to construct confidence regions for the true level curve.

An alternative error criterion which should be more useful in our current setting is related to the per comparison error rate (PCER, cf. Dudoit et al. (2004)). The PCER controls the expected proportion of Type I errors in a multiple testing problem. When constructing confidence regions, it would be useful to ensure that our confidence region can be expected to contain a large proportion of the true level curve. In other words, when conducting hypothesis tests at locations on the level curve, the proportion of Type I errors made should be relatively small. If we can control the expected proportion of Type I errors at locations on the level curve, we can construct a confidence region which is expected to contain a large proportion of the true level curve. A similar error criterion which should be useful is one in which the probability is high that the confidence region contains a large proportion of the true level curve. If we were able to control an error criterion of this type, we can have high confidence that our confidence region contains a large proportion of the level curve. This is a

desirable characteristic for our confidence regions since a confidence region may not be useful if the proportion of the level curve contained in the confidence region is low.

## Chapter 4

### APPLICATION AND COMPARISON OF METHODS

#### 4.1 Introduction

We will apply and compare the box method proposed in Chapter 2 with the hypothesis testing method proposed in Chapter 3 to three different data sets. The first two data sets are simulated, while the third data set is real data related to heavy metal concentrations in the Swiss Jura. For each of the three data sets, the dimensions of the confidence boxes were determined using 1000 realizations of the conditional process  $Z(s) \mid z(s_1), \dots, z(s_N)$  and the critical value used in the hypothesis test method was estimated using 100 realizations of the random field  $\{Z'(s)\}$ . All confidence regions were constructed to satisfy a confidence level of 0.90.

#### 4.2 Application to Simulated Data

The simulated data sets are similar to Example 1 in Lindgren and Rychlik (1995). For the first data set, we simulated 20 observations from a Gaussian random field having mean zero and an isotropic exponential covariance function with  $\sigma^2 = 1$  and range parameter  $\phi = 1$ . The second data set was obtained by adding random “measurement” error to the responses of the first data set. The random errors were generated independently from a Gaussian random variable having mean zero and variance 0.2. Thus, the covariance function for the responses of the second data set is exponential with  $\sigma^2 = 1$ ,  $\phi = 1$ , and  $\tau^2 = 0.2$ . The covariance function of the first data set is shown in Figure 4.1. By comparing the results for these two data sets, we can assess

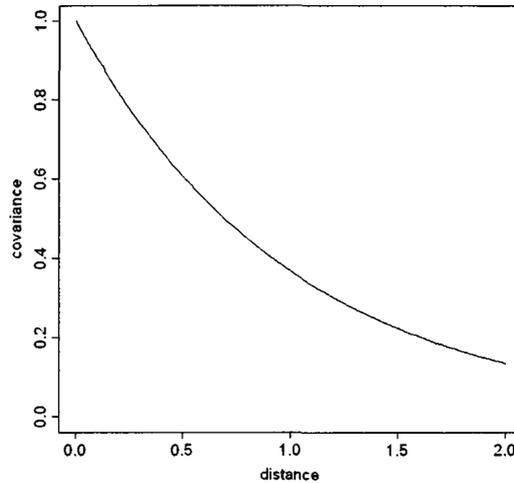
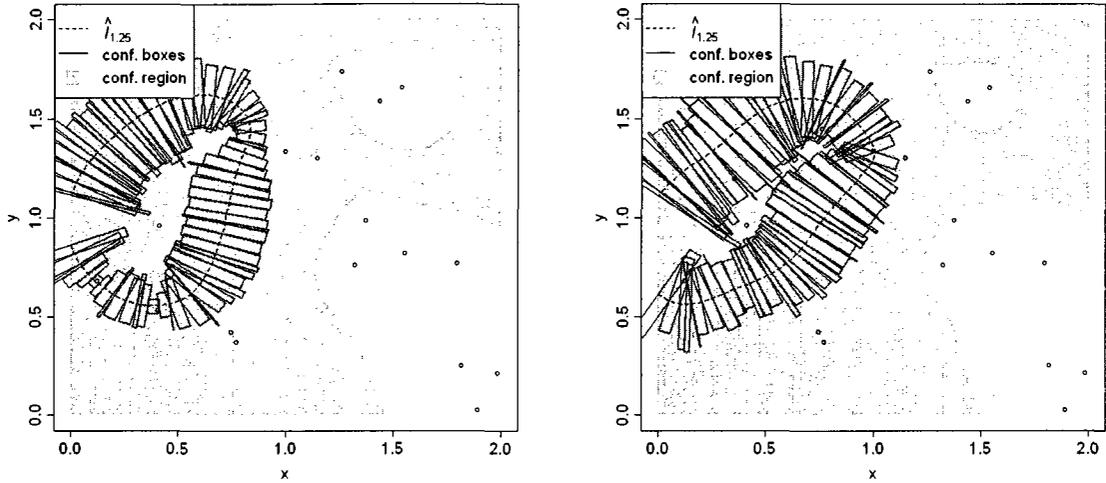


Figure 4.1: Covariance function used to generate simulated data. The covariance function used to generate the simulated data is exponential with  $\sigma^2 = 1$  and  $\phi = 1$ .

the impact of the nugget on the confidence regions produced by the two methods. For both simulated data sets the level curve of study was for the level  $u = 1.25$ , and the estimated level curve was constructed using simple kriging.

The results for the simulated data sets can best be summarized visually. The results of both methods for the first data set are shown in Figure 4.2 (a) and for the second data set in Figure 4.2 (b). Before comparing the results for the two figures, we note that the estimated level curve is different for the two data sets because of the measurement error in the second data set.

There are two notable features of the results in Figures 4.2 (a) and (b). The first is that the union of the confidence boxes is much smaller than the confidence region produced using the hypothesis test method. This is expected since the confidence region produced by the hypothesis test method controls for multiple comparisons while the confidence boxes are constructed independently of each other and no multiple comparison adjustment is made. The next notable feature of the figures is that for both methods, measurement error leads to larger confidence regions. The confidence regions in Figure 4.2 (a) for both methods are noticeably smaller than the corresponding confidence regions in Figure 4.2 (b) when the responses are observed with



(a) Results for data simulated without measurement error.

(b) Results for data simulated with measurement error.

Figure 4.2: Application of methods to simulated data.

measurement error. For the confidence boxes method, the measurement error in the observed responses of the second data set causes the variability in the realizations of  $Z(s) \mid z(s_1), \dots, z(s_{20})$  to be greater than when no measurement error is present. Because the variability in the conditional process is greater, the confidence boxes must extend farther to ensure that the boxes intersect the true level curve at the appropriate confidence level. For the hypothesis test method, the measurement error causes the kriging variance to be larger throughout the domain of interest. Thus, the test statistic  $Z'(s) = |\hat{Z}(s) - 1.25|/\hat{\sigma}(s)$  tends to be smaller, and it becomes more difficult to reject the null hypothesis that the response value at a location is different than  $u = 1.25$ .

### 4.3 Application to Jura Data

The third example in this chapter involves data collected by the Swiss Federal Institute of Technology in a 14.5 km<sup>2</sup> region in the Swiss Jura. Concentrations of cadmium, cobalt, chromium, copper, nickel, lead, and zinc in the topsoil were measured at 359 locations; additional covariates related to land use and geology were also

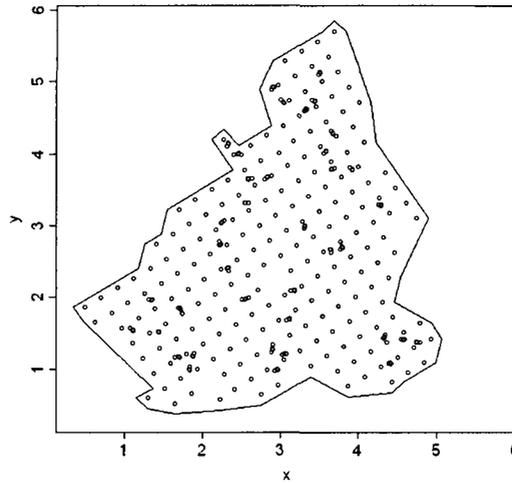
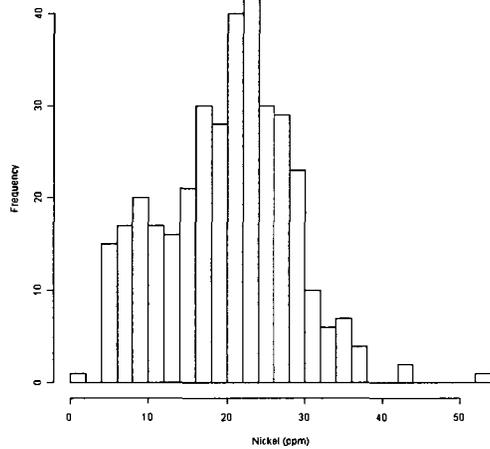


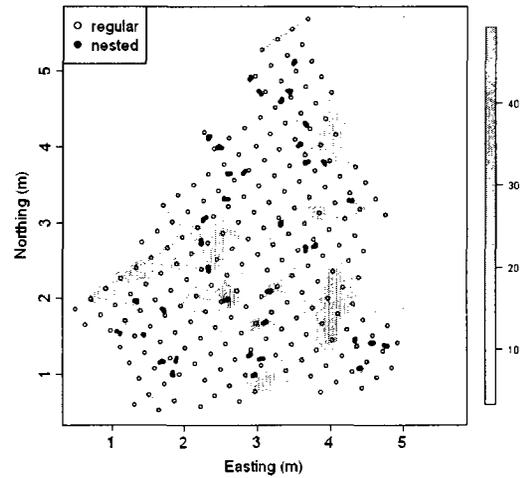
Figure 4.3: Map of sampled locations for Swiss Jura data.

measured. A map of the sampled locations is shown in Figure 4.3. Our analysis will focus on the nickel variable (Ni) measured in parts per million (ppm). The sampling and experimental details of this data are described in Atteia et al. (1994) and Webster et al. (1994).

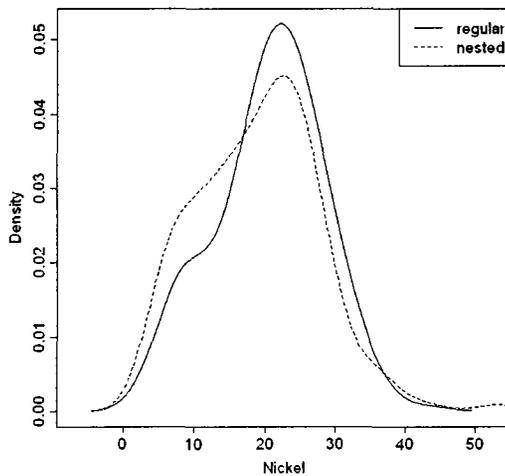
We begin by checking whether the Jura data satisfies the conditions necessary to use the two proposed methods. The necessary conditions are that the data come from a Gaussian distribution and also be stationary. We first check whether the data appears to come from a Gaussian distribution. A histogram of the nickel variable, shown in Figure 4.4 (a), is positively skewed and does not appear to follow a Gaussian distribution. In the description of the sampling design found in Atteia et al. (1994), we learn that the observed locations are a combination of locations on a regular grid and nested sampling locations. Preferential sampling can skew the results of exploratory data analysis (cf. Goovaerts (1997, p. 76, 77)), so we continue by dividing the observed responses into locations on the regular grid and nested locations. Plotting these responses on an image plot of the nickel variable, shown in Figure 4.4 (b), we see that the responses at nested locations tend to be lower than responses at locations on the regular grid. For more formal confirmation of this observation, the kernel density



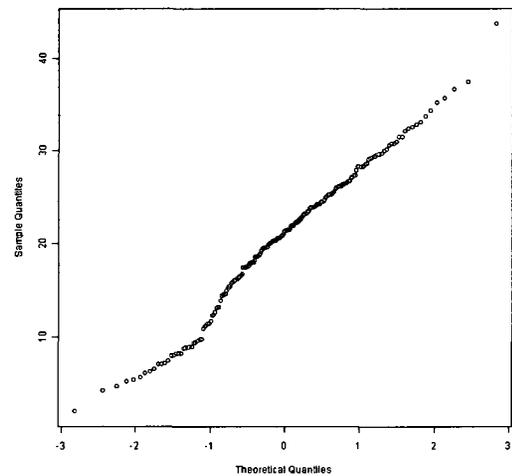
(a) Histogram of nickel variable.



(b) Image plot of nickel variable with observed locations separated into locations on the regular grid and nested locations.



(c) Kernel density estimates of the nickel variable for locations on the regular grid and nested locations.



(d) Normal probability plot of transformed data at locations on the regular grid.

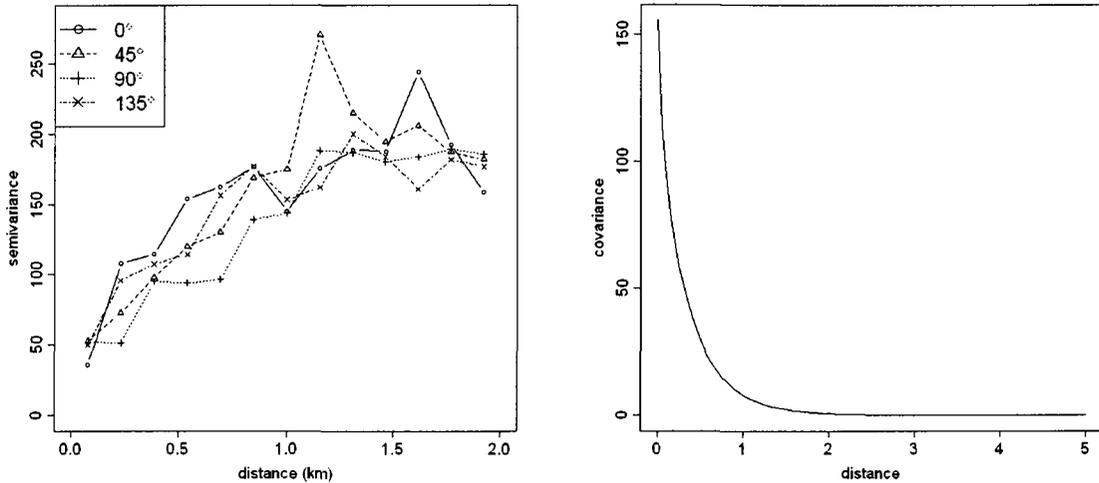
Figure 4.4: Exploratory plots for Jura data.

estimates of the nickel variable were plotted for the two types of locations, as shown in Figure 4.4 (c). The density estimates clearly indicate that lower valued responses are observed more commonly at the nested locations. The responses at nested locations give the impression that the nickel variable is non-Gaussian, whereas the values at locations on the regular grid are much closer to a Gaussian distribution. Because the nested locations can skew analysis, for the purpose of assessing normality we proceed by working with only data values found at locations on the regular grid (cf. Goovaerts (1997, p. 79)). To normalize the data at regular grid locations, we considered the Box-Cox transformation

$$z_i^{(\lambda)} = \frac{y_i^\lambda - 1}{\lambda}, \quad \lambda \neq 0,$$

proposed by Box and Cox (1964), where  $y = (y_1, \dots, y_n)$  is the original data vector under consideration. A power parameter value of 1.15 brings the data at regular locations to approximate normality, as shown by the normal probability plot in Figure 4.4 (d). We proceed under the assumption that the nickel variable follows a Gaussian distribution after performing a Box-Cox transform with  $\lambda = 1.15$ . Subsequent analysis is performed on the transformed data.

To check whether the data is stationary, we must check that the mean is constant and that the covariance is finite and depends only on a displacement vector  $\mathbf{h}$ . To verify that the mean appears constant over the region of interest, we constructed an image plot of the transformed data (not shown). The resulting plot was very similar to the one shown in Figure 4.4 (b). There was no clear indication of a long range trend in the data, and as done by Atteia et al. (1994), we assume a constant mean over the domain of interest. To check to whether the covariance was finite and did not depend on the absolute locations, we plotted directional variograms at angles  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ , where a variogram is a plot of the distance  $h$  against an estimate of  $\frac{1}{2}\text{Var}(Z(s+\mathbf{h}) - Z(s))$ . These are shown in Figure 4.5 (a). The variograms level off as the distance increases, so it is reasonable to assume finite covariance. Because the



(a) Directional variograms of transformed nickel variable. Directions are indicated in the legend. (b) Estimated Matérn covariance function using Restricted Maximum Likelihood.

Figure 4.5: Spatial variation of Jura data.

structure of the directional variograms is similar in each direction, we can conclude that the covariance is isotropic.

Following the recommendation of Stein (1999), we fit a Matérn covariance model to the data and estimated the corresponding parameters using Restricted Maximum Likelihood (REML). The estimated covariance function is shown in Figure 4.5 (b) and the best fitting covariance model is given by

$$C(h) = 155.39 \left( \frac{h}{0.427} \right)^{0.293} \mathcal{K}_{0.293} \left( -\frac{h}{0.427} \right), \quad h \geq 0, \quad (4.1)$$

where the partial sill  $\sigma^2 = 155.39$ , the range parameter  $\phi = 0.427$ , the smoothness parameter  $\nu = .293$ , and the nugget  $\tau^2 = 0$ . The smoothness parameter is quite small, so the underlying Gaussian process is believed to be quite rough. The process will not be differentiable since the smoothness parameter  $\nu < 3$ .

Having verified the appropriate assumptions and estimated the covariance function of the data, we proceed by implementing our methods. Because the true mean is unknown for this data, we estimated the mean using REML, and then used this mean as if it were the true mean. The REML estimate of the mean was 27.9. Simple

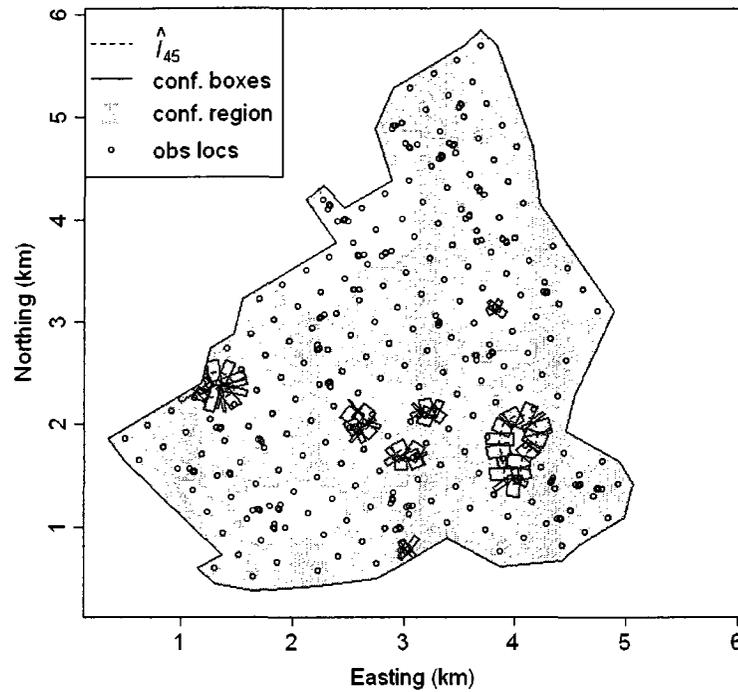


Figure 4.6: Confidence regions for  $I_{45}$  for the Jura data.

kriging was performed over the domain of interest using the covariance specified in (4.1), and the prediction surface was used to construct the estimated level curve  $\hat{I}_{45}$  (the value 45 is around the 93rd percentile of the transformed data). Using the two different methods to construct confidence regions for the true level curve results in the confidence regions shown in Figure 4.6.

The estimated level curve  $\hat{I}_{45}$  appears at many locations in the lower half of the domain of interest. As might be expected, the confidence boxes are substantially smaller than the confidence region constructed using the hypothesis test method. The confidence boxes are fairly tight around the estimated level curve and it is reasonable to conclude that responses values of 45 lie very close to the estimated level curve. The confidence region for the hypothesis test method is larger and includes a large portion of the domain of interest. Because the underlying process is so rough (rougher than a process with exponential covariance function), the kriging variance tends to be high

throughout the domain of interest (typically 20 or greater), making it more difficult to conclude that the response value at a location is different than  $u = 45$ .

The confidence regions produced by the two methods give complementary information about the Jura data. If one is interested in finding some locations having a response value of 45, the confidence boxes indicate that these locations can be found close to the estimated level curve. On the other hand, if one is interested in finding all locations where the response value is 45, based on the available information, the search region must be much larger.

#### 4.4 Comparison of the Two Methods

Having described and applied both the confidence box and hypothesis test methods for constructing confidence regions for level curves, we now elaborate on some of their strengths and weaknesses.

One of the strengths of the confidence box method is that the confidence boxes are often relatively small, especially in comparison to the confidence regions produced by the hypothesis test method. When we are only interested in getting a general idea of the uncertainty in our estimated level curves, the confidence boxes may be more useful than larger confidence regions controlling a stricter error criterion. Another strength is that the confidence boxes can often be constructed in a short period of time. Though the actual time will depend on factors such as the covariance structure, response level of the level curve, and the size of the domain of interest, the confidence boxes require less computational effort. The main weakness of the confidence box method is that it fails to adjust for multiple comparisons. Though this could be done by using a Bonferroni-like adjustment, the resulting boxes would be extremely wide and likely be of little practical use. Another weakness is that the confidence box method only produces boxes around the estimated level curve and may not find parts of the true level curve not near the estimated level curve. It is very plausible that part of the true level curve may reside in regions where predicted responses are close

to the level  $u$  and the mean squared prediction error is relatively high. However, since the estimated level curve may not be near this region, one may never know that part of the level curve may be in this region.

One of the most obvious strengths of the hypothesis test method is that it constructs a confidence region for the entire level curve. When someone is interested in the location of the entire level curve, the hypothesis test method is more appropriate than the confidence box method which only requires that the confidence boxes intersect the level curve with high probability. Another advantage of the hypothesis test method is that it can “find” the level curve in places where the confidence box method cannot. If part of the true level curve is in a region far from the estimated level curve, it is likely the confidence boxes will never intersect the level curve in this region. However, by taking into account both the predicted response and the mean squared prediction error, the confidence region produced by the hypothesis test method will contain this part of the level curve with high confidence. One of the weaknesses of the hypothesis test method is that construction of the confidence region requires more computational effort than construction of the confidence boxes. It took about 15 minutes to construct a confidence region for the hypothesis test method on a computer with a Pentium IV processor running at 2.4 GHz. Another downside of the hypothesis test method is that it is limited to Gaussian processes. Because many data sets do not follow a Gaussian distribution, it will not always be appropriate to use this method. Another weakness of the hypothesis test method is that the error criterion might be unnecessarily strict. Controlling the Type I error rate for the entire level curve can lead to large confidence regions. As discussed in Section 3.5, methods using alternative error criteria may be more useful in some situations.

## Chapter 5

### THE ASYMPTOTIC DISTRIBUTION OF THE MAXIMA OF A TRIANGULAR SEQUENCE OF GAUSSIAN RANDOM FIELDS

#### 5.1 Introduction

In this chapter we develop theory related to the asymptotic distribution of the maxima of a triangular sequence of Gaussian random fields satisfying certain conditions. This result is an extension of the result in Hsing et al. (1996) to two dimensions.

The behavior of the extremes of Gaussian random variables has long been of interest. Theory related to the excursion probability

$$\mathbb{P}\left(\sup_{t \in T} Z(t) \geq u\right), \quad (5.1)$$

has a rich literature history, where  $Z$  is a real-valued Gaussian random process over some domain of interest  $T$  (which may be continuous or discrete and multidimensional). One-dimensional results include those by Rice (1944), Berman (1964), Rootzen (1983), and Hsing et al. (1996). For a helpful overview of research involving the extremal behavior of Gaussian random fields before 2000, one may consult Adler (2000). More recent research includes Takemura and Kuriki (2002, 2003); Taylor et al. (2005), and Taylor (2006), which are important works leading to a very general result by Adler and Taylor (2007) for the probability in (5.1) when  $T \subset \mathbb{R}^d$ , for  $d \geq 1$ .

Our goal in this chapter is to extend the results of Hsing et al. (1996) to the two-dimensional setting and approximate (5.1) when  $T$  is an  $n \times n$  lattice in  $\mathbb{R}^2$ . As stated in Hsing et al. (1996), the objective of their paper is to, "... develop an

asymptotic theory for the extremes of a normal sequence which addresses the issue of clustering.” They consider a triangular array of standard normal random variables  $\xi_{n,i}$ ,  $i = 0, 1, 2, \dots$ , and  $n = 1, 2, 3, \dots$ , such that for each  $n$ ,  $\{\xi_{n,i}, i \geq 0\}$  is a stationary normal sequence. Each row is weakly dependent so that for fixed  $n$ , the extremes of  $\xi_{n,1}, \dots, \xi_{n,N}$  do not cluster as  $N \rightarrow \infty$ , but on the other hand, as the row number  $n$  increases, the correlation between neighboring observations increases to one so that local dependence cannot be ignored in the limit. Under certain conditions on the covariance function  $\rho_{n,j} = \mathbb{E}(\xi_{n,i}\xi_{n,i+j})$ , and assuming that

$$(1 - \rho_{n,j}) \log n \rightarrow \delta_j \in (0, \infty] \quad \text{for all } j \geq 1 \text{ as } n \rightarrow \infty,$$

where  $\delta_j$  is some function depending on  $j$ , Hsing et al. (1996) show that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\max_{i=1, \dots, n} \xi_{n,i} \leq u_n(x)\right) = \exp(-\vartheta \exp(-x)), \quad -\infty < x < \infty,$$

where

$$\vartheta = \mathbb{P}(E/2 + \sqrt{\delta_k}W_k \leq \delta_k \text{ for all } k \geq 1 \text{ such that } \delta_k < \infty),$$

with

$$u_n(x) = \sqrt{2 \log n} + \frac{x}{\sqrt{2 \log n}} - \frac{\log \log n + \log 4\pi}{\sqrt{8 \log n}}.$$

The quantity  $\vartheta$  is called the extremal index and depends on a standard exponential random variable  $E$  which is independent of the  $W_k$ , a Gaussian sequence with mean zero and correlation

$$\mathbb{E}(W_i W_j) = \frac{\delta_i + \delta_j - \delta_{|i-j|}}{2\sqrt{\delta_i \delta_j}}.$$

Using analogues of the conditions in Hsing et al. (1996) for random fields, and making some additional assumptions, we are able to produce a similar result in two-dimensions.

## 5.2 Motivating Example

Consider a Gaussian random field  $\{Y(s)\}$ , where  $s$  represents location in the domain, and assume that  $Y(s)$  has isotropic correlation function  $\exp(-h^2/\phi^2)$ . Suppose we rescale the domain by  $\sqrt{\log n}$  in the  $x$ - and  $y$ -directions. Define  $Y_n(s) = Y(s/\sqrt{\log n})$ . Then

$$\begin{aligned} \text{corr}(Y_n(s), Y_n(t)) &= \text{corr}(Y(s/\sqrt{\log n}), Y(t/\sqrt{\log n})) \\ &= \exp\left(-\frac{|s-t|^2}{\phi^2 \log n}\right). \end{aligned}$$

For fixed  $n$ , the process  $Y_n(s)$  is just an isotropic random field. However, as  $n \rightarrow \infty$ , the correlation structure of  $Y_n(s)$  satisfies the conditions necessary to use the limit result in Hsing et al. (1996).

## 5.3 Setup

Let  $X_n = \{\xi_{i,j,n}\}_{i,j=1}^n$  be a stationary Gaussian random field on an  $n \times n$  lattice, with  $\xi_{i,j,n}$  having mean zero and variance one.  $\xi_{i,j,n}$  is the random variable located at position  $(i, j)$  of the lattice. Assume the covariance function of  $X_n$  is isotropic and let  $\rho_{i,j,n}$  denote the correlation between two random variables separated by  $i$  units in the horizontal direction and  $j$  units in the vertical direction. By isotropy, we have that  $\rho_{i,j,n} = \mathbb{E}(\xi_{k,l,n} \xi_{k+i,l+j,n})$ . Define  $\lim_{n \rightarrow \infty} (1 - \rho_{i,j,n}) \log n = \delta_{i,j}$ , where  $\delta_{i,j}$  is some function of  $i$  and  $j$ .

We define  $M_n$  to be the maximum of the  $n \times n$  lattice of stationary random variables, i.e.,

$$\begin{aligned} M_n &= \max(X_n) \\ &= \max\{\xi_{i,j,n}\}_{i,j=1}^n. \end{aligned}$$

Lastly, we define

$$u_n(x) = \sqrt{2 \log n} + \frac{x}{\sqrt{2 \log n}} - \frac{\log \log n + \log 4\pi}{\sqrt{8 \log n}},$$

a normalizing sequence used in subsequent results. For ease of notation,  $u_n(x)$  will generally be denoted as  $u_n$ , except when the exact value of  $x$  is important.

## 5.4 Main Results

The main theorem of this chapter, stated below, describes the limiting distribution of the maximum of a stationary Gaussian random field on an  $n \times n$  lattice.

**Theorem 5.1.** *Let  $X_n = \{\xi_{i,j,n}\}_{i,j=1}^n$  be a stationary Gaussian random field on an  $n \times n$  lattice, with  $\xi_{i,j,n}$  having mean zero and variance one. Assume the covariance function of  $X_n$  is isotropic and let  $\rho_{i,j,n}$  denote the correlation between two random variables separated by  $i$  units in the horizontal direction and  $j$  units in the vertical direction. Define*

$$A_{\alpha,n} = \{1, \dots, q_n\}^2 \setminus \{\{1, 2, \dots, \alpha\} \times \{1\}\},$$

$$G_{\alpha,m,n} = \{(i, j) \in A_{\alpha,n} : |i - \alpha| \text{ and } (j - 1) \leq m\},$$

$$H_{\alpha,m,n} = A_{\alpha,n} \setminus G_{\alpha,m,n},$$

and let  $p_n = o(n)$ ,  $l_n = o(p_n)$ , and  $q_n = p_n + l_n$ .

Assume that:

$$(A1) \quad \rho_{i,j,n} \geq 0,$$

$$(A2) \quad \rho_{1,0,n} \geq \rho_{i,j,n} \text{ for all } (i, j) \in \{0, 1, 2, \dots\}^2 \setminus \{(0, 0)\},$$

$$(A3) \quad (1 - \rho_{i,j,n}) \log n \rightarrow \delta_{i,j} \in (\epsilon, \infty) \quad \text{for all } (i, j) \in \{0, 1, 2, \dots\}^2 \setminus \{(0, 0)\} \text{ and} \\ \text{some } \epsilon > 0,$$

$$(A4) \quad \lim_{n \rightarrow \infty} \sup_{\sqrt{i^2 + j^2} \geq l_n} |\rho_{i,j,n}| \log n = 0,$$

$$(A5) \quad \frac{l_n^3 (\log n)^2}{q_n n^2} \rightarrow 0,$$

(A6) *uniformly for all  $\alpha$ ,*

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{(i,j) \in H_{\alpha,m,n}} n^{-2 \frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}} \frac{(\log n)^{-\frac{\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}}}{\sqrt{1-\rho_{|i-\alpha|,j-1,n}^2}} = 0.$$

Then

$$\lim_{n \rightarrow \infty} \mathbb{P}(\max\{\xi_{i,j,n}\}_{i,j=1}^n \leq u_{n^2}) = \exp(-\vartheta \exp(-x)),$$

where

$$\vartheta = \mathbb{P}\left(E/4 + \sqrt{\frac{1}{2}} \delta_{i,j} W_{i,j} \leq \delta_{i,j}, (i,j) \in K\right),$$

$$K = \{\mathbb{N} \times \{0\}\} \cup \{\mathbb{Z} \times \mathbb{N}\},$$

$E$  is a standard exponential random variable, and  $\{W_{i,j}\}$  is a mean zero, variance one, Gaussian random field independent of  $E$  with correlation

$$\mathbb{E}(W_{i,j} W_{k,l}) = \frac{\delta_{i,j} + \delta_{k,l} - \delta_{|i-k|,|j-l|}}{2\sqrt{\delta_{i,j}\delta_{k,l}}}. \quad (5.2)$$

**Theorem 5.2.** *Let  $\rho(h)$  be an isotropic, monotonically decreasing correlation function satisfying*

- (i)  $\rho(h) = 1 - C|h|^\beta + o(|h|^\beta)$  as  $h \downarrow 0$ , for some  $0 < \beta \leq 2$ ,
- (ii)  $0 \leq \rho(h) \leq K e^{-h^a}$  for some constants  $K, a > 0$  and all  $h \geq 0$ .

Suppose  $X_n$  is a mean zero, variance one Gaussian random field on an  $n \times n$  lattice with isotropic correlation function given by

$$\rho_n(h) = \rho\left(\frac{h}{(\log n)^{1/\beta}}\right), \quad 0 < \beta \leq 2.$$

Then the conclusion of Theorem 5.1 holds.

*Proof.* Let  $l_n = (\log n)^{1/\beta} (\log \log n)^{2/a}$  and  $p_n = (\log n)^{2/\beta}$ . We proceed to verify conditions A1-A6 of Theorem 5.1.

**Verification of condition A1:**

$\rho_n(h) \geq 0$  since  $\rho(h) \geq 0$ .

**Verification of condition A2:**

First, we note that the distance  $h$  between any two random variables separated by  $i$  horizontal units and  $j$  vertical units will have the form  $h = c\sqrt{i^2 + j^2}$ , where  $c > 0$ . Denote the  $h$  corresponding to specific  $i$  and  $j$  by  $h_{i,j}$ . Because  $(i, j) \in \{0, 1, 2, \dots\}^2 \setminus \{(0, 0)\}$ , we have that  $h_{i,j} = c\sqrt{i^2 + j^2} \geq c = c\sqrt{1^2 + 0^2} = h_{1,0}$ . Since  $\rho(h)$  is monotonically decreasing, we have

$$\rho_n(h_{1,0}) = \rho_{1,0,n} \geq \rho_{i,j,n} = \rho_n(h_{i,j}), \quad \text{for all } (i, j) \in \{0, 1, 2, \dots\}^2 \setminus \{(0, 0)\}.$$

**Verification of condition A3:**

Notice that

$$1 - \rho_n(h) = 1 - \rho\left(\frac{h}{(\log n)^{1/\beta}}\right) \sim C \frac{|h|^\beta}{\log n}.$$

Thus,  $(1 - \rho_n(h)) \log n \rightarrow C|h|^\beta$ . Again noting that  $h = c\sqrt{i^2 + j^2}$  for some  $c > 0$ ,  $(i, j) \in \{0, 1, 2, \dots\}^2 \setminus \{(0, 0)\}$  implies that  $h \geq c$  and that  $C|h|^\beta \geq C|c|^\beta \in (0, \infty)$ .

**Verification of condition A4:**

Notice that

$$\begin{aligned} \sup_{h \geq l_n} \rho_n(h) \log n &= \sup_{h \geq l_n / (\log n)^{1/\beta}} \rho(h) \log n \\ &\leq \sup_{h \geq l_n / (\log n)^{1/\beta}} K e^{-h^\alpha} \log n \\ &\leq K e^{-(\log \log n)^2 + \log \log n} \\ &\rightarrow 0. \end{aligned}$$

**Verification of condition A5:**

Since  $q_n \sim p_n$ , we have

$$\frac{l_n^3 p_n^2}{q_n n^2} \sim \frac{l_n^3 p_n}{n^2} \rightarrow 0.$$

**Verification of condition A6:**

To simplify our derivations, we note that  $\rho_n(h) = \rho_{|i-\alpha|,j-1,n}$ , where  $h = c\sqrt{|i-\alpha|^2 + (j-1)^2}$ . Without loss of generality, assume  $c = 1$ .

Case 1: Assume that  $h/(\log n)^{1/\beta} > \epsilon$ . Then  $1 - \rho_n(h) \geq \epsilon^* > 0$  and we have that

$$n^{-2\frac{1-\rho_n(h)}{1+\rho_n(h)}} \frac{(\log n)^{-\frac{\rho_n(h)}{1+\rho_n(h)}}}{\sqrt{1-\rho_n^2(h)}} \leq \frac{n^{-\epsilon^*} \cdot 1}{\sqrt{\epsilon^*}}.$$

Thus, the sum in question is bounded by

$$\begin{aligned} \sum_{(i,j) \in H_{\alpha,m,n}} \frac{n^{-\epsilon^*}}{\sqrt{2\epsilon^*}} &\leq q_n^2 \frac{n^{-\epsilon^*}}{\sqrt{2\epsilon^*}} \\ &\rightarrow 0. \end{aligned}$$

Case 2: Assume that  $h/(\log n)^{1/\beta} \leq \epsilon$ . Then for  $\epsilon$  small,

$$(1 - \delta) \frac{Ch^\beta}{\log n} < 1 - \rho\left(\frac{h}{(\log n)^{1/\beta}}\right) \leq (1 + \delta) \frac{Ch^\beta}{\log n},$$

where  $\delta > 0$  is small. Without loss of generality, assume  $C = 1$ . Hence,

$$\begin{aligned} &n^{-2\frac{1-\rho_n(h)}{1+\rho_n(h)}} \frac{(\log n)^{-\frac{\rho_n(h)}{1+\rho_n(h)}}}{\sqrt{1-\rho_n^2(h)}} \\ &\leq n^{-2\frac{1-\rho_n(h)}{1+\rho_n(h)}} \frac{(\log n)^{\frac{1-\rho_n(h)}{1+\rho_n(h)}}}{\sqrt{(1+\rho_n(h))(1-\rho_n(h)) \log n}} \\ &\leq \frac{\exp\left(- (1-\delta) \frac{h^\beta}{\log n} \log n\right) \exp\left(\frac{1-\rho_n(h)}{1+\rho_n(h)} \log \log n\right)}{\sqrt{(1-\delta)h^\beta}} \\ &\leq Ke^{-(1-\delta)h^\beta} h^{-\beta/2} \exp\left((1+\delta)h^\beta \frac{\log \log n}{\log n}\right) \\ &\leq Ke^{-bh^\beta} h^{-\beta/2}, \end{aligned}$$

where  $b > 0$ . Thus, making the substitution  $h = \sqrt{|i - \alpha|^2 + (j - 1)^2}$ , the sum in question is bounded by

$$\begin{aligned}
& K \sum_{(i,j) \in H_{\alpha,m,n}} \frac{e^{-(|i-\alpha|^2+(j-1)^2)^{\beta/2} b}}{\sqrt{(|i-\alpha|^2+(j-1)^2)^{\beta/2}}} \\
& \leq K \sum_{i=m+1}^{q_n} \frac{e^{-i^{\beta} b}}{i^{\beta/2}} + 4K \sum_{i=m+1}^{q_n} \sum_{j=1}^{q_n} \frac{e^{-b(i^2+j^2)^{\beta/2}}}{\sqrt{(i^2+j^2)^{\beta/2}}} \\
& \leq K \sum_{i=m+1}^{\infty} e^{-bi^{\beta}} + 4K \left( \sum_{i=m+1}^{\infty} e^{-\frac{1}{2}i^{\beta}} \right) \left( \sum_{j=1}^{\infty} e^{-\frac{1}{2}j^{\beta}} \right), \tag{5.3}
\end{aligned}$$

since

$$\begin{aligned}
(i^2 + j^2)^{\beta/2} &= \frac{1}{2}(i^2 + j^2)^{\beta/2} + \frac{1}{2}(i^2 + j^2)^{\beta/2} \\
&\geq \frac{1}{2}i^{\beta} + \frac{1}{2}j^{\beta}.
\end{aligned}$$

Noting that (5.3) goes to zero as  $m \rightarrow \infty$ , we see that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{(i,j) \in H_{\alpha,m,n}} n^{-2\frac{1-\rho_n(h)}{1+\rho_n(h)}} \frac{(\log n)^{-\frac{\rho_n(h)}{1+\rho_n(h)}}}{\sqrt{1-\rho_n^2(h)}} = 0.$$

□

**Remark 1.** Condition (ii) of Theorem 5.2 can be relaxed, but a different choice of  $l_n$  would be needed.

**Remark 2.** The following covariance functions satisfy the conditions of Theorem 5.2:

- (i) Gaussian with  $\beta = 2$ ,
- (ii) exponential with  $\beta = 1$ ,
- (iii) spherical with  $\beta = 1$ ,
- (iv) Matérn with  $\beta = 1$  when  $\nu \in (0, 1)$  and with  $\beta = 2$  when  $\nu \geq 1$ .

## 5.5 Approximation of Maximum on a Fixed Domain

We would like to use the limiting result in Theorem 5.1 to approximate the maximum of a stationary Gaussian random field on an  $n \times n$  lattice in a fixed domain. Specifically, we would like to approximate

$$\mathbb{P}\left(\max_{1 \leq i, j \leq n} Z_{i,j} \leq u\right), \quad (5.4)$$

where  $\{Z_{i,j}\}$  is a stationary Gaussian random field on a square  $n \times n$  lattice of equally spaced locations in a fixed domain, and  $Z_{i,j}$  represents the random variable at position  $(i, j)$ . Letting  $\{Z_i\}_{i=1}^n$  represent a stationary Gaussian sequence of random variables, both Hsing et al. (1996) and Hannig and Marron (2006) suggest that instead of using the limiting Gumbel distribution to approximate  $\mathbb{P}(\max_{i=1, \dots, n} Z_i \leq u)$ , the Gaussian power distribution  $\Phi(u)^{\vartheta n}$  should instead be used. The reasoning behind this approach is that the Gaussian power distribution converges to the limiting Gumbel distribution and the empirical fact that the Gaussian power distribution often performs better than the limiting Gumbel distribution. Thus, instead of directly using the result in Theorem 5.1 to approximate the probability in (5.4), we instead wish to use the approximation  $\Phi(u)^{\vartheta n^2}$ , where

$$\vartheta = \mathbb{P}\left(E/4 + \sqrt{\frac{1}{2}}\delta_{i,j}W_{i,j} \leq \delta_{i,j}, (i, j) \in K\right), \quad (5.5)$$

and

$$K = \{\mathbb{N} \times \{0\}\} \cup \{\mathbb{Z} \times \mathbb{N}\}.$$

Except in some special cases, there are no explicit expressions for  $\vartheta$  and we must estimate this value using simulation.

### 5.5.1 Structure of Simulation

The accuracy of the proposed approximation is considered over a region  $\mathcal{D}$  of size  $[0, 20] \times [0, 20]$ . Our goal is to approximate  $\mathbb{P}(\max_{1 \leq i, j \leq n} Z_{i,j} \leq u)$  over an  $n \times n$

$s_{1,3}$	$s_{2,3}$	$s_{3,3}$
$s_{1,2}$	$s_{2,2}$	$s_{3,2}$
$s_{1,1}$	$s_{2,1}$	$s_{3,1}$

Figure 5.1: The relationship between  $s_{i,j}$  and a  $3 \times 3$  grid of squares in  $\mathcal{D}$ . The random variable  $Z_{i,j}$  is located at the center point of the square in position  $i$  of the  $x$ -direction and position  $j$  of the  $y$ -direction.

lattice contained in  $\mathcal{D}$ . Dividing  $\mathcal{D}$  into an  $n \times n$  grid of squares, the random variable  $Z_{i,j}$  is located at the center point of the square in position  $i$  of the  $x$ -direction and position  $j$  of the  $y$ -direction. This relationship is illustrated in Figure 5.1 for a for a  $3 \times 3$  grid of locations.

Suppose that the random field  $\{Z_{i,j}\}$  has isotropic covariance function  $\rho(\cdot)$  satisfying the conditions of Theorem 5.2. In order to come up with a standard with which to compare our approximation, we empirically simulated  $\mathbb{P}(\max_{1 \leq i,j \leq n} Z_{i,j} \leq u)$  using 1000 realizations of a random field  $\{Z_{i,j}\}$  having covariance function  $\rho(\cdot)$ . The estimate of  $\mathbb{P}(\max_{1 \leq i,j \leq n} Z_{i,j} \leq u)$  is then given by the proportion of realizations for which  $\max_{1 \leq i,j \leq n} \{Z_{i,j}\} \leq u$ . This quantity will be referred to as the empirical probability. We approximate the probability in (5.4) using the Gaussian power distribution  $\Phi(u)^{\tilde{\vartheta}n^2}$ , where  $\tilde{\vartheta}$  is the estimate of  $\vartheta$  obtained via simulation. Details of the simulation used to estimate  $\tilde{\vartheta}$  are given below.

### 5.5.2 Estimating $\tilde{\vartheta}$

We would like to approximate  $\vartheta$  by

$$\tilde{\vartheta} = \mathbb{P}\left(E/4 + \sqrt{\frac{1}{2}\delta_{i,j}}W_{i,j} \leq \delta_{i,j}, (i,j) \in \tilde{K}\right) \quad (5.6)$$

where

$$\tilde{K} = \{\{1, \dots, q\} \times \{0\}\} \cup \{\{-q, \dots, -2, -1, 0, 1, 2, \dots, q\} \times \{1, \dots, q\}\}.$$

Because we cannot estimate  $\vartheta$  using an infinite set  $K$  as defined in (5.5), we substitute a finite set  $\tilde{K}$  in its place. We first generate a realization of  $\{W_{i,j}, (i,j) \in \tilde{K}\}$ , where each  $W_{i,j} \sim N(0, 1)$  and the correlation between random variables is given in (5.2). Next, a realization of a standard exponential random variable  $E$  is generated independently of the  $\{W_{i,j}\}$ . Lastly, the quantity  $E/4 + \sqrt{\frac{1}{2}\delta_{i,j}}W_{i,j}$  is compared to  $\delta_{i,j}$  for all  $(i,j) \in \tilde{K}$ . The estimated value of  $\tilde{\vartheta}$  is given by the proportion of realizations for which the event  $\{E/4 + \sqrt{\frac{1}{2}\delta_{i,j}}W_{i,j} \leq \delta_{i,j}, (i,j) \in \tilde{K}\}$  occurred.

Due to time and resource considerations, the value of  $q$  was chosen to be either 15 or 25. The estimation of  $\tilde{\vartheta}$  was computationally expensive for large values of  $q$  and the speed of the simulation also depended on the covariance function  $\rho(\cdot)$ . The value of  $q$  was chosen accordingly. More information about this is given in Section 5.5.3.

### 5.5.3 Covariance Functions and Other Details

Three isotropic correlation functions were considered in the simulations: Gaussian, spherical, and exponential. Letting  $h$  denote the distance between the two random variables under consideration, the Gaussian correlation function is given by

$$\rho(h) = \exp(-h^2/\phi^2), \quad h \geq 0, \quad (5.7)$$

the spherical correlation function is given by

$$\rho(h) = \left(1 - \frac{3h}{2\phi} + \frac{1}{2}\left(\frac{h}{\phi}\right)^3\right) \mathbf{1}_{[0,\phi)}(h), \quad h \geq 0, \quad (5.8)$$

and the exponential correlation function by

$$\rho(h) = \exp(-h/\phi), \quad h \geq 0. \quad (5.9)$$

For each covariance function we looked at estimating  $\mathbb{P}(\max_{1 \leq i, j \leq n} Z_{i,j} \leq u)$  for  $n = 12, 36, 108$ , and  $324$ . These values of  $n$  were chosen because the locations of  $\{Z_{i,j}\}$  for the  $12 \times 12$  lattice of locations is a subset of the locations of the  $36 \times 36$  lattice, the locations of the  $36 \times 36$  lattice are a subset of the locations of the  $108 \times 108$  lattice, and so on. Thus, observations from a realization of  $\{Z_{i,j}\}$  over the  $324 \times 324$  lattice could be used as the observations for lattices of size  $108, 36$ , and  $12$ . Consequently, this reduced the amount of time needed to complete each simulation.

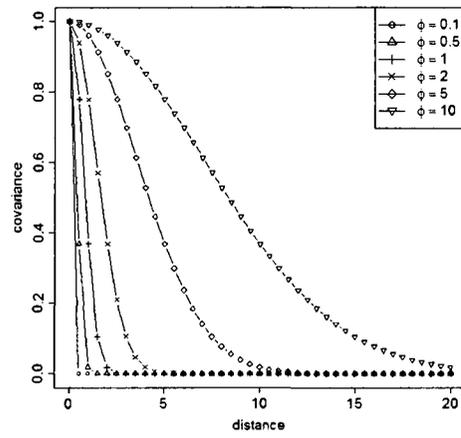
In order to see how the accuracy of our approximation changed with the degree of dependence among  $\{Z_{i,j}\}$ , we looked at the accuracy of our approximations for  $\phi = 0.1, 0.5, 1, 2, 5$ , and  $10$  (except for the spherical covariance function where  $\phi = 0.1$  was not considered due to time constraints). In order to get a sense of the level of dependence for the differing correlation functions, we have provided comparisons in Figures 5.2 (a)-(c). The Gaussian correlation function tends to exhibit the strongest local dependence and the spherical correlation function the least.

Similar to the motivating example in Section 5.2, we imbed our fixed domain random field into a triangular array of random fields. Following the example in Hannig and Marron (2006), a natural way of doing this is to assume that  $c = t/(\log n)^{1/\beta}$ , where  $c$  is the horizontal/vertical spacing of the lattice locations for the random variables of the fixed domain random field,  $t > 0$  is some constant, and  $(\log n)^{1/\beta}$  is a scaling sequence determined by the covariance function.

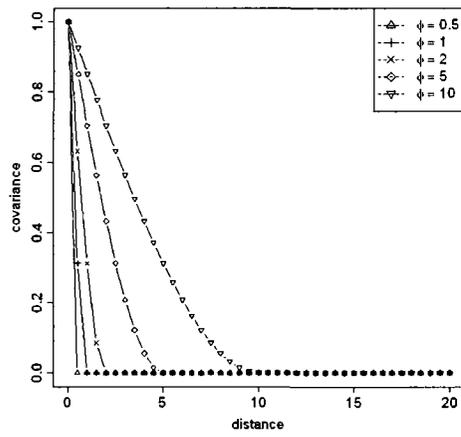
The last detail of the simulation procedure which should be discussed is the size of  $\tilde{K}$  in our simulations. As previously noted

$$\tilde{K} = \{\{1, \dots, q\} \times \{0\}\} \cup \{\{-q, \dots, -2, -1, 0, 1, 2, \dots, q\} \times \{1, \dots, q\}\}.$$

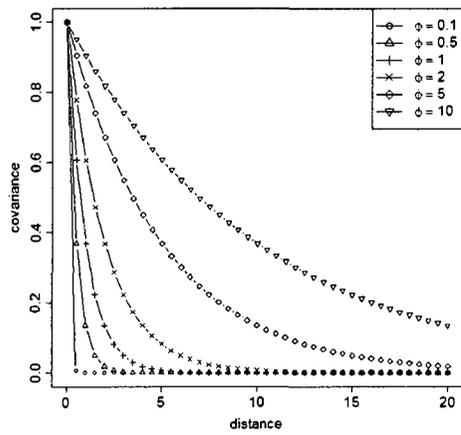
The value of  $q$  for simulations having Gaussian or exponential covariance functions was chosen to be 25. The value of  $q$  for simulations having spherical covariance functions



(a) Gaussian covariance function.



(b) Spherical covariance function.



(c) Exponential covariance function.

Figure 5.2: A comparison of spatial dependence for various correlation functions.

was 15. The reason for this is that simulations for the spherical covariance function took longer to run, and the value of  $q$  was lowered to reduce the time necessary to complete the simulations.

#### 5.5.4 Summary of results

Graphical results of the simulations are provided in Appendix A at the end of this dissertation for select combinations of the covariance functions and the range parameter  $\phi$ . In these figures the empirical probability described in Section 5.5.1 is shown as a solid line and labeled “empirical”. The approximation is shown as a dashed line and labeled “approximation”. To more easily assess the accuracy of the approximation in the upper tail of the distribution, we have provided lines running horizontally at probabilities of 0.90 and 0.95.

#### Summary of results for Gaussian covariance function

The results for simulations involving the Gaussian covariance function are shown at the beginning of Appendix A.

The approximation works best when  $n$  and  $\phi$  are both small. When both  $n$  and  $\phi$  are small the approximation appears to be a fairly accurate estimate of the empirical probability for all  $u$ . However, as  $\phi$  increases, the accuracy of the approximation begins to decrease for larger  $n$ . Once  $\phi = 2, 5,$  and  $10$ , the approximation for  $n = 324$  or  $n = 108$  does not become accurate until far in the upper tail. The approximation performs poorly when both  $n$  and  $\phi$  are large. For smaller  $n$ , the approximation does moderately well in the upper tail when  $\phi$  is close to one or two, but begins to lose accuracy as  $\phi$  increases. Regardless of the value of  $n$  or  $\phi$ , the approximation tends to be a conservative approximation of the empirical probability.

#### Summary of results for spherical covariance function

The results for the spherical covariance function are shown after the results for the Gaussian covariance function in Appendix A. The results of these approximations

are more mixed than the results for the Gaussian covariance function. Looking at the approximation across all  $u$ , the approximation tends to perform better when  $n$  is small. The approximation is liberal for small  $\phi$  and large  $n$ , but tends to get better as  $\phi$  increases. The best approximation of the tail probability occurs when both  $\phi$  and  $n$  are large. On the other hand, the approximation seems to fluctuate between liberal and conservative for larger values of  $\phi$  and smaller  $n$ .

### Summary of results for exponential covariance function

Similar to the results of the Gaussian covariance function, the approximation for the exponential covariance function is most accurate when  $n$  and  $\phi$  are both small. However, when  $n$  is small, even for larger values of  $\phi$  the approximation is a relatively good estimate of the empirical tail probability. For larger  $n$  and  $\phi$ , the asymptotic approximation is conservative and does not accurately estimate the empirical probability until far into the upper tail.

#### 5.5.5 A Brief Comparison with Adler (2008)

Adler (2008) proposes a method of approximating the excursion probability of a wide range of random processes. One of the specific results he provides is an approximation to

$$\mathbb{P}\left(\sup_{t \in T} Z(t) \geq u\right), \quad (5.10)$$

where  $Z$  is a centered Gaussian random process satisfying certain assumptions over a rectangular parameter space  $T$  which is bounded in  $\mathbb{R}^N$ . The approximation we have proposed in this chapter can be used to approximate

$$\mathbb{P}\left(\max_{1 \leq i, j \leq n} Z_{i,j} > u\right), \quad (5.11)$$

where  $\{Z_{i,j}\}$  is a Gaussian random field on an  $n \times n$  lattice in a bounded region of interest  $\mathcal{D} \subset \mathbb{R}^2$ . When  $n$  is large and the domain  $\mathcal{D}$  is equivalent to the parameter space  $T$ , it seems reasonable to use the estimate of the excursion probability in (5.10)

proposed by Adler as an approximation of (5.11). Thus, we will compare the approximation of (5.11) proposed in this chapter with the estimate of (5.10) proposed by Adler when  $n$  is large and taking  $T$  to be the same region as  $\mathcal{D}$ . As mentioned in Section 3.4, Adler's result requires the random field to be twice-differentiable. Of the three covariance functions discussed in this chapter, only the Gaussian covariance gives processes meeting the required conditions. Thus, we will only compare the approximations for the Gaussian covariance function. Additional details about Adler's approximation are provided in provided Section 3.4, as well as the explicit expression used to approximate the excursion probability when  $T$  is a rectangle in  $\mathbb{R}^2$  and the covariance function of  $Z$  is isotropic and Gaussian.

For our comparison, we let the domain  $\mathcal{D}$  be a  $[0, 20] \times [0, 20]$  region and assumed that the lattice of locations in  $\mathcal{D}$  was of size  $324 \times 324$ . In order to have a standard of comparison for the two approximations, we simulated 1000 realizations of  $\{Z_{i,j}\}_{i,j=1}^{324}$  and estimated the excursion probability empirically. Following the same details specified in Sections 5.5.1-5.5.3, we proceeded to estimate (5.11) using the approximation proposed in this chapter, and then estimated the same probability using Adler's method. The results are shown in Figures 5.3 (a)-(d) for four different values of the range parameter  $\phi$ . Horizontal lines are provided at vertical positions of 0.05 and 0.10 for reference.

Except for situations where the range parameter is small (e.g.,  $\phi = 0.1$  or 0.5), the approximation proposed by Adler is closer to the empirical estimate of  $\mathbb{P}(\max_{1 \leq i,j \leq n} Z(s_{i,j}) > u)$  than the approximation proposed in this chapter. Unless the range parameter  $\phi$  is small, when  $n$  is large, it is better to use Adler's approximation of (5.10) to estimate the excursion probability in (5.11) for the Gaussian covariance function. Another benefit of Adler's method is that the approximation is a simple calculation which can be computed instantly; one does not need to perform any numerical or simulation procedures to obtain the estimate.

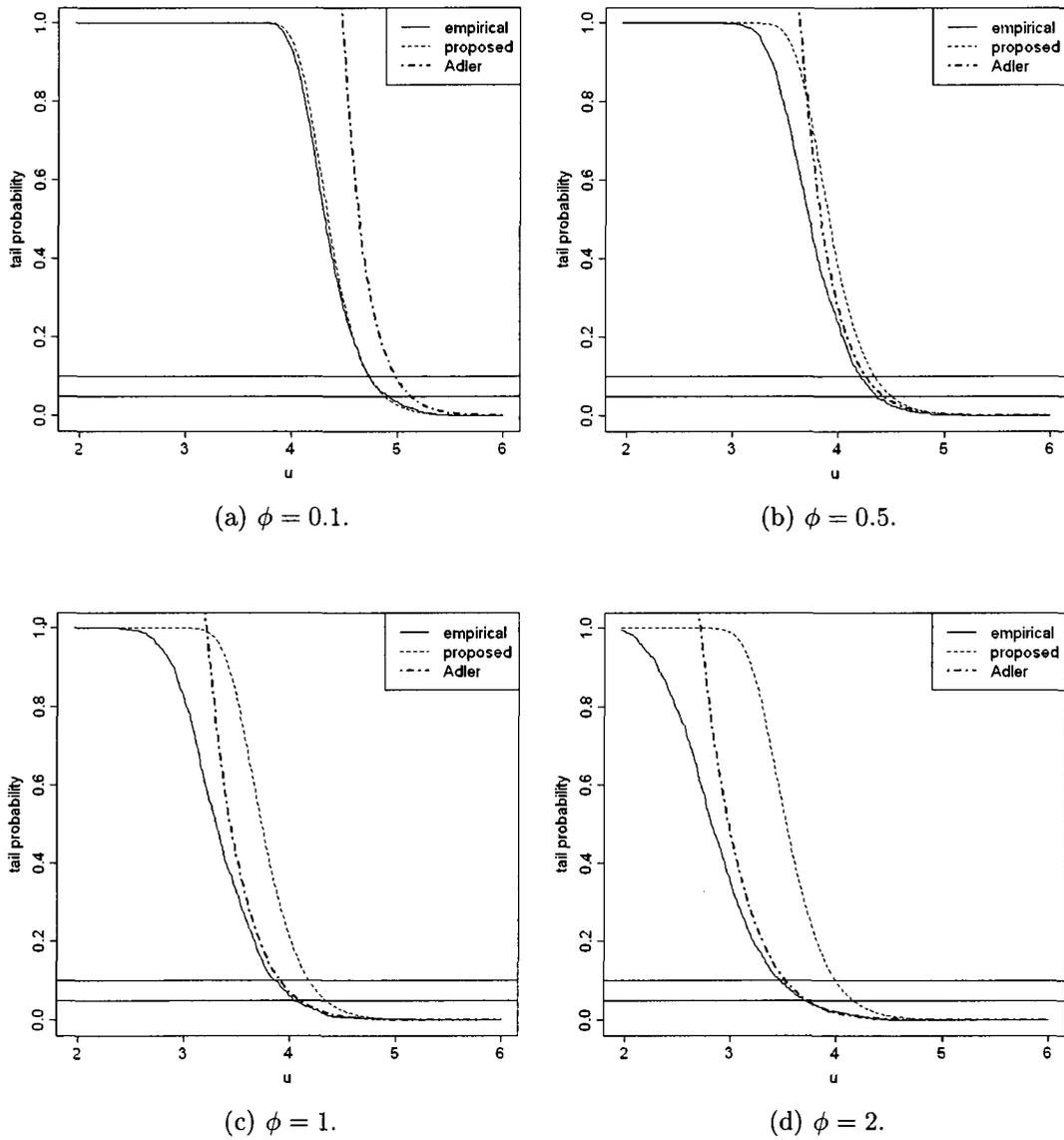


Figure 5.3: A comparison of estimated excursion probabilities using two methods. The empirical excursion probability is shown as a solid line and serves as the reference probability. The dash-dotted line represents the approximation using Adler's method. The dashed line represents the approximation using the method proposed in this chapter.

### 5.5.6 Overall results

In general, the proposed asymptotic approximation of

$$\mathbb{P}\left(\max_{1 \leq i, j \leq n} Z(s_{i,j}) \leq u\right) \quad (5.12)$$

works best when  $n$  and  $\phi$  are small. The approximation also appears to be useful in estimating the upper tail of  $\mathbb{P}(\max_{1 \leq i, j \leq n} Z(s_{i,j}) \leq u)$  when  $n$  is large and  $\phi$  is moderately large for both the spherical and exponential covariance functions. The approximation does not accurately estimate the upper tail for the Gaussian covariance function when both  $n$  and  $\phi$  are large. When  $n$  is large (and the lattice is very dense within the domain  $\mathcal{D}$ ), the simulation results discussed in Section 5.5.5 indicate that a better estimate can be obtained using the approximation in Adler (2008). Though the result presented by Adler is actually estimating

$$\mathbb{P}\left(\sup_{s \in \mathcal{D}} Z(s) \leq u\right), \quad (5.13)$$

it can be expected that for smooth random fields and relatively dense lattices in  $\mathcal{D}$  that (5.12) and (5.13) will have similar values.

In general, the accuracy of the asymptotic approximation seems to decrease as the range parameter  $\phi$  increases. This is consistent with the results of a simulation study performed by Wilhem (2002). In this study, Wilhelm compared the asymptotic estimator proposed in Hsing et al. (1996) (having the form  $\Phi(u)^{n^{\bar{\nu}}}$ ) to two other estimators of the maxima of a stationary sequence of Gaussian random variables. None of the approaches gave reliable answers for highly dependent stationary series. Since our proposed asymptotic approximation  $\Phi(u)^{n^{2\bar{\nu}}}$  is very similar to the estimator proposed by Hsing et al., it is not surprising that our approximation is also less reliable as the dependence in the random field increases.

The proposed approximation tends to work best in the upper tail when  $u$  is large. This is consistent with the limiting result in Theorem 5.1 since we are considering the

limiting distribution of  $\mathbb{P}(M_n \leq u_{n^2})$ , where  $u_{n^2} \rightarrow \infty$  as  $n \rightarrow \infty$ . It seems reasonable then that the asymptotic approximation also would perform better for larger values of  $u$ .

Conceptually, it seems that the asymptotic approximation proposed in this chapter should work better for large values of  $n$ . To the contrary, it was somewhat surprising that the asymptotic approximation typically performed better for smaller  $n$  than for densely spaced lattices where  $n$  was large. The explanation for this apparent inconsistency appears to be the difficulty in accurately approximating  $\tilde{\vartheta}$  when  $n$  is large. The accuracy of  $\tilde{\vartheta}$  is more important when  $n$  is large since the value of  $\tilde{\vartheta}$  has a greater impact on the value of the approximation  $\Phi(u)^{n^2\tilde{\vartheta}}$  than when  $n$  is small. The approximation of  $\tilde{\vartheta}$  depends on

$$\tilde{K} = \{\{1, \dots, q\} \times \{0\}\} \cup \{\{-q, \dots, -2, -1, 0, 1, 2, \dots, q\} \times \{1, \dots, q\}\},$$

where  $q \in \mathbb{N}$ . Computational limitations did not allow  $q$  to get much bigger than 30 in our simulations, which does not appear to be large enough to consistently obtain a good approximation of  $\tilde{\vartheta}$  when  $n$  is large. As the size of  $\tilde{K}$  increases (i.e., as  $q$  becomes larger), the value of  $\tilde{\vartheta}$  will monotonically decrease. We previously noted that the asymptotic approximation tends to be conservative for the Gaussian and exponential covariance functions when  $n$  is large. The approximation will become less conservative when  $\tilde{\vartheta}$  is smaller. Because  $\tilde{\vartheta}$  should decrease as the size of  $\tilde{K}$  increases, we can expect that the approximations would become more accurate for larger  $n$  if the simulation procedure for estimating  $\tilde{\vartheta}$  was performed using larger sizes of  $\tilde{K}$ .

## 5.6 Proof of Theorem 5.1 and Complements

### 5.6.1 Preliminaries

It is not possible to directly follow the argumentation of Hsing et al. (1996) when trying to extend their result to two dimensions. The main reason is that their result

depends on the work of O'Brien (1987), which is not applicable in our situation. To extend the result of Hsing et al., we will first need to create some of our own tools and then use them to get the desired extension.

As in Theorem 5.1, we define sequences  $p_n = o(n)$ ,  $l_n = o(p_n)$ , and  $q_n = p_n + l_n$ . We also define  $r_n = n(p_n + l_n)^{-1} = nq_n^{-1}$ . These definitions will stay consistent throughout the sequel.

For fixed  $n$ , we have previously defined  $M_n = \max\{\xi_{i,j,n}\}_{i,j=1}^n$ , the maximum of a stationary Gaussian random field on an  $n \times n$  lattice. For simplicity, we define  $M_{q_n}$  to be the maximum of a  $q_n \times q_n$  subset of this random field, i.e.,  $M_{q_n} = \max\{\xi_{i,j,n}\}_{i,j=1}^{q_n}$ . At times, we will need to consider the maximum of a subset of the random variables of  $X_n$ . Thus, we define

$$M_{(i,j):(k,l)}^a = \max\{\xi_{i,j,n}, \dots, \xi_{a,j,n}, \xi_{1,j+1,n}, \dots, \xi_{a,j+1,n}, \dots, \xi_{1,l,n}, \dots, \xi_{k,l,n}\}.$$

To describe  $M_{(i,j):(k,l)}^a$  in more detail, suppose we have a lattice of random variables with  $a$  columns (the number of rows is not required to equal  $a$ ) having the same covariance structure as  $X_n$ .  $M_{(i,j):(k,l)}^a$  represents the maximum of the following set of random variables:

- All random variables with vertical position  $j$  which have horizontal position greater than or equal to  $i$  and less than or equal to  $a$ . i.e., the set of variables  $\{\xi_{i,j,n}, \xi_{i+1,j,n}, \dots, \xi_{a,j,n}\}$ .
- All random variables with vertical positions  $i + 1, i + 2, \dots, l - 1$ .
- All random variables with vertical position  $l$  which have a horizontal position between 1 and  $k$ , inclusive. i.e., the set of variables  $\{\xi_{1,l,n}, \xi_{2,l,n}, \dots, \xi_{k,l,n}\}$ .

A visual representation of the set  $M_{(3,1):(2,6)}^5$  is shown in Figure 5.4. Also, it will be useful to define the following equivalence relationship. For  $j < l$  define

$$M_{(a+1,j):(k,l)}^a = M_{(1,j+1):(k,l)}^a. \quad (5.14)$$

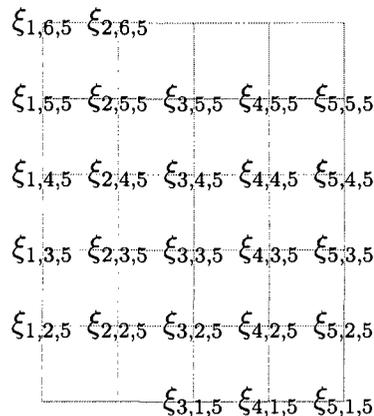


Figure 5.4: A visual representation of  $M_{(3,1):(2,6)}^5$ .

To complete this section of introductory information, we conclude by presenting Corollary 2.2 of Li and Shao (2002), which is used extensively in this chapter.

**Corollary 5.3** (Corollary 2.2 of Li and Shao (2002)). *Let  $n \geq 3$  and let  $(\xi_j, 1 \leq j \leq n)$  be standard normal random variables with covariance matrix  $R = (r_{ij})$ . Assume that  $r_{ij} \geq 0$ . Then*

$$\begin{aligned}
& \mathbb{P}\left(\bigcap_{j=1}^m \{\xi_j \leq u\}\right) \mathbb{P}\left(\bigcap_{m < j \leq n} \{\xi_j \leq u\}\right) \\
& \leq \mathbb{P}\left(\bigcap_{j=1}^n \{\xi_j \leq u\}\right) \\
& \leq \mathbb{P}\left(\bigcap_{j=1}^m \{\xi_j \leq u\}\right) \mathbb{P}\left(\bigcap_{m < j \leq n} \{\xi_j \leq u\}\right) \\
& \quad \times \exp\left\{\sum_{i=1}^m \sum_{j=m+1}^n \log\left(\frac{\pi}{\pi - 2 \arcsin(r_{ij})}\right) \exp(-u^2/r_{ij})\right\}
\end{aligned}$$

for  $1 \leq m \leq n - 1$  and  $u \geq 0$ .

### 5.6.2 Bounding $\mathbb{P}(M_n \leq u_{n^2})$

In this section, we develop bounds on  $\mathbb{P}(M_n \leq u_{n^2})$  that will later be used in the proof of Theorem 5.1.

**Lemma 5.4.** *Assume condition A1 of Theorem 5.1. Then*

$$\mathbb{P}(M_{q_n} \leq u_{n^2})^{r_n^2} \leq \mathbb{P}(M_n \leq u_{n^2}).$$

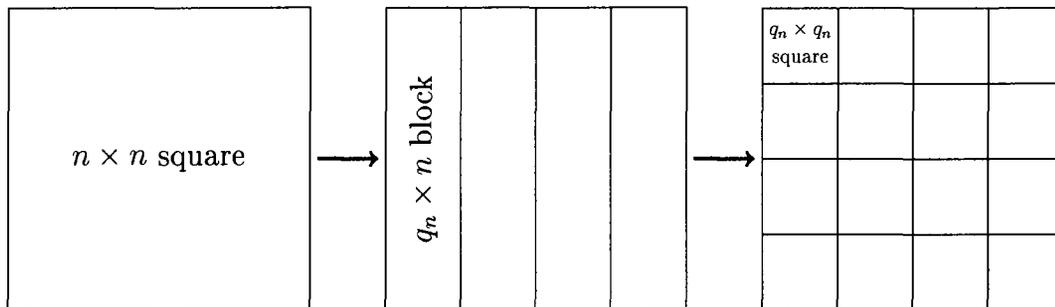


Figure 5.5: A visual approach to Lemma 5.4. The original set of  $n^2$  random variables on a lattice of size  $n \times n$  is first broken into  $n/q_n$  blocks of size  $q_n \times n$ , and then each of these blocks is broken into squares of size  $q_n \times q_n$ .

*Proof.* To complete this proof, we will use the left-hand side of Corollary 5.3. The initial set of  $n^2$  random variables is first broken into  $n/q_n = r_n$  blocks of random variables of size  $q_n \times n$ . Each of these blocks is then further broken into  $r_n$  squares of random variables of size  $q_n \times q_n$ . This process is illustrated graphically in Figure 5.5.

Through direct application of the left-hand side of Corollary 5.3, we have that

$$\mathbb{P}(M_n \leq u_{n^2}) \geq \mathbb{P}(M_{(1,1):(q_n,n)}^{q_n} \leq u_{n^2})^{r_n}.$$

Similarly, we see that

$$\mathbb{P}(M_{(1,1):(q_n,n)}^{q_n} \leq u_{n^2}) \geq \mathbb{P}(M_{q_n} \leq u_{n^2})^{r_n}.$$

Combining these two facts, it follows that

$$\mathbb{P}(M_n \leq u_{n^2}) \geq \mathbb{P}(M_{q_n} \leq u_{n^2})^{r_n^2}.$$

□

**Lemma 5.5.** *Assume conditions A1-A5 of Theorem 5.1. Then for large enough  $n$ ,*

$$\mathbb{P}(M_n \leq u_{n^2}) \leq \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} (1 + o(1)).$$

*Proof.* We will prove this upperbound for  $\mathbb{P}(M_n \leq u)$  by breaking our  $n \times n$  lattice of random variables into blocks of random variables of size  $q_n \times q_n$ , and using the

right-hand side of Corollary 5.3. To complete this proof, we will introduce notation that will only be used in the proof of this particular lemma. Define

$$M_n^{-(1,1):(a,b)} = \bigvee_{j=1}^b M_{(a+1,j):(n,j)}^n \vee \bigvee_{j=a+1}^n M_{(1,j):(n,j)}^n.$$

More descriptively,  $M_n^{-(1,1):(a,b)}$  is the maximum over all random variables in the  $n \times n$  lattice of random variables, excluding a section of size  $a \times b$  in the lower left-hand corner. Noting that  $\log(\pi/(\pi - 2 \arcsin(r))) \leq \log(1/(1 - r))$  for  $r > 0$ , from the right-hand side of Corollary 5.3 we see that

$$\begin{aligned} & \mathbb{P}(M_n \leq u_{n^2}) \\ & \leq \mathbb{P}(M_{q_n} \leq u_{n^2}) \mathbb{P}(M_n^{-(1,1):(q_n,q_n)} \leq u_{n^2}) \\ & \quad \times \exp \left\{ \sum_{i=1}^{q_n} \sum_{j=1}^{q_n} \sum_{k=1}^{q_n} \sum_{l=q_n+1}^n \log \left( \frac{1}{1 - \rho_{|i-k|,|j-l|,n}} \right) \exp \left( -\frac{u_{n^2}^2}{\rho_{|i-k|,|j-l|,n}} \right) \right. \\ & \quad \left. + \sum_{i=1}^{q_n} \sum_{j=1}^{q_n} \sum_{k=q_n+1}^n \sum_{l=1}^{q_n} \log \left( \frac{1}{1 - \rho_{|i-k|,|j-l|,n}} \right) \exp \left( -\frac{u_{n^2}^2}{\rho_{|i-k|,|j-l|,n}} \right) \right\}. \end{aligned} \quad (5.15)$$

Let  $\rho^* = \sup_{1 \leq i,j < l_n} \rho_{i,j,n}$  and  $\rho' = \sup_{\max(i,j) \geq l_n} \rho_{i,j,n}$ .

We will split the sums in (5.15) into those sets where  $1 \leq |i - k|, |j - l| < l_n$  and those where the  $\max(|i - k|, |j - l|) \geq l_n$ . Consider the  $q_n \times q_n$  block of random variables associated with  $M_{q_n}$ . Denote this block by  $A$ . Denote the random variables associated with  $M_n^{-(1,1):(q_n,q_n)}$  by  $B$ . Notice that there will be less than  $2l_n q_n$  random variables in  $A$  that are within a distance of  $l_n$  units of the random variables in  $B$ . For each of these variables in  $A$ , there are at most  $3l_n^2$  random variables in  $B$  that will be within a distance of  $l_n$  units. Thus,  $6l_n^3 q_n$  is an upper bound for  $\#\{\xi_{i,j,n} \in A, \xi_{k,l,n} \in B : |i - k|, |j - l| < l_n\}$ . This is illustrated in Figure 5.6. Conversely, there will be less than  $n^2 - q_n^2$  random variables in  $B$  that will be at least  $l_n$  units away from the

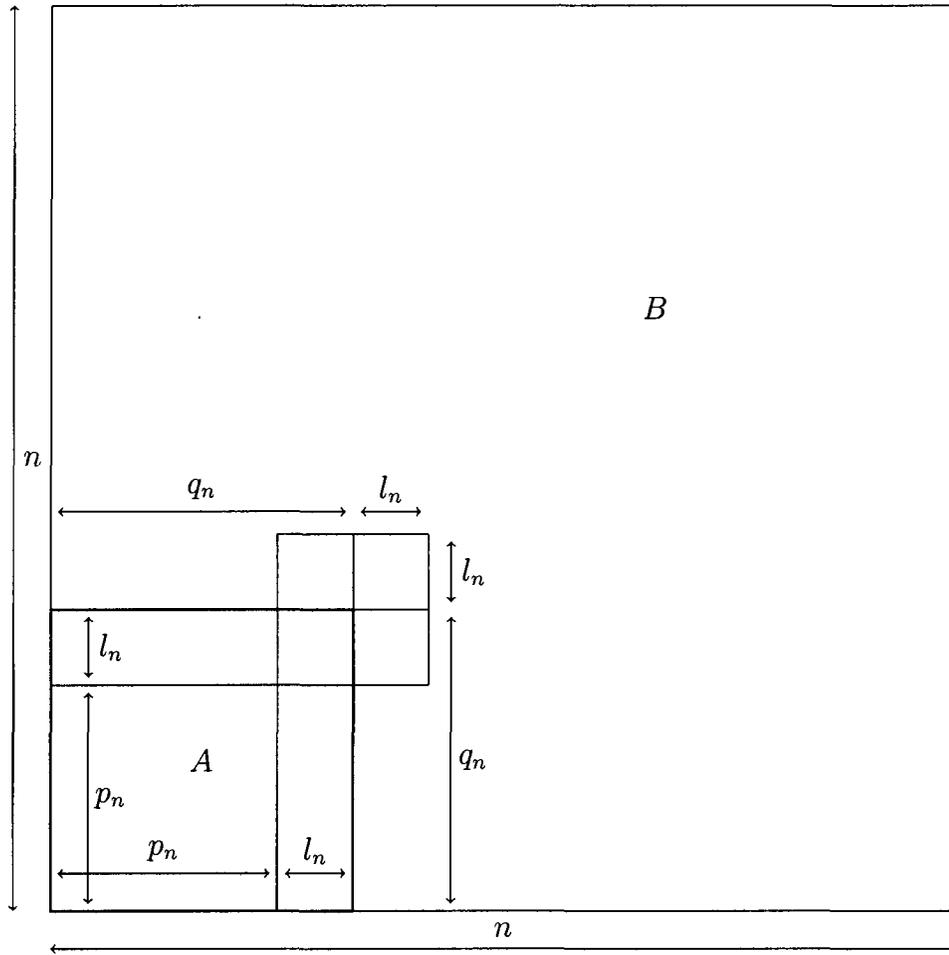


Figure 5.6: The number of variables separated by  $l_n$  units in  $A$  and  $B$ . Region  $A$  is shown by the shaded region. There are less than  $2l_n q_n$  random variables in  $A$  that are within  $l_n$  units of  $B$ . For each of these variables in  $A$ , there are no more than  $3l_n^2$  random variables in  $B$  that are within  $l_n$  units.

$q_n^2$  random variables in  $A$ . Applying these facts to (5.15), we see that

$$\begin{aligned}
& \mathbb{P}(M_n \leq u_{n^2}) \\
& \leq \mathbb{P}(M_{q_n} \leq u_{n^2}) \mathbb{P}(M_n^{-(1,1):(q_n, q_n)} \leq u_{n^2}) \exp \left\{ \sum_{|i-k|, |j-l| < l_n} \log \left( \frac{1}{1-\rho^*} \right) \exp(-u_{n^2}^2) + \right. \\
& \quad \left. + \sum_{\max(|i-k|, |j-l|) \geq l_n} \log \left( \frac{1}{1-\rho'} \right) \exp \left( -\frac{u_{n^2}^2}{\rho'} \right) \right\} \\
& \leq \mathbb{P}(M_{q_n} \leq u) \mathbb{P}(M_n^{-(1,1):(q_n, q_n)} \leq u_{n^2}) \\
& \quad \times \exp \left\{ 6l_n^3 q_n \log \left( \frac{1}{1-\rho^*} \right) \exp(-u_{n^2}^2) + q_n^2 n^2 \log \left( \frac{1}{1-\rho'} \right) \exp \left( -\frac{u_{n^2}^2}{\rho'} \right) \right\}. \quad (5.16)
\end{aligned}$$

We will continue by breaking off blocks of size  $q_n \times q_n$  from the original  $n \times n$  matrix of random variables and applying Corollary 5.3. Following this pattern and referring back to (5.16), we have that

$$\begin{aligned}
& \mathbb{P}(M_n \leq u_{n^2}) \\
& \leq \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} \exp \left\{ 6r_n^2 l_n^3 q_n \log \left( \frac{1}{1-\rho^*} \right) \exp(-u_{n^2}^2) \right. \\
& \quad \left. + r_n^2 q_n^2 n^2 \log \left( \frac{1}{1-\rho'} \right) \exp \left( -\frac{u_{n^2}^2}{\rho'} \right) \right\} \\
& = \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} \exp \left\{ 6r_n^2 l_n^3 q_n \log \left( \frac{1}{1-\rho^*} \right) \exp(-u_{n^2}^2) \right. \\
& \quad \left. + n^4 \log \left( \frac{1}{1-\rho'} \right) \exp \left( -\frac{u_{n^2}^2}{\rho'} \right) \right\}. \quad (5.17)
\end{aligned}$$

Remembering that  $\rho' = \sup_{\max(i,j) \geq l_n} \rho_{i,j,n}$ , condition A4 gives us that  $\rho' = o(1/\log n)$ , and hence

$$\log \left( \frac{1}{1-\rho'} \right) \rightarrow 0. \quad (5.18)$$

We can also see that

$$\begin{aligned}
\exp \left( -\frac{u_{n^2}^2}{\rho'} \right) n^4 &= \exp \left( -\frac{u_{n^2}^2 - 4\rho' \log n}{\rho'} \right) \\
&= \exp \left( -\frac{u_{n^2}^2 - o(1)}{\rho'} \right) \\
&\leq \exp(-u_{n^2}^2 - o(1)) \\
&\rightarrow 0. \quad (5.19)
\end{aligned}$$

Next, we notice that

$$\begin{aligned}\exp(-u_{n^2}^2) &= \exp(-4 \log n + \log(2 \log n) + O(1)) \\ &\leq K \frac{\log n}{n^4},\end{aligned}\tag{5.20}$$

for some constant  $K$ . Using (5.20), we see that

$$\begin{aligned}r_n^2 l_n^3 q_n \log\left(\frac{1}{1-\rho^*}\right) \exp(-u_{n^2}^2) &\leq K r_n^2 l_n^3 q_n \log\left(\frac{1}{1-\rho^*}\right) \frac{\log n}{n^4} \\ &= K \frac{l_n^3 \log n}{q_n n^2} \log\left(\frac{1}{1-\rho^*}\right),\end{aligned}\tag{5.21}$$

since  $r_n = n/q_n$ . By condition A2, we know that  $\rho^* = \sup_{1 \leq i, j < l_n} \rho_{i, j, n} = \rho_{1, 0, n}$ . Thus, from condition A3, we have that for large enough  $n$  and some  $\epsilon > 0$ ,

$$\begin{aligned}(1 - \rho^*) \log n &= (1 - \rho_{1, 0, n}) \log n \\ &\geq \epsilon/2.\end{aligned}$$

Hence, for large enough  $n$ ,  $(1 - \rho^*)^{-1} \leq 2 \log n / \epsilon$  and

$$\log\left(\frac{1}{1-\rho^*}\right) \leq \log n.\tag{5.22}$$

Combining (5.21) and (5.22), we see that

$$\begin{aligned}r_n^2 l_n^3 q_n \log\left(\frac{1}{1-\rho^*}\right) \exp(-u_{n^2}^2) &\leq K \frac{l_n^3 (\log n)^2}{q_n n^2} \\ &\rightarrow 0,\end{aligned}\tag{5.23}$$

by condition A5. Using (5.18), (5.19), and (5.23), we see that

$$\begin{aligned}\exp\left\{4r_n^2 l_n^3 q_n \log\left(\frac{1}{1-\rho^*}\right) \exp(-u_{n^2}^2) + n^4 \log\left(\frac{1}{1-\rho'}\right) \exp\left(-\frac{u_{n^2}^2}{\rho'}\right)\right\} \\ = \exp\{o(1)\} \\ = 1 + o(1).\end{aligned}\tag{5.24}$$

Lastly, combining (5.17) and (5.24) gives us that

$$\mathbb{P}(M_n \leq u_{n^2}) \leq \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} (1 + o(1)).$$

□

**Lemma 5.6.** *Suppose  $1 \leq b \leq a$  and define  $A = \{1 \dots, a\} \times \{1, \dots, b\} \setminus \{(a, b)\}$ . Then*

$$\mathbb{P}(M_{(1,1):(a,b)}^a > u) = \mathbb{P}(\xi_{a,b,n} > u) + \sum_{(i,j) \in A} \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,b)}^a \leq u)). \quad (5.25)$$

Furthermore, if  $b < a$ , then

$$\mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) = \sum_{i=1}^a \sum_{j=1}^b \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,a)}^a \leq u)). \quad (5.26)$$

*Proof.* The main trick used in this proof is to intersect the initial probability with the event  $\{(\xi_{i,j,n} > u) \cup (\xi_{i,j,n} \leq u)\}$  for the appropriate  $(i, j)$ , and then to break up the probability by summing over the disjoint events. We will start with position  $(i, j) = (a, b)$ , and then continue to decrement the  $i$  position until we reach the first position of the row. We will then continue by decrementing the  $j$  position by 1 (to get  $b - 1$ ) and then continuing this pattern starting in position  $(a, b - 1)$ . We will first prove the latter assertion in the case when  $b < a$ . We have that

$$\begin{aligned} & \mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) \\ &= \mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u) \cap (\xi_{a,b,n} > u)) \\ & \quad + \mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u) \cap (\xi_{a,b,n} \leq u)) \\ &= \mathbb{P}((\xi_{a,b,n} > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) \\ & \quad + \mathbb{P}((M_{(1,1):(a-1,b)}^a > u) \cap (M_{(a,b):(a,a)}^a \leq u)). \end{aligned} \quad (5.27)$$

To clearly see the unfolding pattern, we continue by intersecting the last term in (5.27) with the event  $\{(\xi_{i,j,n} > u) \cup (\xi_{i,j,n} \leq u)\}$  for  $(i, j)$  equal to  $(a - 1, b)$ . Thus,

we see that

$$\begin{aligned}
& \mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) \\
&= \mathbb{P}((\xi_{a,b,n} > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) \\
&\quad + \mathbb{P}((M_{(1,1):(a-1,b)}^a > u) \cap (M_{(a,b):(a,a)}^a \leq u) \cap (\xi_{a-1,b,n} > u)) \\
&\quad + \mathbb{P}((M_{(1,1):(a-1,b)}^a > u) \cap (M_{(a,b):(a,a)}^a \leq u) \cap (\xi_{a-1,b,n} \leq u)) \\
&= \mathbb{P}((\xi_{a,b,n} > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) \\
&\quad + \mathbb{P}((\xi_{a-1,b,n} > u) \cap (M_{(a,b):(a,a)}^a \leq u)) \\
&\quad + \mathbb{P}((M_{(1,1):(a-2,b)}^a > u) \cap (M_{(a-1,b):(a,a)}^a \leq u)).
\end{aligned}$$

Continuing to intersect the last term with  $\{(\xi_{i,j,n} > u) \cup (\xi_{i,j,n} \leq u)\}$  for the appropriate  $(i, j)$  and then simplifying, this pattern gives us that

$$\mathbb{P}((M_{(1,1):(a,b)}^a > u) \cap (M_{(1,b+1):(a,a)}^a \leq u)) = \sum_{i=1}^a \sum_{j=1}^b \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,a)}^a)).$$

Hence, (5.26) is proved.

We will now look at the more general situation when  $1 \leq b \leq a$ . Conditioning on the event  $\{(\xi_{i,j,n} > u) \cup (\xi_{i,j,n} \leq u)\}$  similar to before, we note that

$$\begin{aligned}
\mathbb{P}(M_{(1,1):(a,b)}^a > u) &= \mathbb{P}(\xi_{a,b,n} > u) \\
&\quad + \mathbb{P}((M_{(1,1):(a-1,b)}^a > u) \cap (\xi_{a,b,n} \leq u)) \\
&= \mathbb{P}(\xi_{a,b,n} > u) \\
&\quad + \mathbb{P}((\xi_{a-1,b,n} > u) \cap (M_{(a,b):(a,b)}^a \leq u)) \\
&\quad + \mathbb{P}((M_{(1,1):(a-2,b)}^a > u) \cap (M_{(a-1,b):(a,b)}^a \leq u)) \\
&= \mathbb{P}(\xi_{a,b,n} > u) \\
&\quad + \mathbb{P}((\xi_{a-1,b,n} > u) \cap (M_{(a,b):(a,b)}^a \leq u)) \\
&\quad + \mathbb{P}((\xi_{a-2,b,n} > u) \cap (M_{(a-1,b):(a,b)}^a \leq u)) \\
&\quad + \mathbb{P}((M_{(1,1):(a-3,b)}^a > u) \cap (M_{(a-2,b):(a,b)}^a \leq u)).
\end{aligned}$$

Continuing this pattern and intersecting the last term with  $\{(\xi_{i,j,n} > u) \cup (\xi_{i,j,n} \leq u)\}$  for the appropriate  $(i, j)$  in row  $b$ , we have

$$\begin{aligned} & \mathbb{P}(M_{(1,1):(a,b)}^a > u) \\ &= \mathbb{P}(\xi_{a,b,n} > u) + \sum_{i=1}^{a-1} \mathbb{P}((\xi_{i,b,n} > u) \cap (M_{(i+1,b):(a,b)}^a \leq u)) \\ & \quad + \mathbb{P}((M_{(1,1):(a,b-1)}^a > u) \cap (M_{(1,b):(a,b)}^a \leq u)). \end{aligned} \quad (5.28)$$

Using (5.26), we have

$$\begin{aligned} & \mathbb{P}((M_{(1,1):(a,b-1)}^a > u) \cap (M_{(1,b):(a,b)}^a \leq u)) \\ &= \sum_{i=1}^a \sum_{j=1}^{b-1} \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,b)}^a \leq u)). \end{aligned} \quad (5.29)$$

Substituting the equality from (5.29) into the latter portion of (5.28), we see that

$$\begin{aligned} & \mathbb{P}(M_{(1,1):(a,b)}^a > u) \\ &= \mathbb{P}(\xi_{a,b,n} > u) + \sum_{i=1}^{a-1} \mathbb{P}((\xi_{i,b,n} > u) \cap (M_{(i+1,b):(a,b)}^a \leq u)) \\ & \quad + \sum_{i=1}^a \sum_{j=1}^{b-1} \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,b)}^a \leq u)) \\ &= \mathbb{P}(\xi_{a,b,n} > u) + \sum_{(i,j) \in A} \mathbb{P}((\xi_{i,j,n} > u) \cap (M_{(i+1,j):(a,b)}^a \leq u)). \end{aligned}$$

□

**Lemma 5.7.**

$$\mathbb{P}(M_{q_n} > u_{n^2}) \geq q_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u_{n^2})).$$

*Proof.* First, noting that  $\mathbb{P}(M_{q_n} > u_{n^2}) = \mathbb{P}(M_{(1,1):(q_n,q_n)}^{q_n} > u_{n^2})$  and defining  $Q = \{1, \dots, q_n\}^2 \setminus \{(q_n, q_n)\}$ , Lemma 5.6 gives us that

$$\mathbb{P}(M_{q_n} > u_{n^2}) = \mathbb{P}(\xi_{q_n,q_n,n} > u_{n^2}) + \sum_{(i,j) \in Q} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})).$$

We next note that for all  $(i, j) \in Q$ , by stationarity,

$$\mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \leq \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})), \quad (5.30)$$

and that

$$\mathbb{P}((\xi_{q_n,1,n} > u_{n^2}) \cap (M_{(1,2):(q_n,q_n)}^{q_n} \leq u_{n^2})) \leq \mathbb{P}(\xi_{q_n,q_n,n} > u_{n^2}). \quad (5.31)$$

For horizontal position  $1 \leq i < q_n$ , we use the inequality in (5.30) for each vertical position  $2, \dots, q_n$  to see that

$$\begin{aligned} & \sum_{j=1}^{q_n} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & \geq q_n \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u_{n^2})), \end{aligned} \quad (5.32)$$

and the inequality in (5.31) to see that

$$\begin{aligned} & \sum_{j=1}^{q_n-1} \mathbb{P}((\xi_{q_n,j,n} > u_{n^2}) \cap (M_{(1,j+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) + \mathbb{P}(\xi_{q_n,q_n,n} > u_{n^2}) \\ & \geq q_n \mathbb{P}((\xi_{q_n,1,n} > u_{n^2}) \cap (M_{(1,2):(q_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.33)$$

Obtaining inequalities like (5.32) and (5.33) for all  $1 \leq i \leq q_n$ , we have

$$\begin{aligned} & \mathbb{P}(\xi_{q_n,q_n,n} > u_{n^2}) + \sum_{(i,j) \in Q} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i,j+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & \geq q_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned}$$

□

**Lemma 5.8.**

$$\mathbb{P}(M_{q_n} > u_{n^2}) \leq p_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,l_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right).$$

*Proof.* Similar to the technique used in Lemma 5.6, we will intersect the event  $\{M_{q_n} > u_{n^2}\}$  with the event  $\{(M_{(1,p_n+1):(q_n,q_n)}^{q_n} > u_{n^2}) \cup (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})\}$ . We will then

break up the probability in question by summing over the disjoint events. Thus, we have

$$\begin{aligned}
& \mathbb{P}(M_{q_n} > u_{n^2}) \\
&= \mathbb{P}((M_{(1,1):(q_n,q_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} > u_{n^2})) \\
&\quad + \mathbb{P}((M_{(1,1):(q_n,q_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\
&= \mathbb{P}(M_{(1,p_n+1):(q_n,q_n)}^{q_n} > u_{n^2}) \\
&\quad + \mathbb{P}((M_{(1,1):(q_n,p_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})). \tag{5.34}
\end{aligned}$$

Notice that  $M_{(1,p_n+1):(q_n,q_n)}^{q_n}$  is the maximum over a  $q_n \times l_n$  block of variables with  $q_n$  horizontal positions and  $q_n - p_n = l_n$  vertical positions. Thus, the distribution of  $M_{(1,p_n+1):(q_n,q_n)}^{q_n}$  is the same as  $M_{(1,1):(q_n,l_n)}^{q_n}$ . Using this fact in (5.34), we have

$$\begin{aligned}
\mathbb{P}(M_{q_n} > u_{n^2}) &= \mathbb{P}(M_{(1,1):(q_n,l_n)}^{q_n} > u_{n^2}) \\
&\quad + \mathbb{P}((M_{(1,1):(q_n,p_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})). \tag{5.35}
\end{aligned}$$

Also, notice that

$$\begin{aligned}
\mathbb{P}(M_{(1,1):(q_n,l_n)}^{q_n} > u_{n^2}) &= \mathbb{P}\left(\bigcup_{i=1}^{q_n} \bigcup_{j=1}^{l_n} \xi_{i,j,n} > u_{n^2}\right) \\
&\leq l_n q_n \mathbb{P}(\xi_{1,1,n} > u_{n^2}) \\
&= \frac{l_n q_n}{n^2} n^2 \mathbb{P}(\xi_{1,1,n} > u_{n^2}).
\end{aligned}$$

Using the fact that

$$\begin{aligned}
n(1 - \Phi(u_n(x))) &\rightarrow \exp(-x) \\
&= O(1),
\end{aligned}$$

we have

$$\mathbb{P}(M_{(1,1):(q_n,l_n)}^{q_n} > u_{n^2}) \leq O\left(\frac{l_n q_n}{n^2}\right). \tag{5.36}$$

Combining (5.35) and (5.36) gives us

$$\mathbb{P}(M_{q_n} > u_{n^2}) \leq \mathbb{P}((M_{(1,1):(q_n,p_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right). \quad (5.37)$$

From Lemma 5.6 we know that

$$\begin{aligned} & \mathbb{P}((M_{(1,1):(q_n,p_n)}^{q_n} > u_{n^2}) \cap (M_{(1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ &= \sum_{i=1}^{q_n} \sum_{j=1}^{p_n} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.38)$$

By stationarity, we see that for  $(i, j) \in \{1, \dots, q_n\} \times \{1, \dots, p_n\}$ ,

$$\begin{aligned} & \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & \leq \mathbb{P}((\xi_{i,p_n+1,n} > u_{n^2}) \cap (M_{(i+1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.39)$$

Once again appealing to stationarity and carefully noting the sizes of the regions in question, we see that

$$\begin{aligned} & \mathbb{P}((\xi_{i,p_n+1,n} > u_{n^2}) \cap (M_{(i+1,p_n+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ &= \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(l_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.40)$$

Fixing the horizontal position to be  $i$ , we use (5.39) and (5.40) to see that

$$\begin{aligned} & \sum_{j=1}^{p_n} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & \leq p_n \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(l_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.41)$$

Applying the inequality in (5.41) for each  $1 \leq i \leq q_n$ , we have

$$\begin{aligned} & \sum_{i=1}^{q_n} \sum_{j=1}^{p_n} \mathbb{P}((\xi_{i,j,n} > u_{n^2}) \cap (M_{(i+1,j):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & \leq p_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(l_n,q_n)}^{q_n} \leq u_{n^2})). \end{aligned} \quad (5.42)$$

Combining the information in (5.37), (5.38), and (5.42), we see that

$$\mathbb{P}(M_{q_n} > u_{n^2}) \leq p_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(l_n,q_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right).$$

□

**Lemma 5.9.** *Assume conditions A1-A5 of Theorem 5.1. Then for large enough  $n$ ,*

$$\begin{aligned} & \left(1 - p_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,l_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right)\right)^{r_n^2} \\ & \leq \mathbb{P}(M_n \leq u_{n^2}) \\ & \leq \left(1 - q_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u_{n^2}))\right)^{(r_n-1)^2} + o(1). \end{aligned}$$

*Proof.* Using Lemmas 5.4 and 5.8, we have that

$$\begin{aligned} & \mathbb{P}(M_n \leq u_{n^2}) \\ & \geq \mathbb{P}(M_{q_n} \leq u_{n^2})^{r_n^2} \\ & \geq \left(1 - p_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u_{n^2}) \cap (M_{(i+1,1):(q_n,l_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right)\right)^{r_n^2}. \end{aligned}$$

Similarly, using Lemmas 5.5 and 5.7, we see that

$$\begin{aligned} & \mathbb{P}(M_n \leq u_{n^2}) \\ & \leq \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} (1 + o(1)) \\ & \leq \left(1 - q_n \sum_{i=1}^{q_n} \mathbb{P}((\xi_{i,1,n} > u) \cap (M_{(i+1,1):(q_n,q_n)}^{q_n} \leq u))\right)^{(r_n-1)^2} (1 + o(1)). \end{aligned}$$

□

### 5.6.3 Convergence Results

Now that we have constructed bounds for  $\mathbb{P}(M_n \leq u_{n^2})$ , we need to prove certain convergence results for use in the proof of Theorem 5.1.

**Lemma 5.10.** *Let*

$$K \subset \{\mathbb{N} \times \{0\}\} \cup \{\mathbb{Z} \times \mathbb{N}\}$$

*be a bounded set. Assume condition A3 of Theorem 5.1. Then for fixed  $x$ ,*

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{P}\left(\max_{(i,j) \in K} \xi_{i,j,n} \leq u_{n^2} \mid \xi_{0,0,n} > u_{n^2}\right) \\ & = \mathbb{P}\left(E/4 + \sqrt{\frac{1}{2}} \delta_{i,j} W_{i,j} \leq \delta_{i,j}, (i,j) \in K\right), \end{aligned}$$

where  $E$  is a standard exponential random variable and  $\{W_{i,j}\}$  is a mean zero, variance one, Gaussian random field with correlation

$$\mathbb{E}(W_{i,j}W_{k,l}) = \frac{\delta_{i,j} + \delta_{k,l} - \delta_{|i-k|,|j-l|}}{2\sqrt{\delta_{i,j}\delta_{k,l}}}.$$

*Proof.* First, we determine that

$$\begin{aligned} & \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i,j) \in K \mid \xi_{0,0,n} > u_{n^2}) \\ &= \frac{\int_{u_{n^2}}^{\infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i,j) \in K \mid \xi_{0,0,n} = x) \phi(x) dx}{1 - \Phi(u_{n^2})} \\ &= \int_0^{\infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i,j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ &\quad \times \frac{\phi(u_{n^2} + y/u_{n^2})}{u_{n^2}(1 - \Phi(u_{n^2}))} dy \\ &= \int_0^{\infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i,j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ &\quad \times \frac{\phi(u_{n^2} + y/u_{n^2})}{\phi(u_{n^2})} \times \frac{\phi(u_{n^2})/u_{n^2}}{1 - \Phi(u_{n^2})} dy \\ &= \int_0^{\infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i,j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ &\quad \times \exp\left(-y - \frac{y^2}{2u_{n^2}^2}\right) \times \frac{\phi(u_{n^2})/u_{n^2}}{1 - \Phi(u_{n^2})} dy. \end{aligned} \tag{5.43}$$

Let  $\{\eta_{i,j,n}, (i,j) \in K\}$  denote random variables with distribution given by the conditional distribution for

$$\{\xi_{i,j,n}, (i,j) \in K \mid (\xi_{0,0,n} = u_{n^2} + y/u_{n^2})\}.$$

Then

$$(\eta_{i,j,n}, (i,j) \in K)^T \sim N(\mu, \Sigma),$$

where

$$\mu = (\rho_{i,j,n}(u_{n^2} + y/u_{n^2}), (i,j) \in K)^T$$

and

$$\Sigma = (\rho_{|i-k|,|j-l|,n} - \rho_{i,j,n}\rho_{k,l,n})_{(i,j),(k,l) \in K}.$$

Define  $Z_{i,j,n}$  to be the standardized  $\eta_{i,j,n}$  so that

$$Z_{i,j,n} = \frac{\eta_{i,j,n} - \rho_{i,j,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{i,j,n}^2}}.$$

Thus,

$$(Z_{i,j,n}, (i, j) \in K)^T \sim N(\mathbf{0}, R)$$

where

$$R = \left( \frac{\rho_{|i-k|,|j-l|,n} - \rho_{i,j,n}\rho_{k,l,n}}{\sqrt{(1 - \rho_{i,j,n}^2)(1 - \rho_{k,l,n}^2)}} \right)_{(i,j),(k,l) \in K}$$

Standardizing  $\eta_{i,j,n}$ , we obtain

$$\begin{aligned} & \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i, j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ &= \mathbb{P}(\eta_{i,j,n} \leq u_{n^2}, (i, j) \in K) \\ &= \mathbb{P}\left(\frac{y\rho_{i,j,n}}{4} + \sqrt{\frac{1 + \rho_{i,j,n}}{4}} \sqrt{\frac{u_{n^2}^2(1 - \rho_{i,j,n})}{4}} Z_{i,j,n} \leq \frac{u_{n^2}^2(1 - \rho_{i,j,n})}{4}, (i, j) \in K\right). \end{aligned} \quad (5.44)$$

Since  $u_{n^2} \sim 2\sqrt{\log n}$ , we see from condition A3 that

$$\frac{u_{n^2}^2(1 - \rho_{i,j,n})}{4} \rightarrow \delta_{i,j}. \quad (5.45)$$

We also see from condition A3 that  $\rho_{i,j,n} = 1 - \frac{\delta_{i,j}}{\log n} + o\left(\frac{1}{\log n}\right)$  and  $\rho_{i,j,n} \rightarrow 1$ , giving us that

$$\begin{aligned} & (\rho_{|i-k|,|j-l|,n} - \rho_{i,j,n}\rho_{k,l,n}) \log n = (\log n - \delta_{|i-k|,|j-l|} + o(1)) \\ & \quad - (\log n - \delta_{i,j} + o(1)) \left(1 - \frac{\delta_{k,l}}{\log n} + o\left(\frac{1}{\log n}\right)\right) \\ &= \delta_{i,j} + \delta_{k,l} - \delta_{|i-k|,|j-l|} + \frac{\delta_{i,j}\delta_{k,l}}{\log n} + o(1) \\ & \rightarrow \delta_{i,j} + \delta_{k,l} - \delta_{|i-k|,|j-l|}, \end{aligned} \quad (5.46)$$

and

$$\begin{aligned} & (1 - \rho_{i,j,n}^2) \log n (1 - \rho_{k,l,n}^2) \log n = (1 - \rho_{i,j,n}) \log n (1 - \rho_{k,l,n}) \log n \\ & \quad \times (1 + \rho_{i,j,n})(1 + \rho_{k,l,n}) \\ & \rightarrow 4\delta_{i,j}\delta_{k,l}. \end{aligned} \quad (5.47)$$

We use (5.46) and (5.47) to show that

$$\begin{aligned}\mathbb{E}(Z_{i,j,n}Z_{k,l,n}) &= \frac{\rho_{|i-k|,|j-l|,n} - \rho_{i,j,n}\rho_{k,l,n}}{\sqrt{(1 - \rho_{i,j,n}^2)(1 - \rho_{k,l,n}^2)}} \\ &\rightarrow \frac{\delta_{i,j} + \delta_{k,l} - \delta_{|i-k|,|j-l|}}{2\sqrt{\delta_{i,j}\delta_{k,l}}}.\end{aligned}\quad (5.48)$$

Remembering that  $\rho_{i,j,n} \rightarrow 1$  and using (5.45) and (5.48) with (5.44), we have

$$\begin{aligned}\mathbb{P}\left(\frac{y\rho_{i,j,n}}{4} + \sqrt{\frac{1 + \rho_{i,j,n}}{4}}\sqrt{\frac{u_{n^2}^2(1 - \rho_{i,j,n})}{4}}Z_{i,j,n} \leq \frac{u_{n^2}^2(1 - \rho_{i,j,n})}{4}, (i, j) \in K\right) \\ \rightarrow \mathbb{P}\left(\frac{y}{4} + \sqrt{\frac{1}{2}}\delta_{i,j}W_{i,j} \leq \delta_{i,j}, (i, j) \in K\right).\end{aligned}\quad (5.49)$$

Via l'Hôpital's rule we also see that

$$\frac{\phi(u_{n^2})/u_{n^2}}{1 - \Phi(u_{n^2})} \rightarrow 1. \quad (5.50)$$

From equations (5.43), (5.44), (5.49), (5.50), and the dominated convergence theorem we see that

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i, j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ = \int_0^\infty \lim_{n \rightarrow \infty} \mathbb{P}(\xi_{i,j,n} \leq u_{n^2}, (i, j) \in K \mid \xi_{0,0,n} = u_{n^2} + y/u_{n^2}) \\ \quad \times \exp\left(-y - \frac{y^2}{2u_{n^2}^2}\right) \times \frac{\phi(u_{n^2})/u_{n^2}}{1 - \Phi(u_{n^2})} dy \\ = \int_0^\infty \mathbb{P}\left(y/4 + \sqrt{\frac{1}{2}}\delta_{i,j}W_{i,j} \leq \delta_{i,j}, (i, j) \in K\right) \exp(-y) dy \\ = \mathbb{P}\left(E/4 + \sqrt{\frac{1}{2}}\delta_{i,j}W_{i,j} \leq \delta_{i,j}, (i, j) \in K\right),\end{aligned}$$

where  $E$  is a standard exponential random variable independent of  $W_{i,j}$ . □

**Lemma 5.11.** *Fix  $\alpha \in \{1, \dots, q_n\}$ ,  $y > 0$  and let  $m \in \mathbb{N}$ . Define  $A_{\alpha,n}$ ,  $G_{\alpha,m,n}$ , and  $H_{\alpha,m,n}$  as in Theorem 5.1, and assume conditions A3 and A6 of Theorem 5.1 are satisfied. Then for  $(i, j) \in H_{\alpha,m,n}$  and large enough  $m, n$ ,*

$$\frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}} > 0.$$

*Proof.* Nearly identical to the observation in the proof of Theorem 2.1 in Hsing et al. (1996), we see that

$$\begin{aligned}
& \sum_{(i,j) \in H_{\alpha,m,n}} n^{-2 \frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}} \frac{(\log n)^{-\frac{\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}}}{\sqrt{1-\rho_{|i-\alpha|,j-1,n}^2}} \\
&= \sum_{(i,j) \in H_{\alpha,m,n}} n^{-2 \frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}} \frac{(\log n)^{\frac{1-\rho_{|i-\alpha|,j-1,n}}{2(1+\rho_{|i-\alpha|,j-1,n})}}}{\sqrt{(1-\rho_{|i-\alpha|,j-1,n}^2) \log n}} \\
&\geq \sum_{(i,j) \in H_{\alpha,m,n}} \exp\left(-2 \log n \frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}\right) \frac{1}{\sqrt{(1-\rho_{|i-\alpha|,j-1,n}^2) \log n}} \\
&\geq \sum_{(i,j) \in H_{\alpha,m,n}} \frac{\exp(-2(1-\rho_{|i-\alpha|,j-1,n}) \log n)}{\sqrt{2(1-\rho_{|i-\alpha|,j-1,n}) \log n}}. \tag{5.51}
\end{aligned}$$

If there exists  $\epsilon > 0$  such that

$$\bigvee_{(i,j) \in H_{\alpha,m,n}} ((1-\rho_{|i-\alpha|,j-1,n}) \log n)^{-1} > \epsilon,$$

for all  $m, n$ , then from (5.51) we have that

$$\sum_{(i,j) \in H_{\alpha,m,n}} n^{-2 \frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}} \frac{(\log n)^{-\frac{\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}}}{\sqrt{1-\rho_{|i-\alpha|,j-1,n}^2}} \geq e^{-2/\epsilon} \sqrt{\frac{\epsilon}{2}},$$

contradicting condition A6. Thus, for large enough  $m$  and  $n$ ,

$$\bigvee_{(i,j) \in H_{\alpha,m,n}} ((1-\rho_{|i-\alpha|,j-1,n}) \log n)^{-1} \leq \epsilon,$$

giving us that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \bigvee_{(i,j) \in H_{\alpha,m,n}} ((1-\rho_{|i-\alpha|,j-1,n}) \log n)^{-1} = 0, \tag{5.52}$$

and implying that for any  $B > 0$  and  $(i, j) \in H_{\alpha,m,n}$ ,

$$(1-\rho_{|i-\alpha|,j-1,n}) \log n > B \text{ for large enough } m \text{ and } n. \tag{5.53}$$

Using a proof by contradiction, we will now show that for large enough  $m$  and  $n$ ,

$$\frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1-\rho_{|i-\alpha|,j-1,n}^2}} > 0.$$

Suppose that

$$\frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}} \leq 0. \quad (5.54)$$

Then

$$\begin{aligned} 0 &\geq u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2}) \\ &= u_{n^2}(1 - \rho_{|i-\alpha|,j-1,n}) - \rho_{|i-\alpha|,j-1,n}y/u_{n^2} \\ &\geq u_{n^2}(1 - \rho_{|i-\alpha|,j-1,n}) - y/u_{n^2} \\ &\geq u_{n^2}^2(1 - \rho_{|i-\alpha|,j-1,n}) - y. \end{aligned}$$

Noting that  $u_{n^2} \sim 2\sqrt{\log n}$  and picking  $B > y$ , we have from (5.53) that for large enough  $m$  and  $n$ ,  $u_{n^2}^2(1 - \rho_{|i-\alpha|,j-1,n}) \geq 4B > y$  and  $u_{n^2}^2(1 - \rho_{|i-\alpha|,j-1,n}) - y \geq 0$ , contradicting the assumption in (5.54). Hence, for large enough  $m$  and  $n$ ,

$$\frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}} > 0.$$

□

**Lemma 5.12.** Fix  $\alpha \in \{1, \dots, q_n\}$  and let  $m \in \mathbb{N}$ . Define  $A_{\alpha,n}$ ,  $G_{\alpha,m,n}$ , and  $H_{\alpha,m,n}$  as in Theorem 5.1, and assume conditions A3 and A6 of Theorem 5.1 are satisfied.

Then

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} (\xi_{i,j,n} > u_{n^2}) \mid \xi_{\alpha,1,n} > u_{n^2}\right) = 0.$$

*Proof.* The proof of this result is similar to the proof of Theorem 2.1 in Hsing et al. (1996). Using the argument from Lemma 5.10 we see that

$$\begin{aligned} &\mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} (\xi_{i,j,n} > u_{n^2}) \mid \xi_{\alpha,1,n} > u_{n^2}\right) \\ &\sim \int_0^\infty \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} (\xi_{i,j,n} > u_{n^2}) \mid \xi_{\alpha,1,n} = u_{n^2} + y/u_{n^2}\right) \\ &\quad \times \exp\left(-y - \frac{y^2}{2u_{n^2}^2}\right) dy. \end{aligned} \quad (5.55)$$

As in Lemma 5.10, we assume  $\{\eta_{i,j,n}, (i, j) \in H_{\alpha,m,n}\}$  has the same distribution as

$$\{\xi_{i,j,n}, (i, j) \in H_{\alpha,m,n} \mid (\xi_{\alpha,1,n} = u_{n^2} + y/u_{n^2})\}.$$

Then

$$(\eta_{i,j,n}, (i, j) \in H_{\alpha,m,n})^T \sim N(\mu, \Sigma),$$

where

$$\mu = (\rho_{|i-\alpha|,|j-1,n}(u_{n^2} + y/u_{n^2}), (i, j) \in H_{\alpha,m,n})^T$$

and

$$\Sigma = (\rho_{|i-k|,|j-l,n} - \rho_{|i-\alpha|,|j-1,n}\rho_{|k-\alpha|,|l-1,n})_{(i,j),(k,l) \in H_{\alpha,m,n}}.$$

Define  $Z_{i,j,n}$  to be the standardized  $\eta_{i,j,n}$  so that

$$Z_{i,j,n} = \frac{\eta_{i,j,n} - \rho_{|i-\alpha|,|j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1,n}^2}}.$$

Then

$$(Z_{i,j,n}, (i, j) \in H_{\alpha,m,n})^T \sim N(0, R)$$

where

$$R = \left( \frac{\rho_{|i-k|,|j-l,n} - \rho_{|i-\alpha|,|j-1,n}\rho_{|k-\alpha|,|l-1,n}}{\sqrt{(1 - \rho_{|i-\alpha|,|j-1,n}^2)(1 - \rho_{|k-\alpha|,|l-1,n}^2)}} \right)_{(i,j),(k,l) \in R}$$

Using this with (5.55), we have

$$\begin{aligned} & \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} (\xi_{i,j,n} > u_{n^2}) \mid \xi_{\alpha,1,n} > u_{n^2}\right) \\ & \sim \int_{y=0}^{\infty} \exp\left(-y - \frac{y^2}{2u_{n^2}^2}\right) \\ & \quad \times \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} \left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,|j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1,n}^2}}\right)\right) dy. \end{aligned}$$

To prove this lemma, it suffices to show that for each fixed  $y_0 > 0$ ,

$$\begin{aligned} & \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_0^{y_0} \exp(-y) \\ & \quad \times \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m}} \left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,|j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1,n}^2}}\right)\right) dy = 0. \end{aligned} \quad (5.56)$$

First, we notice that

$$\begin{aligned} & \mathbb{P}\left(\bigcup_{(i,j) \in H_{\alpha,m,n}} \left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right)\right) \\ & \leq \sum_{(i,j) \in H_{\alpha,m,n}} \mathbb{P}\left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right). \end{aligned} \quad (5.57)$$

Lemma 5.11 gives us that for large enough  $m$  and  $n$ ,

$$\frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}} > 0,$$

so we use the inequality  $1 - \Phi(x) \leq x^{-1}\phi(x)$  for  $x > 0$  (cf. Adler and Taylor (2007, p. 9)), to obtain that

$$\begin{aligned} & \mathbb{P}\left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right) \\ & \leq \left(\frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right)^{-1} \\ & \quad \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right)^2\right). \end{aligned} \quad (5.58)$$

We will first bound the exponential part of (5.58). Notice that

$$\begin{aligned} & \left(\frac{u_{n^2} - \rho_{|i-\alpha|,|j-1|,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,|j-1|,n}^2}}\right)^2 \\ & = \frac{1}{1 - \rho_{i,j,n}^2} \left(u_{n^2}^2(1 - \rho_{|i-\alpha|,|j-1|,n})^2 + \frac{y^2 \rho_{|i-\alpha|,|j-1|,n}^2}{u_{n^2}^2} - 2y(1 - \rho_{|i-\alpha|,|j-1|,n})\rho_{|i-\alpha|,|j-1|,n}\right) \\ & \geq u_{n^2}^2 \frac{1 - \rho_{|i-\alpha|,|j-1|,n}}{1 + \rho_{|i-\alpha|,|j-1|,n}} - \frac{2y\rho_{|i-\alpha|,|j-1|,n}}{1 + \rho_{|i-\alpha|,|j-1|,n}} \\ & \geq (4 \log n - \log(2 \log n)) \frac{1 - \rho_{|i-\alpha|,|j-1|,n}}{1 + \rho_{|i-\alpha|,|j-1|,n}} + C, \end{aligned} \quad (5.59)$$

since  $u_{n^2}^2 = 4 \log n - \log(2 \log n) + O(1)$  (cf. Hsing et al. 1996, p. 682), where  $C$  is some constant (not necessarily positive) depending on  $x$  and  $y$ . Thus, letting  $K$

denote some positive constant, (5.59) gives us

$$\begin{aligned}
& \exp\left(-\frac{1}{2}\left(\frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}}\right)^2\right) \\
& \leq \exp\left(-\frac{1}{2}\left((4 \log n - \log(2 \log n))\frac{1 - \rho_{|i-\alpha|,j-1,n}}{1 + \rho_{|i-\alpha|,j-1,n}} + C\right)\right) \\
& \leq Kn^{-2\frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}} \log n} \frac{1-\rho_{|i-\alpha|,j-1,n}}{2(1+\rho_{|i-\alpha|,j-1,n})}. \tag{5.60}
\end{aligned}$$

Next, we bound the denominator in (5.58). Starting with the last part of (5.59), we have

$$\begin{aligned}
& (4 \log n - \log(2 \log n))\frac{1 - \rho_{|i-\alpha|,j-1,n}}{1 + \rho_{|i-\alpha|,j-1,n}} + C \\
& \geq (1 - \rho_{|i-\alpha|,j-1,n}^2)\left(\frac{1}{4}(4 \log n - \log(2 \log n))\right) + C \\
& \geq \frac{3}{4}(1 - \rho_{|i-\alpha|,j-1,n}^2) \log n + C. \tag{5.61}
\end{aligned}$$

From (5.53) in Lemma 5.11, we know that for large enough  $m$  and  $n$ ,

$$\frac{3}{4}(1 - \rho_{|i-\alpha|,j-1,n}^2) \log n + C \geq \frac{1}{2}(1 - \rho_{|i-\alpha|,j-1,n}^2) \log n. \tag{5.62}$$

Combining (5.58)-(5.62), we see that

$$\begin{aligned}
& \mathbb{P}\left(Z_{i,j} > \frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}}\right) \\
& \leq Kn^{-2\frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}} \log n} \frac{\log n^{\frac{1-\rho_{|i-\alpha|,j-1,n}}{2(1+\rho_{|i-\alpha|,j-1,n})}}}{\sqrt{(1 - \rho_{|i-\alpha|,j-1,n}^2) \log n}} \\
& = Kn^{-2\frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}} \log n} \frac{(\log n)^{-\frac{\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}}}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}}.
\end{aligned}$$

Thus, by condition A6 we have

$$\begin{aligned}
& \sum_{(i,j) \in H_{\alpha,m,n}} \mathbb{P}\left(Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}}\right) \\
& \leq \sum_{(i,j) \in H_{\alpha,m,n}} Kn^{-2\frac{1-\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}} \log n} \frac{(\log n)^{-\frac{\rho_{|i-\alpha|,j-1,n}}{1+\rho_{|i-\alpha|,j-1,n}}}}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}} \\
& \rightarrow 0. \tag{5.63}
\end{aligned}$$

Using the dominated convergence theorem with (5.56), (5.57) and (5.63), we see that

$$\begin{aligned} & \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_0^{y_0} \exp(-y) \\ & \times \mathbb{P} \left( \bigcup_{(i,j) \in H_{\alpha,m,n}} \left( Z_{i,j,n} > \frac{u_{n^2} - \rho_{|i-\alpha|,j-1,n}(u_{n^2} + y/u_{n^2})}{\sqrt{1 - \rho_{|i-\alpha|,j-1,n}^2}} \right) \right) dy = 0. \end{aligned}$$

□

#### 5.6.4 Proof of Theorem 6.1

*Proof of Theorem 5.1.* In light of conditions A1-A6, we use Lemmas 5.5 and 5.7 to see that

$$\begin{aligned} & \mathbb{P}(M_n \leq u_{n^2}) \\ & \leq \mathbb{P}(M_{q_n} \leq u_{n^2})^{(r_n-1)^2} (1 + o(1)) \\ & \leq \left( 1 - q_n \sum_{\alpha=1}^{q_n} \mathbb{P}((\xi_{1,\alpha,n} > u) \cap (M_{(1,\alpha+1):(q_n,q_n)}^{q_n} \leq u)) \right)^{(r_n-1)^2} (1 + o(1)), \end{aligned} \quad (5.64)$$

where  $r_n = n/q_n$ . Next, we restructure the inner part of (5.64) to see that

$$\begin{aligned} & q_n \sum_{\alpha=1}^{q_n} \mathbb{P}((\xi_{1,\alpha,n} > u_{n^2}) \cap (M_{(1,\alpha+1):(q_n,q_n)}^{q_n} \leq u_{n^2})) \\ & = \frac{n^2 \mathbb{P}(\xi_{1,\alpha,n} > u_{n^2})}{r_n^2} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{(1,\alpha+1):(q_n,q_n)}^{q_n} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}), \end{aligned} \quad (5.65)$$

giving us that

$$\begin{aligned} & \mathbb{P}(M_n \leq u_{n^2}) \\ & \leq \left( 1 - \frac{n^2 \mathbb{P}(\xi_{1,\alpha,n} > u_{n^2})}{r_n^2} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{(1,\alpha+1):(q_n,q_n)}^{q_n} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \right)^{(r_n-1)^2} + o(1). \end{aligned} \quad (5.66)$$

Our immediate goal is to show that

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{(1,\alpha+1):(q_n,q_n)}^{q_n} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = \vartheta. \quad (5.67)$$

Letting  $M_{G_{\alpha,m,n}} = \max_{(i,j) \in G_{\alpha,m,n}} \{\xi_{i,j,n}\}$  and  $M_{H_{\alpha,m,n}} = \max_{(i,j) \in H_{\alpha,m,n}} \{\xi_{i,j,n}\}$ , the result in (5.67) follows from showing

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = \vartheta \quad (5.68)$$

and

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{H_{\alpha,m,n}} > u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = 0. \quad (5.69)$$

By Lemma 5.12, each summand in (5.69) converges to zero, and hence, the average converges to zero. Turning to (5.68), observe that

$$\begin{aligned} & \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \\ &= \frac{1}{q_n} \sum_{\alpha=1}^m \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \\ & \quad + \frac{1}{q_n} \sum_{\alpha=m+1}^{q_n-m} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \\ & \quad + \frac{1}{q_n} \sum_{\alpha=q_n-m+1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}). \end{aligned} \quad (5.70)$$

Notice that for  $\alpha, \beta \in \{m+1, m+2, \dots, q_n-m\}$ , the sets  $G_{\alpha,m,n}$  and  $G_{\beta,m,n}$  have identical structure, and because of stationarity we have

$$\mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = \mathbb{P}(M_{G_{\beta,m,n}} \leq u_{n^2} \mid \xi_{1,\beta,n} > u_{n^2}). \quad (5.71)$$

Thus, defining

$$K_m = \{ \{-m, \dots, m\} \times \{0, \dots, m\} \} \setminus \{ \{-m, \dots, 0\} \times \{0\} \},$$

we see that for  $\alpha \in \{m+1, m+2, \dots, q_n-m\}$ ,

$$\mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = \mathbb{P}(M_{K_m} \leq u_{n^2} \mid \xi_{0,0,n} > u_{n^2}). \quad (5.72)$$

Combining (5.70) and (5.72) we see that

$$\begin{aligned}
& \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \\
&= \frac{1}{q_n} \sum_{\alpha=1}^m \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) \\
&\quad + \frac{q_n - 2m}{q_n} \mathbb{P}(M_{K_m} \leq u_{n^2} \mid \xi_{0,0,n} > u_{n^2}) \\
&\quad + \frac{1}{q_n} \sum_{\alpha=q_n-m+1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}). \tag{5.73}
\end{aligned}$$

The first and last sums in (5.73) go to zero since both sums are less than  $mq_n^{-1} \rightarrow 0$  and

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{q_n - 2m}{q_n} \mathbb{P}(M_{K_m} \leq u_{n^2} \mid \xi_{0,0,n} > u_{n^2}) = \vartheta,$$

from Lemma 5.10. Thus,

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}(M_{G_{\alpha,m,n}} \leq u_{n^2} \mid \xi_{1,\alpha,n} > u_{n^2}) = \vartheta.$$

Since

$$n(1 - \Phi(u_n)) \rightarrow e^{-x}, \tag{5.74}$$

we have from (5.66) and (5.67) that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_{n^2}) \leq e^{-\vartheta e^{-x}}. \tag{5.75}$$

On the other hand, combining Lemmas 5.4 and 5.8 gives us that

$$\begin{aligned}
& \mathbb{P}(M_n \leq u_{n^2}) \\
& \geq \mathbb{P}(M_{q_n} \leq u_{n^2})^{r_n^2} \\
& \geq \left( 1 - p_n \sum_{\alpha=1}^{q_n} \mathbb{P}((\xi_{1,\alpha,n} > u_{n^2}) \cap (M_{(i,\alpha+1):(l_n,q_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right) \right)^{r_n^2}. \tag{5.76}
\end{aligned}$$

Similar to (5.65), we see that

$$\begin{aligned}
& p_n \sum_{\alpha=1}^{q_n} \mathbb{P}((\xi_{1,\alpha,n} > u_{n^2}) \cap (M_{(i,\alpha+1):(l_n,q_n)}^{q_n} \leq u_{n^2})) + O\left(\frac{l_n q_n}{n^2}\right) \\
&= \frac{p_n n^2 \mathbb{P}(\xi_{1,\alpha,n} > u_{n^2})}{r_n^2} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}((M_{(i,\alpha+1):(l_n,q_n)}^{q_n} \leq u_{n^2}) \mid \xi_{1,\alpha,n} > u_{n^2}) + \frac{1}{r_n^2} O\left(\frac{l_n}{q_n}\right). \tag{5.77}
\end{aligned}$$

Using an argument nearly identical to the one used to get (5.67), we have that

$$\lim_{n \rightarrow \infty} \frac{1}{q_n} \sum_{\alpha=1}^{q_n} \mathbb{P}((M_{(i,\alpha+1):(l_n, q_n)}^{q_n} \leq u_{n^2}) \mid \xi_{1,\alpha,n} > u_{n^2}) = \vartheta. \quad (5.78)$$

By assumption,  $p_n \sim q_n$ , so as in establishing (5.75), we obtain

$$\liminf_{n \rightarrow \infty} \mathbb{P}(M_n \leq u_{n^2}) \geq e^{-\vartheta e^{-x}}. \quad (5.79)$$

Combining (5.79) and (5.75) gives the desired result.

□

## Chapter 6

### CONCLUSION AND FUTURE WORK

We have introduced two new methods for constructing confidence regions for level curves. The first is an extension of Lindgren and Rychlik (1995) and Wameling and Saborowski (2001) and constructs rectangular confidence regions in directions perpendicular to estimated level curves. The dimensions of these regions are chosen using simulation of the response process conditional on the observed data. Each confidence box should individually intersect the true level curve with confidence level  $1 - \alpha$ . The second method constructs a confidence region which contains the true level curve with confidence level  $1 - \alpha$ . The confidence regions are constructed through multiple hypothesis testing using a test statistic derived from simple kriging. The critical value of the tests is adjusted to control the simultaneous Type I error rate and is estimated via simulation. Lastly, we have presented a limiting result for the distribution of the maxima of a triangular sequence of stationary Gaussian random fields on an  $n \times n$  lattice. Under certain dependence and limiting conditions, we were able to show that the maximum of the random fields exhibits extremal clustering in the limit.

There are several ideas related to the construction of confidence regions for level curves that can be explored for future work. One topic for exploration is a practical way to implement a multiple comparison adjustment for the confidence box method. One approach would be to “blur” the confidence boxes by connecting the vertices of the confidence boxes to produce a single confidence region. It seems unlikely that this confidence region would contain the entire level curve with high confidence, but it

may be useful in controlling an alternative error criterion such as the probability that the confidence region contains a large proportion of the level curve. Another topic for future research is to look into alternative approximations for the critical value controlling the simultaneous Type I error rate of the hypothesis tests. As previously discussed, Adler (2008) presents an approximation to the excursion probability of a Gaussian random field which does not require stationarity or isotropy. It might be possible to use this approximation to conservatively estimate the appropriate critical value. The major advantage of this would be that the critical value could be estimated using a closed form expression instead of a comparatively lengthy simulation. An additional topic which may be considered for further study is the use of alternative error criteria in the construction of confidence regions using the hypothesis test method. Currently, this method is designed to ensure that the entire level curve is contained in the confidence region with high probability, but alternative error criteria may also be appropriate and yield smaller confidence regions.

There are open problems related to the limiting theorem given in Chapter 5 which also deserve consideration. Except in rare, artificial situations, we do not know of any situations where  $\vartheta$  is easy to calculate. One could explore other covariance functions satisfying the appropriate conditions to see whether a simple calculation of  $\vartheta$  may be found. Additionally, one might explore the details of accurately and efficiently estimating  $\vartheta$ . Another open problem is whether one can weaken the conditions necessary to obtain the result in Theorem 5.1. Six conditions were necessary to obtain the result in Theorem 5.1; the analogous result in Hsing et al. (1996) required only three of these conditions. Study could be given to whether these additional conditions could be removed. Of particular interest would be the removal of the condition that the correlation function must be non-negative. This condition was necessary to use the normal comparison inequality in Li and Shao (2002), which was used in some of the lemmas of Chapter 5. If another technique for bounding was used, it may be possible to eliminate this condition.

## Appendix A

### APPENDIX: FIGURES FOR CHAPTER 5 SIMULATIONS

We conclude this dissertation by providing graphical results for the approximation results described in Section 5.5 for  $\mathbb{P}(\max_{1 \leq i,j \leq n} Z_{i,j} \leq u)$ , where  $\{Z_{i,j}\}$  is a Gaussian random field on an  $n \times n$  lattice in a  $[0, 20] \times [0, 20]$  region of interest. The empirical estimate of this probability is shown as a solid line, while the approximation is shown as a dashed line. The results are shown for  $n = 12, 36, 108,$  and  $324$  and  $n$  is noted in the caption of each figure. Figures A.1-A.4 are the results for the Gaussian covariance function when  $\phi = 0.1, 0.5, 1,$  and  $2$ . Figures A.5-A.9 are the results for the spherical covariance function when  $\phi = 0.5, 1, 2, 5,$  and  $10$ . Lastly, Figures A.10-A.13 are the results for the exponential covariance function when  $\phi = 0.1, 0.5, 1,$  and  $2$ . Horizontal lines at probabilities  $0.90$  and  $0.95$  have been provided for reference. A summary of these results is given in Section 5.5.4.

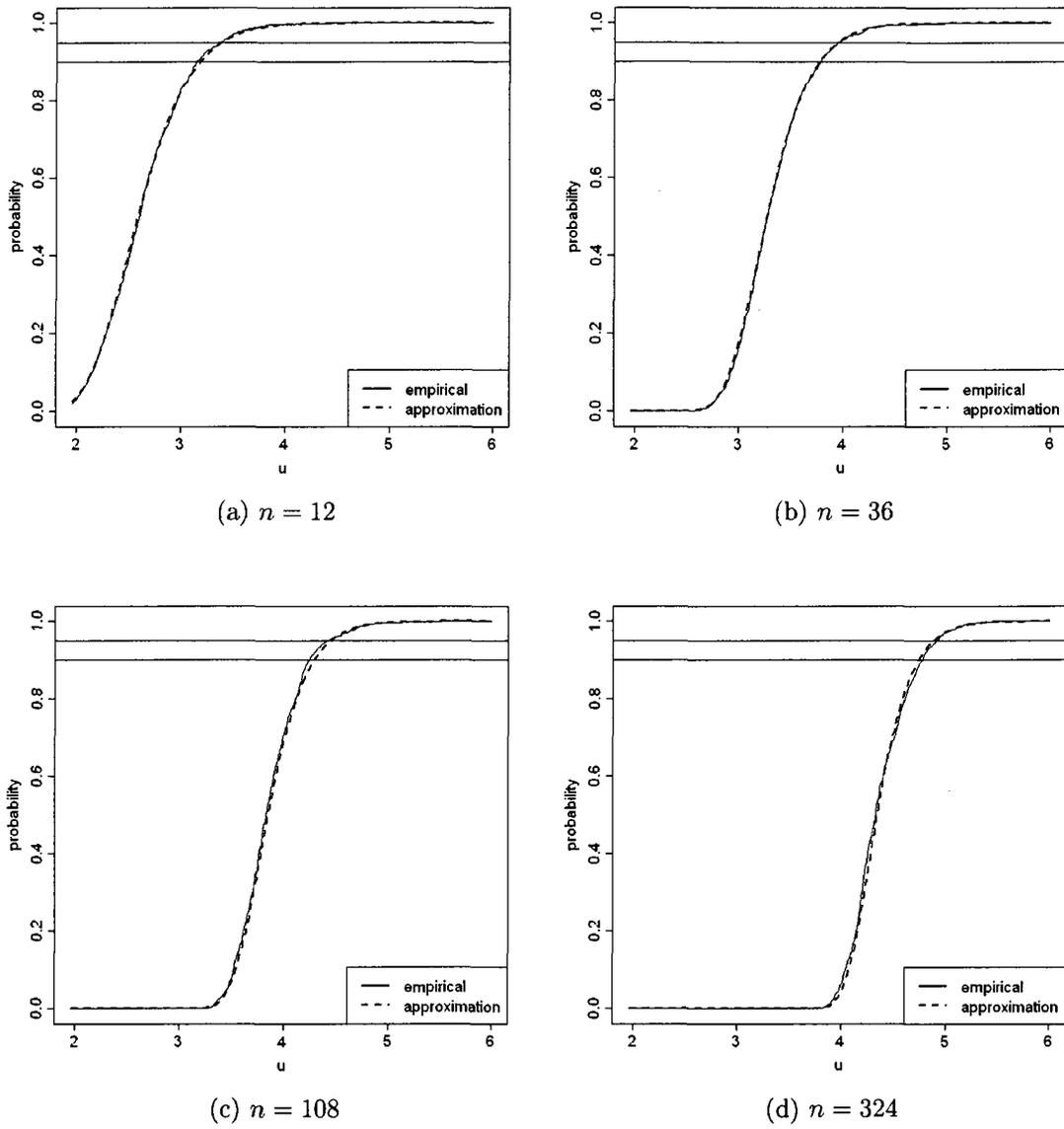


Figure A.1: Approximation for Gaussian covariance function with  $\phi = 0.1$ .

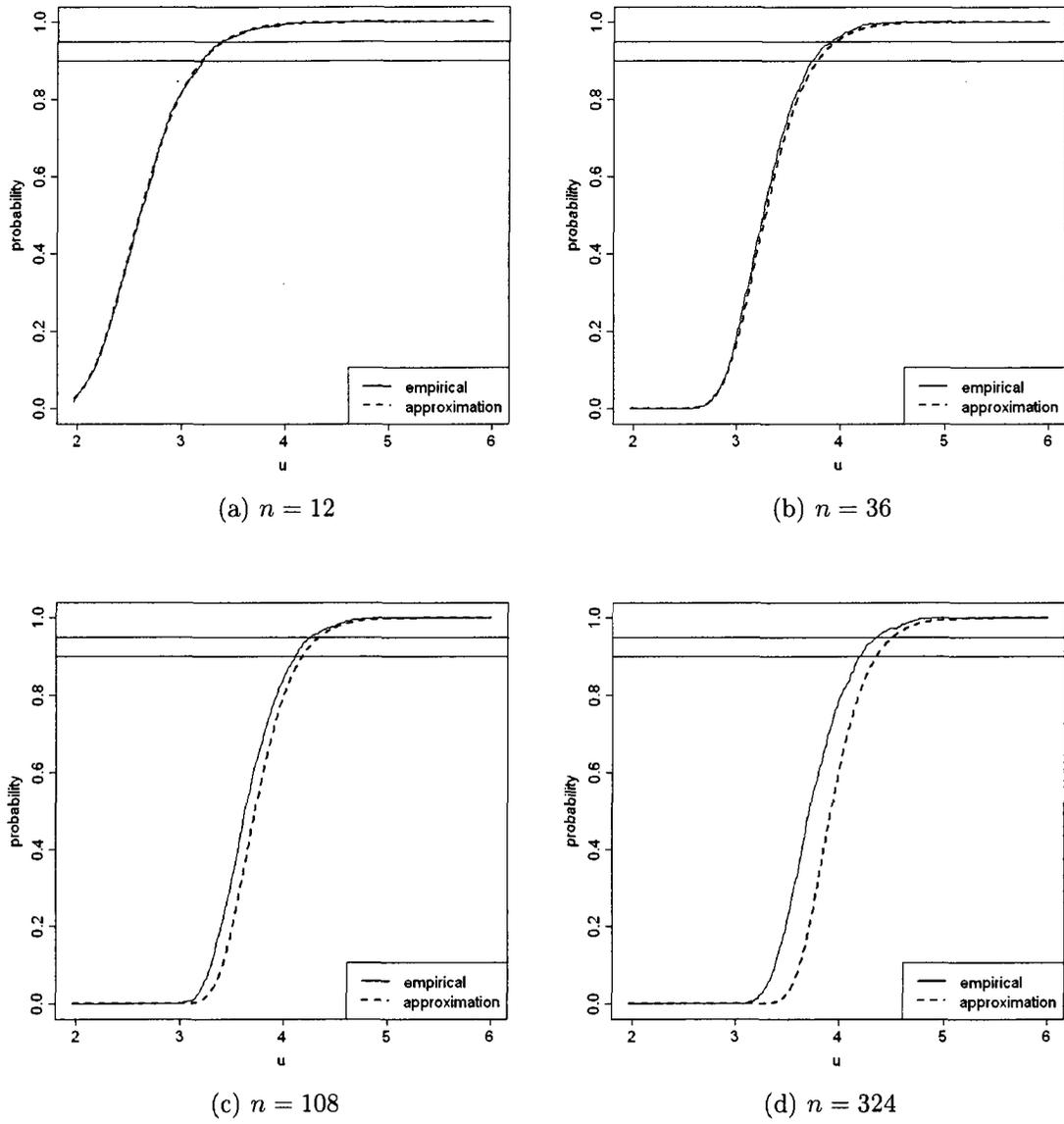


Figure A.2: Approximation for Gaussian covariance function with  $\phi = 0.5$ .

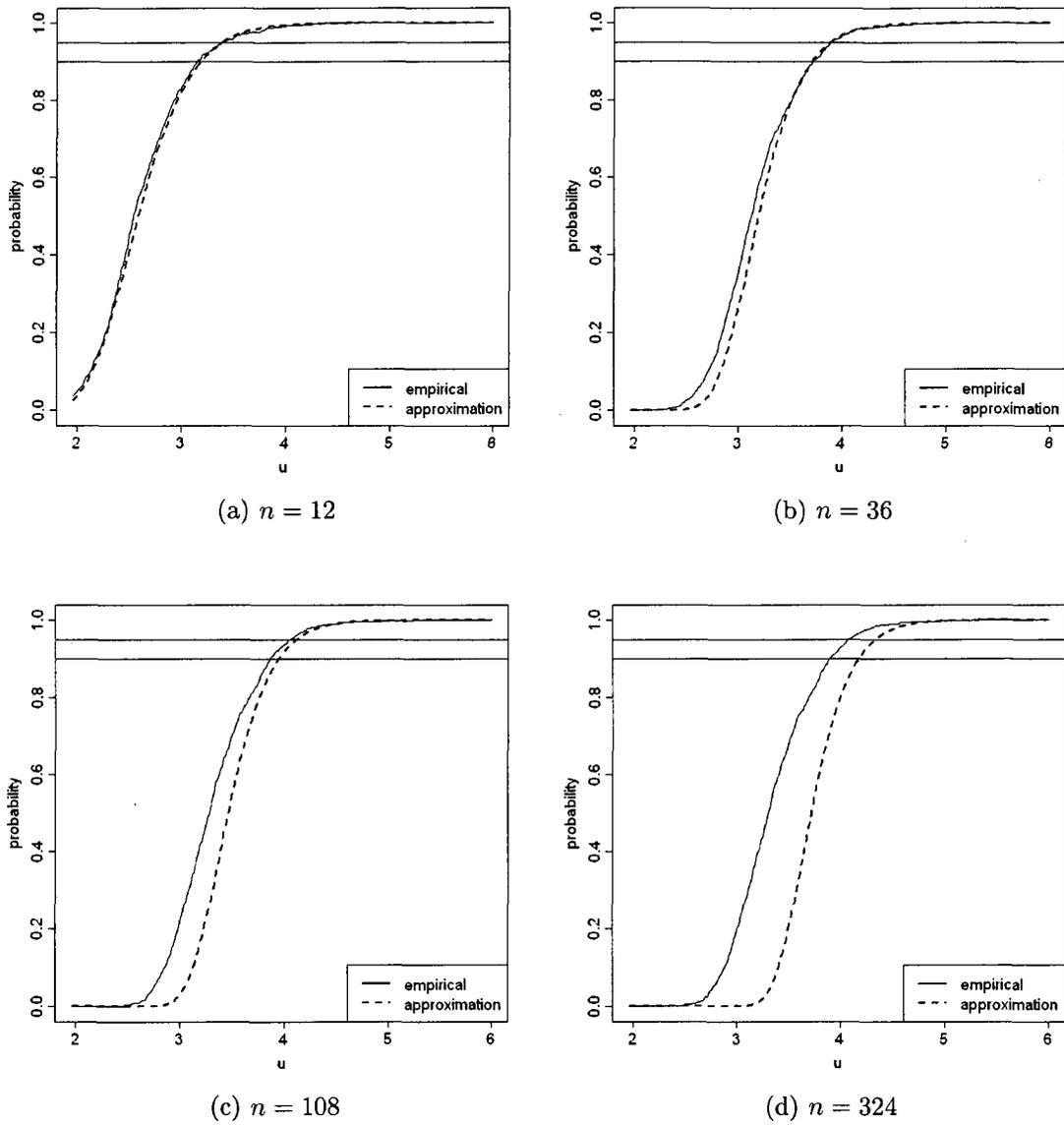


Figure A.3: Approximation for Gaussian covariance function with  $\phi = 1$ .

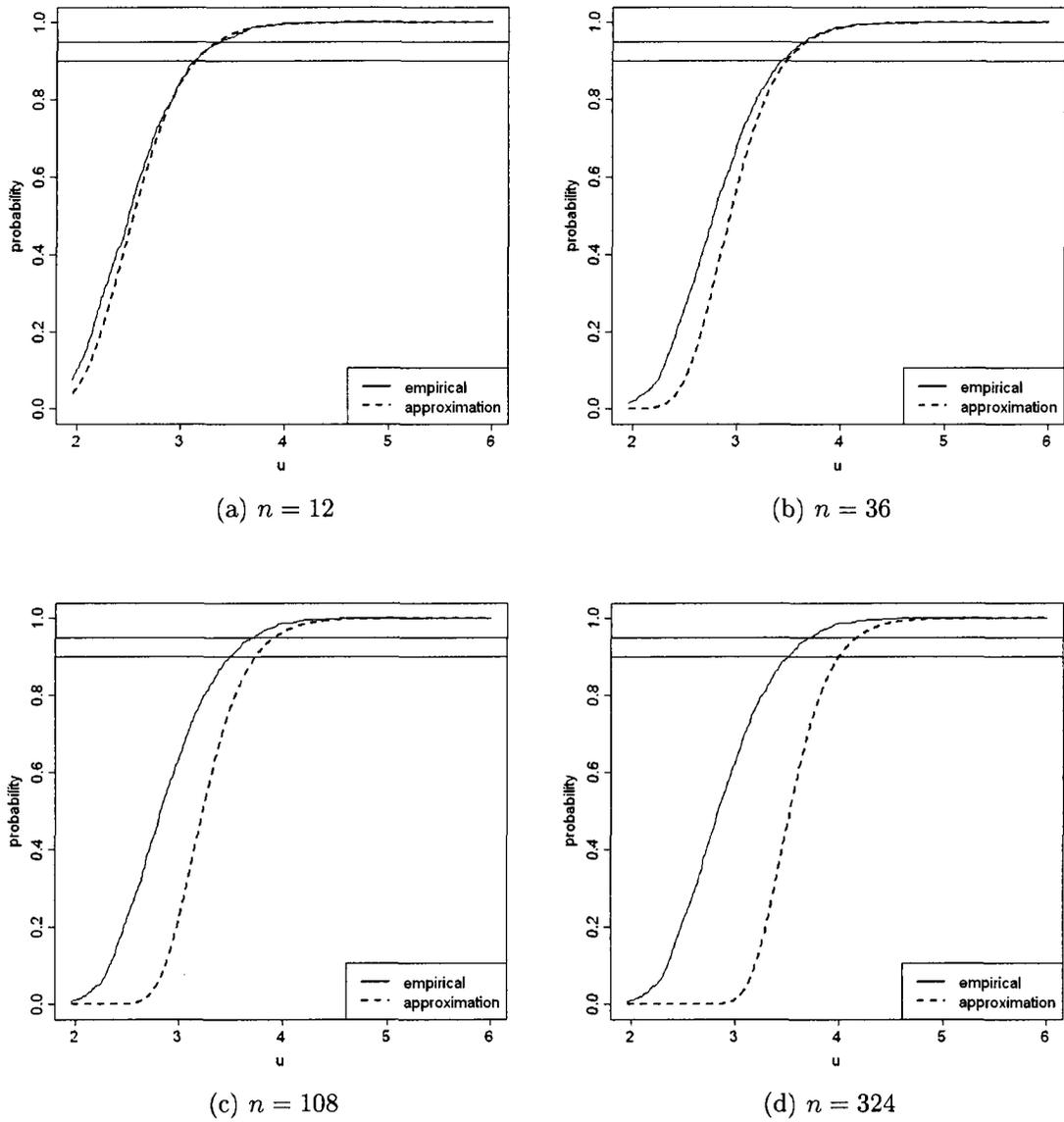


Figure A.4: Approximation for Gaussian covariance function with  $\phi = 2$ .

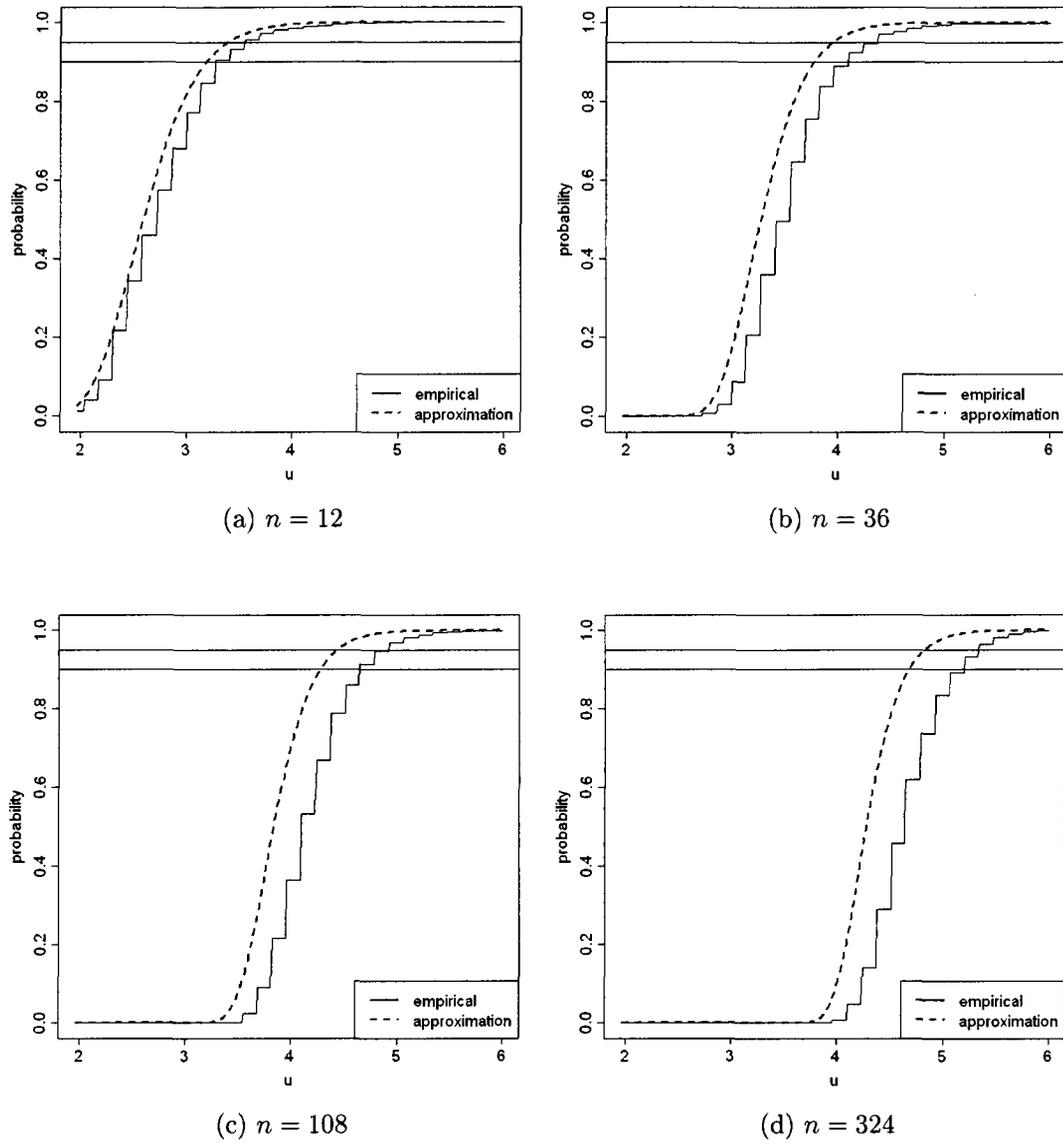


Figure A.5: Approximation for spherical covariance function with  $\phi = 0.5$ .

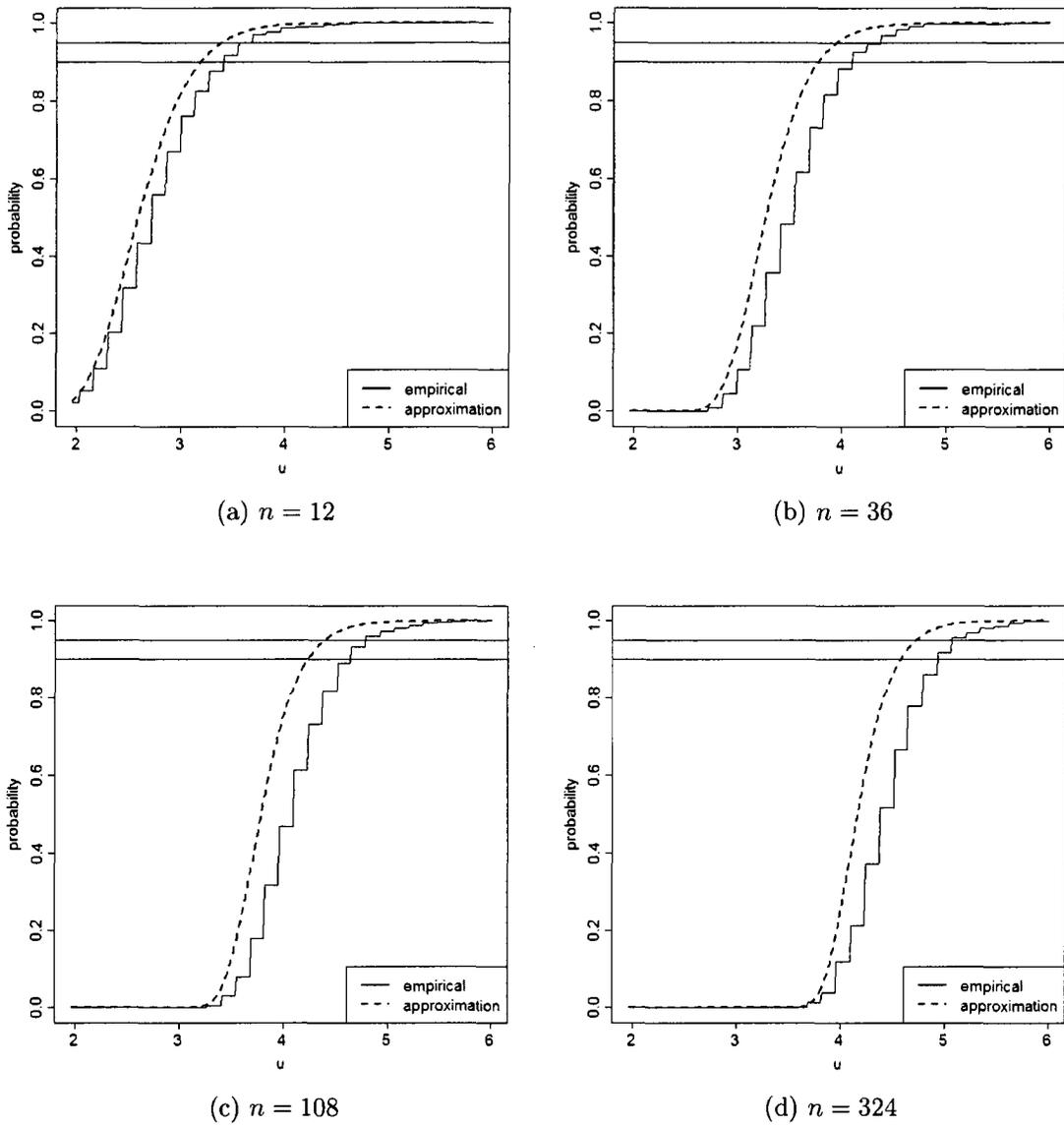


Figure A.6: Approximation for spherical covariance function with  $\phi = 1$ .

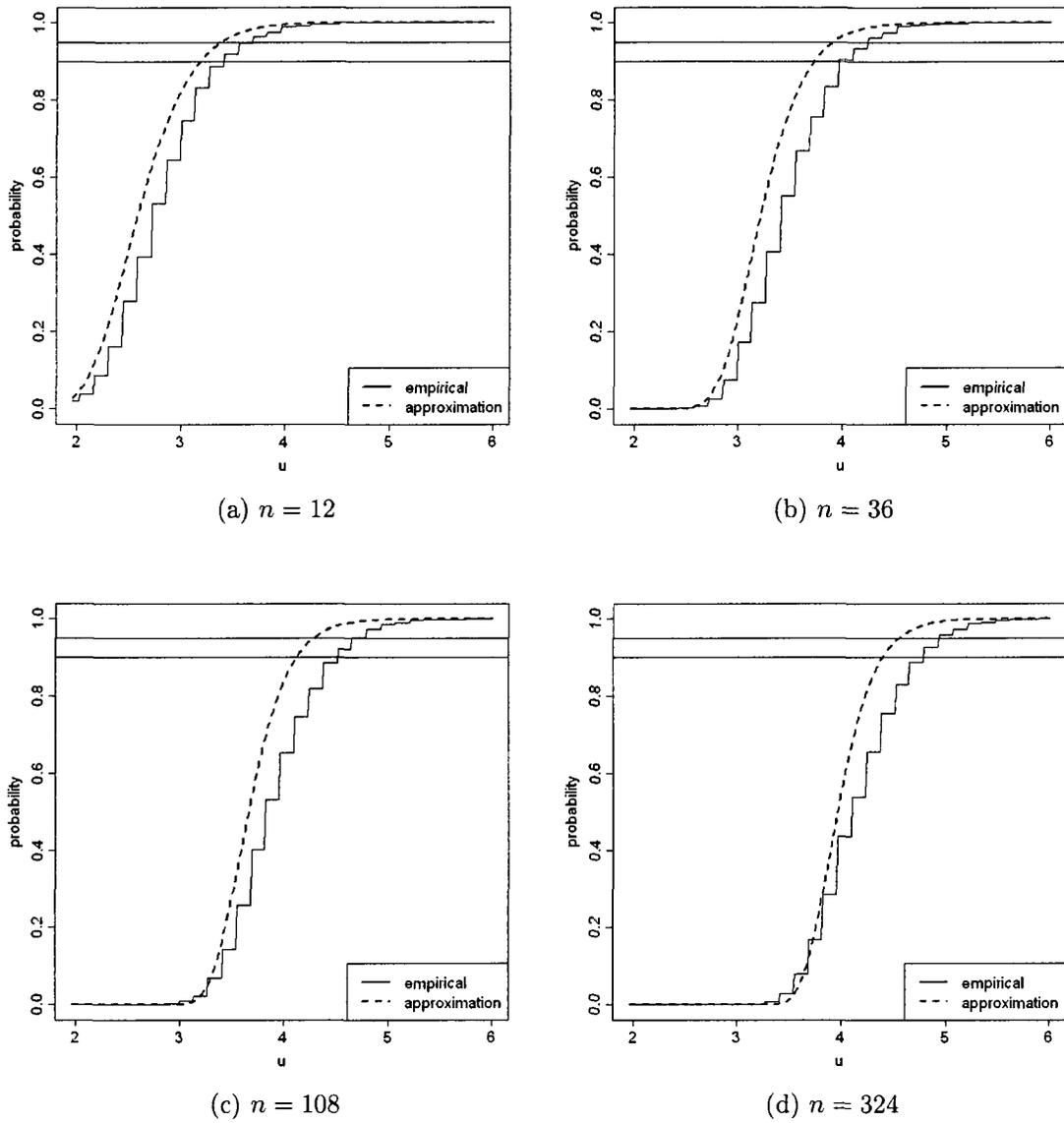
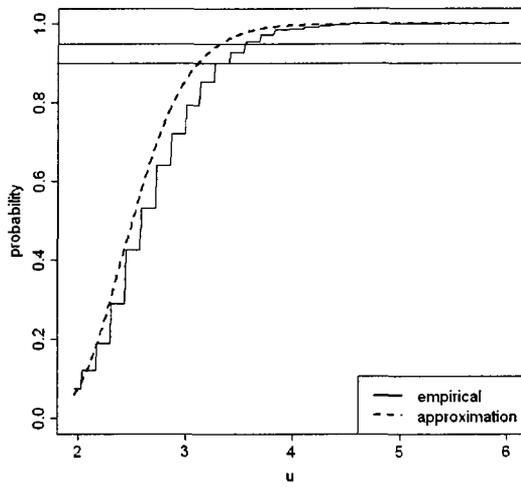
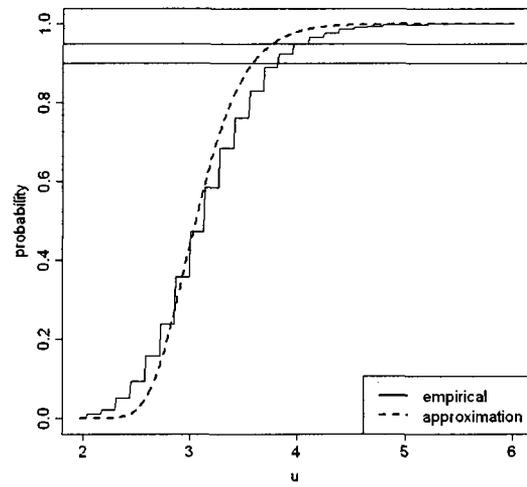
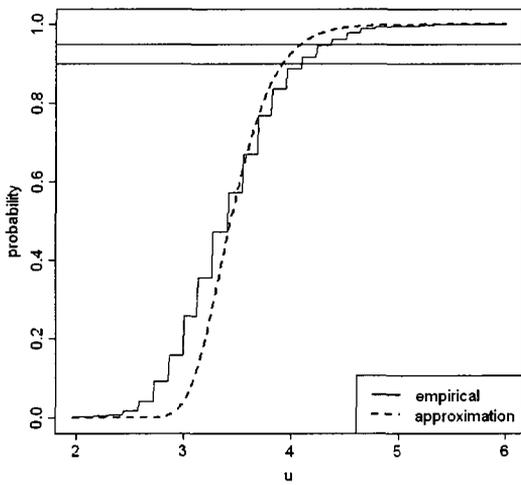
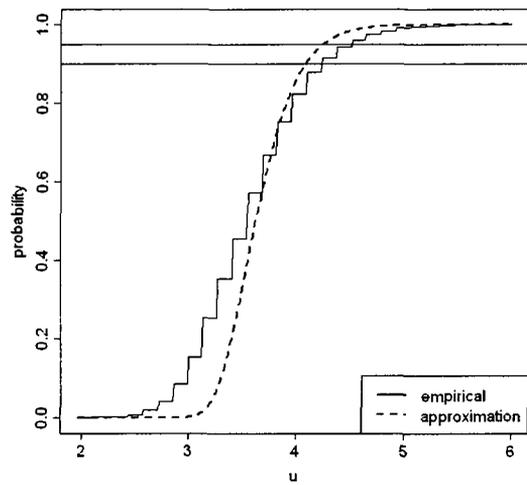


Figure A.7: Approximation for spherical covariance function with  $\phi = 2$ .

(a)  $n = 12$ (b)  $n = 36$ (c)  $n = 108$ (d)  $n = 324$ Figure A.8: Approximation for spherical covariance function with  $\phi = 5$ .

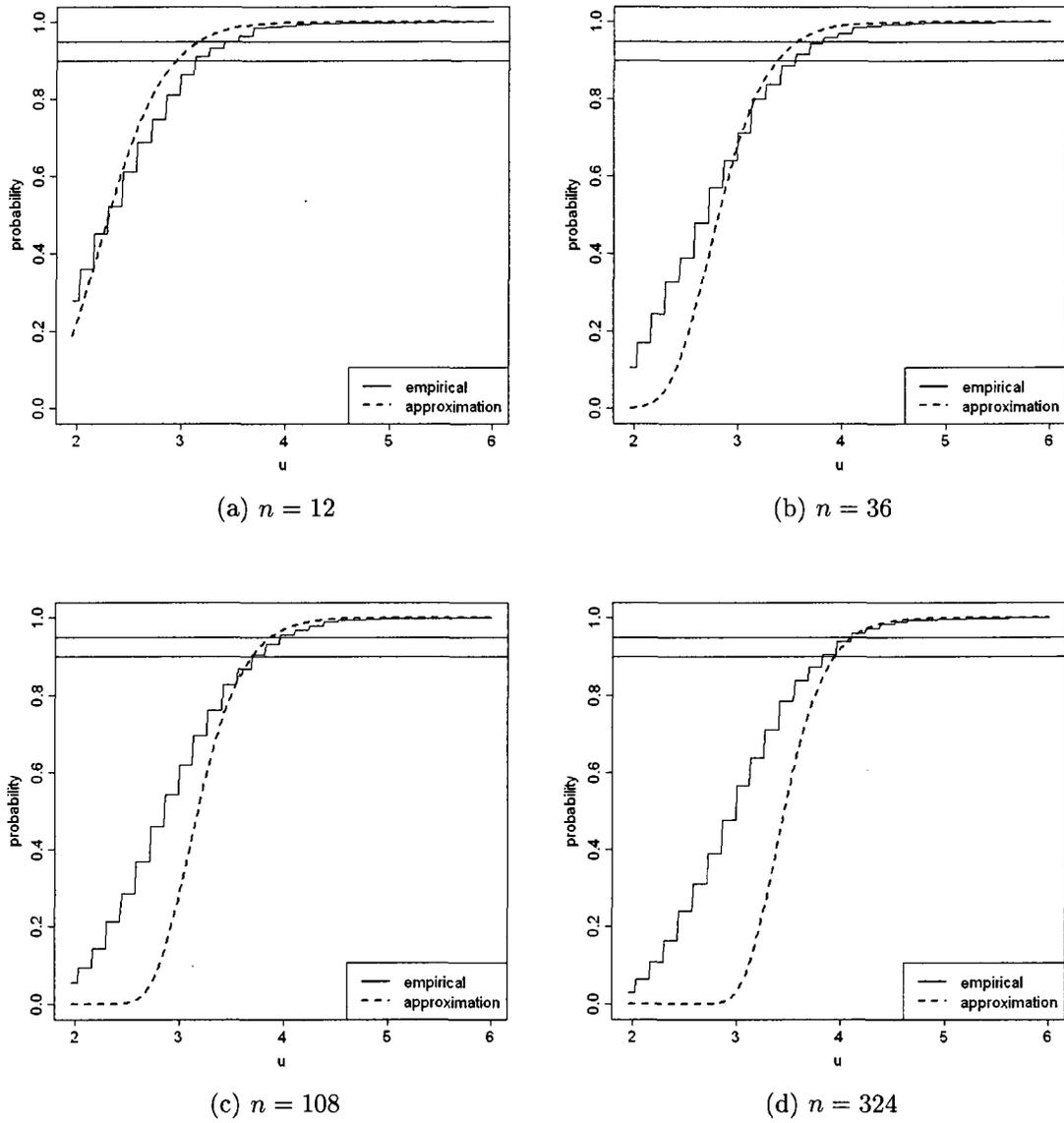
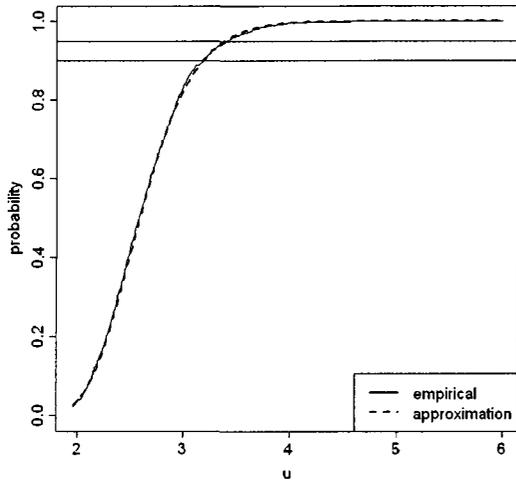
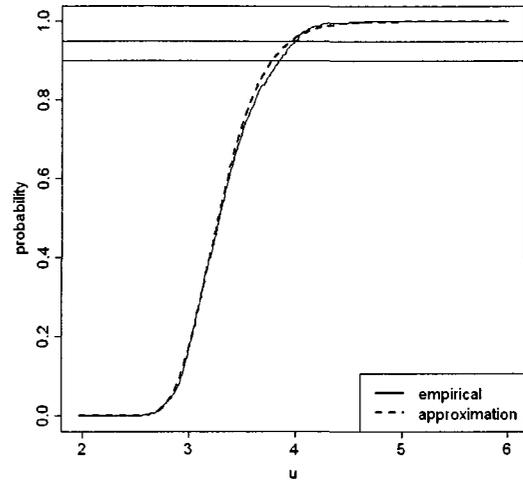
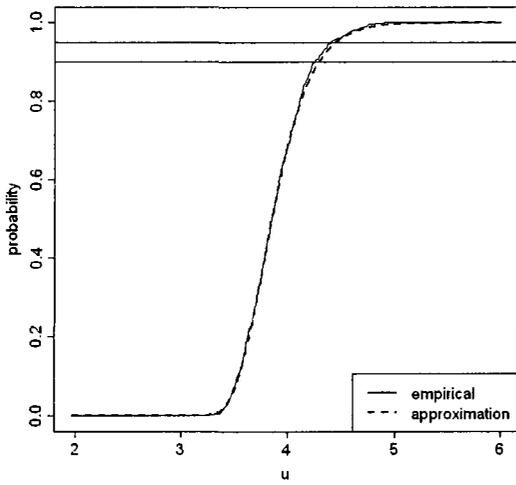
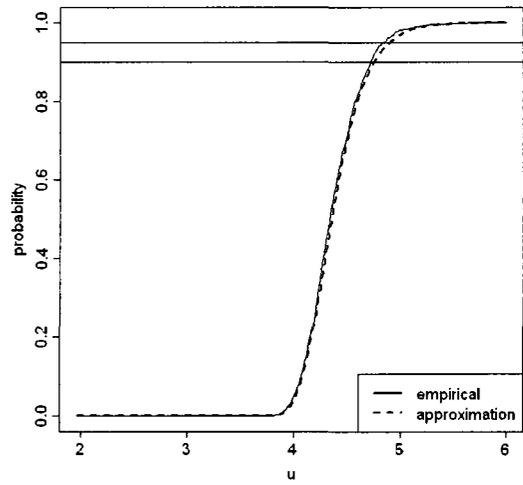


Figure A.9: Approximation for spherical covariance function with  $\phi = 10$ .

(a)  $n = 12$ (b)  $n = 36$ (c)  $n = 108$ (d)  $n = 324$ Figure A.10: Approximation for exponential covariance function with  $\phi = 0.1$ .

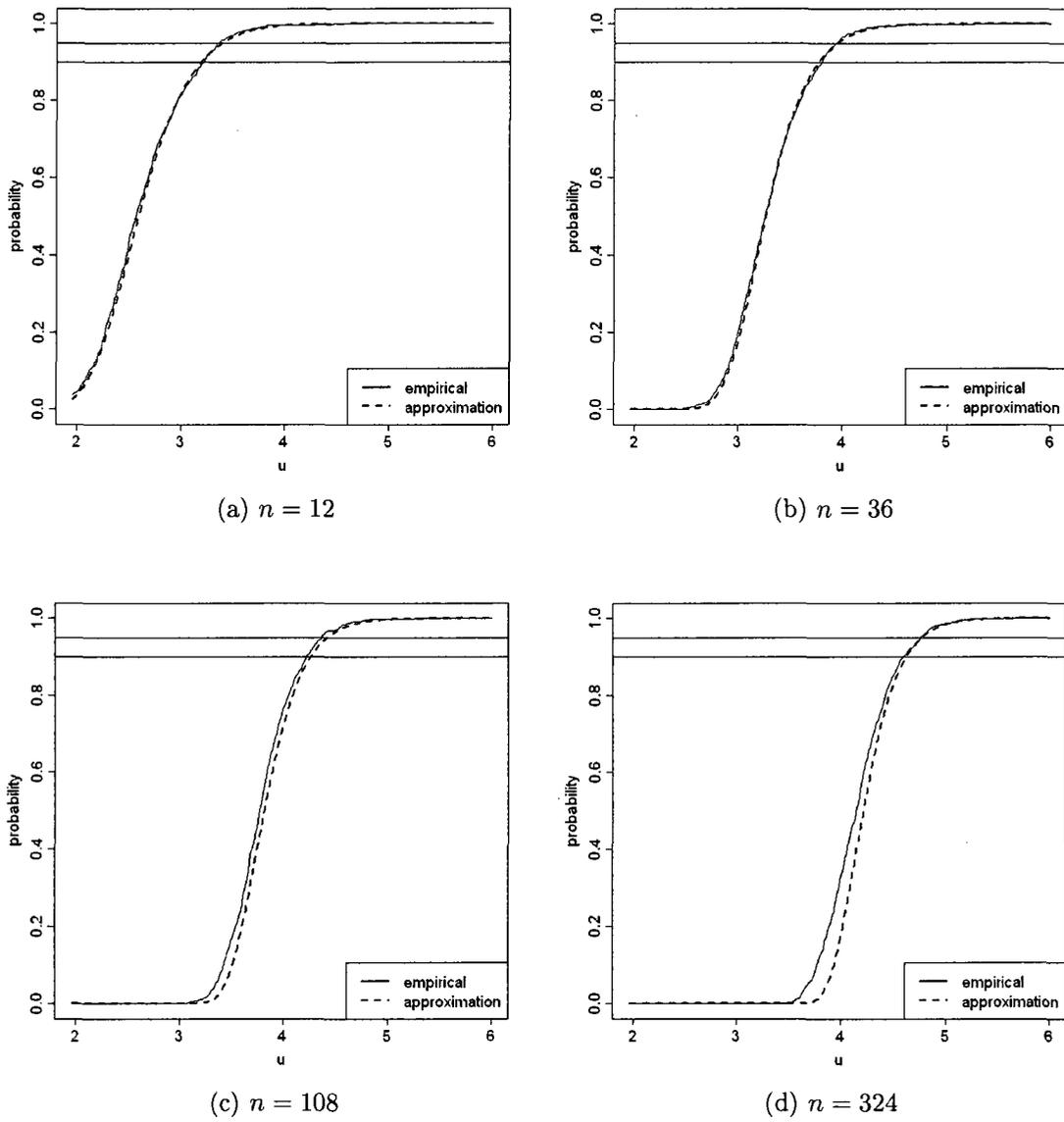


Figure A.11: Approximation for exponential covariance function with  $\phi = 0.5$ .

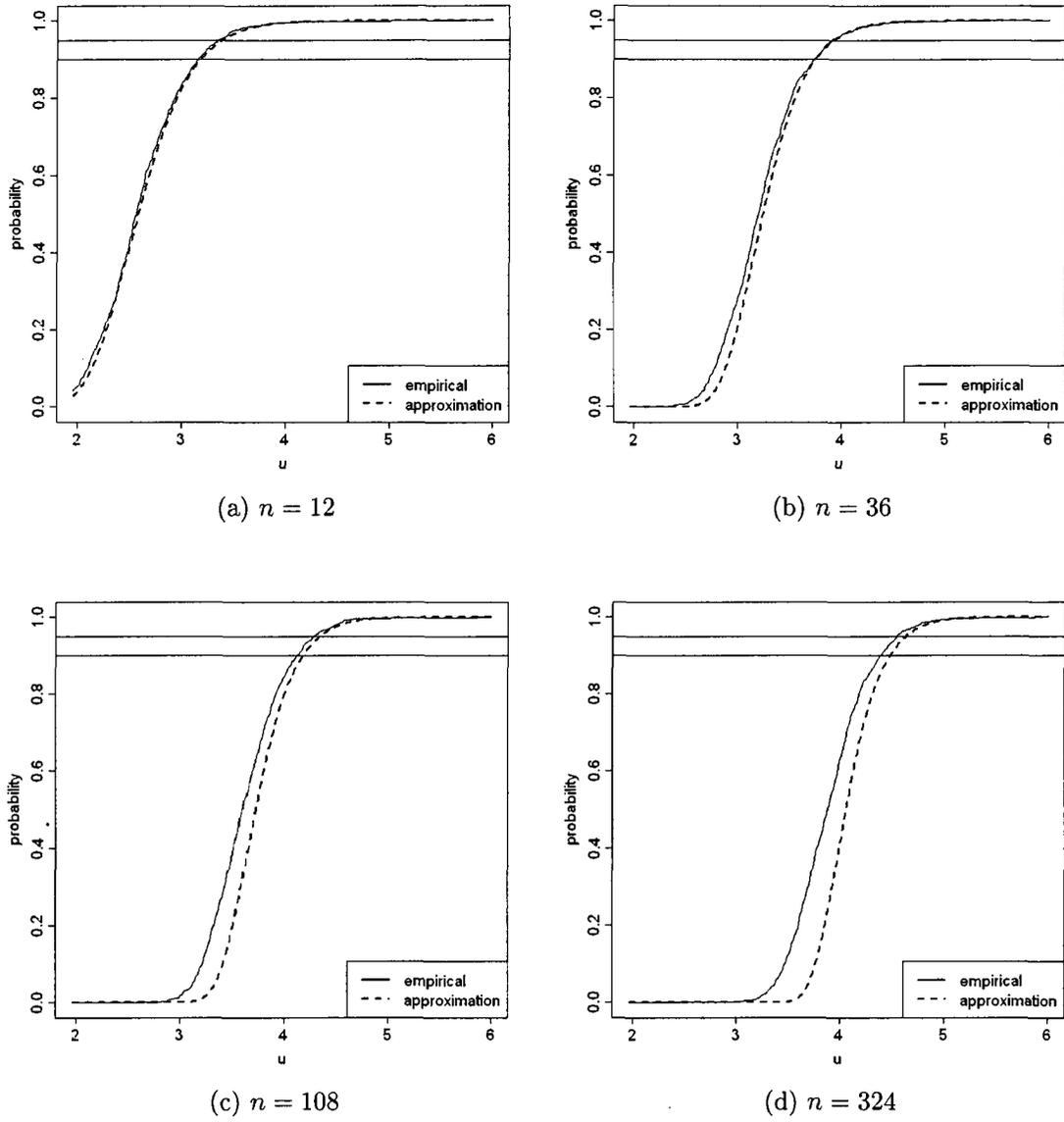


Figure A.12: Approximation for exponential covariance function with  $\phi = 1$ .

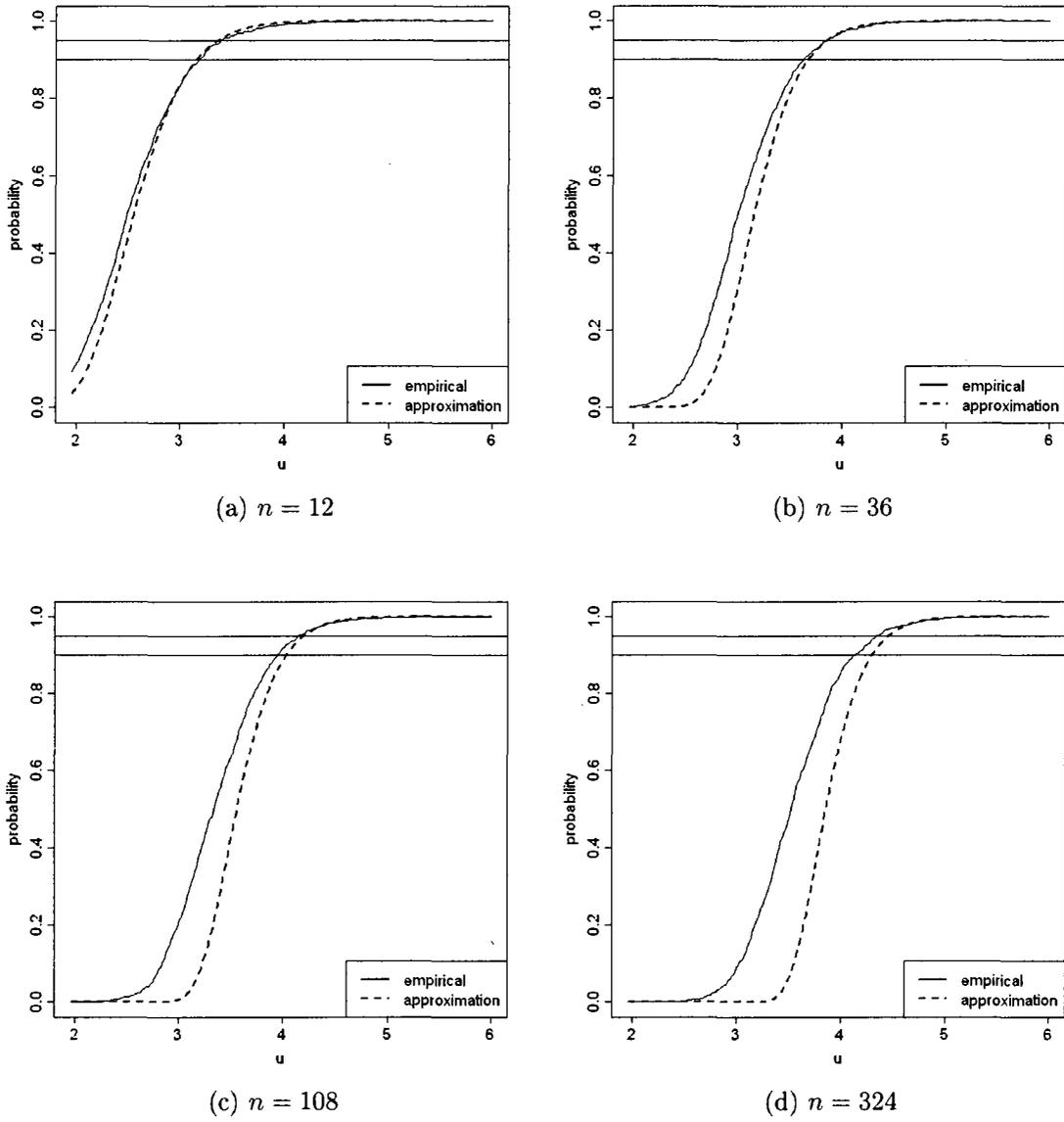


Figure A.13: Approximation for exponential covariance function with  $\phi = 2$ .

## References

- Adler (2000). On excursion sets, tube formulas and maxima of random fields. *The Annals of Applied Probability*, 10:1–74.
- Adler, R. (2008). Some new random field tools for spatial analysis. *Stochastic Environmental Research and Risk Assessment*, 22:809–822.
- Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer.
- Atteia, O., Dubois, J., and Webster, R. (1994). Geostatistical analysis of soil contamination in the Swiss Jura. *Environmental Pollution (United Kingdom)*, 86:315–327.
- Banerjee, S., Carlin, B., and Gelfand, A. (2004). *Hierarchical modeling and analysis for spatial data*. Chapman & Hall.
- Bayarri, M. J., Berger, J. O., Calder, E. S., Dalbey, K., Lunagomez, S., Patra, A. K., Pitman, E. B., Spiller, E., and Wolpert, R. L. (2008). Using Statistical and Computer Models to Quantify Volcanic Hazards. Discussion Paper 2008-07, Duke University Department of Statistical Science, [www.stat.duke.edu/ftp/pub/WorkingPapers/08-07.html](http://www.stat.duke.edu/ftp/pub/WorkingPapers/08-07.html).
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 57(1):289–300.
- Berman, S. (1964). Limit theorems for the maximum term in stationary sequences. *The Annals of Mathematical Statistics*, 35:502–516.
- Box, G. and Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.
- Chiles, J.-P. and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. Wiley.
- Cramer, H. and Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*. John Wiley & Sons.
- Cressie, N. (1993). *Statistics for Spatial Data*. John Wiley & Sons Inc.

- Cullman, A. D. and Saborowski, J. (2005). Estimation of local probabilities for exceeding threshold values in environmental monitoring. *European Journal of Forest Research*, 124:67–71.
- Dudoit, S., van der Laan, M., and Pollard, K. (2004). Multiple testing. Part I. Single-step procedures for control of general type I error rates. *Statistical Applications in Genetics and Molecular Biology*, 3(1):1040.
- Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
- Hannig, J. and Marron, J. (2006). Advanced Distribution Theory for SiZer. *Journal of the American Statistical Association*, 101(474):484–499.
- Hsing, T., Husler, J., and Reiss, R.-D. (1996). The Extremes of a Triangular Array of Normal Random Variables. *The Annals of Applied Probability*, 6:671–686.
- Li, W. V. and Shao, Q.-M. (2002). A Normal Comparison Inequality and its Applications. *Probability Theory and Related Fields*, 122:494–508.
- Lindgren, G. (2004). *Lectures on Stationary Stochastic Processes*. Lund University. Page 35.
- Lindgren, G. and Rychlik, I. (1995). How Reliable Are Contour Curves? Confidence Sets for Level Contours. *Bernoulli*, 1(4):301–319.
- Lindgren, G. and Rychlik, I. (1996). Confidence sets for level contours on sparsely observed images. In P., L. and G., S., editors, *Proceedings symposium on image analysis, Lund, Sweden*, pages 159–163.
- O’Brien, G. L. (1987). Extreme Values for Stationary and Markov Sequences. *The Annals of Probability*, 15(1):281–291.
- Paciorek, C. (2003). *Nonstationary Gaussian processes for regression and spatial modelling*. PhD thesis, Carnegie Mellon University.
- Polfeldt, T. (1999). On the Quality of Contour Maps. *Environmetrics*, 10:785–790.
- Rice, S. O. (1944). Mathematical Analysis of Random Noise. *Bell Systems Tech. J.*, 23:282–332.
- Rootzen, H. (1983). The rate of convergence of extremes of stationary normal sequences. *Advances in Applied Probability*, 15(1):54–80.
- Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC.
- Stein, M. (1999). *Interpolation of spatial data: some theory for kriging*. Springer.

- Takemura, A. and Kuriki, S. (2002). On the equivalence of the tube and Euler characteristic methods for the distributions of the maximum of Gaussian fields over piecewise smooth domains. *Annals of Applied Probability*, 12(2):768–796.
- Takemura, A. and Kuriki, S. (2003). Tail probability via the tube formula when the critical radius is zero. *Bernoulli*, 9(3):535–558.
- Taylor, J. (2006). A Gaussian kinematic formula. *Annals of Probability*, 34(1):122–158.
- Taylor, J., Takemura, A., and Adler, R. (2005). Validity of the expected Euler characteristic heuristic. *Annals of Probability*, 33(4):1362–1396.
- Wameling, A. (2003a). Accuracy of geostatistical prediction of yearly precipitation in Lower Saxony. *Environmetrics*, 14:699–709.
- Wameling, A. (2003b). Vertical and Horizontal Spatial Variation of Geostatistical Prediction. In Foody, G. M. and Atkinson, P. M., editors, *Uncertainty in Remote Sensing and GIS*, chapter 13, pages 209–222. Wiley.
- Wameling, A. and Saborowski, J. (2001). *Construction of local confidence intervals for contour lines*. Greenwich: University of Greenwich.
- Webster, R., Atteia, O., and Dubois, J. (1994). Coregionalization of trace metals in the soil in the Swiss Jura. *European Journal of Soil Science*, 45(2):205–218.
- Webster, R. and Oliver, M. (2007). *Geostatistics for Environmental Scientists*. Wiley.
- Wilhem, J. (2002). A Simulation Study on Competing Distributions for the Maxima of Stationary Normal Processes. Unpublished Master’s Thesis, Colorado State University.