

Detecting conjunctions using cluster volumes

Moh'd Alodat

Department of Mathematics and Statistics
McGill University
Montréal, Québec
Canada

A Thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

© Moh'd Alodat, 2004



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-06266-5

Our file Notre référence

ISBN: 0-494-06266-5

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Contents

Abstract	4
Résumé	5
Statement of Originality	6
Acknowledgments	7
1 Introduction	8
1.1 Brain mapping	8
1.2 Detecting activation	9
1.2.1 Value of the random field	9
1.2.2 Volume of clusters of the excursion set	9
1.3 Conjunctions	13
1.4 Conjunction cluster volume	13
1.5 Outline of the thesis	15
2 Random field theory	20
2.1 Random fields	20
2.2 Continuity and differentiability of random fields	22
2.3 Horizontal window conditioning	24
2.4 Random field near local maxima	24
2.5 Integral geometry	26
2.6 EC densities of the conjunction	29
2.7 Poisson clumping heuristic	30
3 Volumes of clusters of conjunctions	31
3.1 Distribution of the volume of a cluster	31
3.2 $E\{V\}$ using the balls model	36
3.3 $E\{V\}$ using the Poisson clumping heuristic	37

3.4	Distribution of the maximum cluster volume	40
3.5	Correcting the mean of the distribution of V	41
3.6	Example of a one-dimensional stationary Gaussian random field	41
3.7	The χ^2 random field	43
4	Simulation	44
4.1	Simulation of V from the balls model	44
4.2	Simulation of V and V_{max} from random fields	46
5	Application	52
6	Conclusion	56

Abstract

In brain mapping, the regions of the brain that are ‘activated’ by a task or external stimulus are detected by thresholding an image of test statistics. Often the experiment is repeated on several different subjects or for several different stimuli on the same subject, and the researcher is interested in the common points in the brain where ‘activation’ occurs in all test statistic images. The conjunction is thus defined as those points in the brain that show ‘activation’ in all images. We are interested in which parts of the conjunction are noise, and which show true activation in all test statistic images. We would expect truly activated regions to be larger than usual, so our test statistic is based on the volume of clusters (connected components) of the conjunction. Our main result is an approximate P-value for this in the case of the conjunction of two Gaussian or χ^2 test statistic images. The results are applied to a functional magnetic resonance experiment in pain perception.

Résumé

En cartographie cérébrale, les régions du cerveau qui sont activées par une tâche ou un stimulus externe sont détectées par seuillage d'une image de statistiques de test. L'expérience est souvent répétée sur plusieurs sujets différents ou selon différents stimuli sur le même sujet. Le chercheur est alors intéressé par les points d'activation communs à toutes les images de statistiques de test. La conjonction est définie comme étant l'ensemble des points dans le cerveau qui démontre une activation dans toutes les images. Nous désirons départager les parties de la conjonction qui sont réelles de celles qui sont formées de bruit parmi toutes les images de statistiques de test. Puisque nous nous attendons à ce qu'une activation réelle suscite une conjonction plus grande qu'une conjonction due au hasard, notre statistique de test est basée sur le volume des amas (c'est-à-dire des composantes connectées) de la conjonction. Notre résultat principal est une valeur p approximative pour cette statistique de test dans le cas d'une conjonction de deux images gaussiennes ou χ^2 . Les résultats sont appliqués à des images par résonance magnétique fonctionnelle obtenues dans le cadre d'une expérience sur la perception de la douleur.

Statement of Originality

In this thesis, I have derived an approximation to the distribution of the volume of one cluster of the excursion set of the conjunction of two independent, smooth, stationary Gaussian (χ^2) random fields. I have solved this problem for general dimension. I have used some tools from integral geometry to find the mean value of this distribution. This mean value is simplified to a closed form. I have applied these results to an fMRI experiment in pain perception.

Acknowledgments

I would like to thank my supervisors, Professors Alain Vandal and Keith Worsley for their help, guidance and encouragement through the time of working in this thesis at McGill University. I am thankful to all professors, staff and students in the Department of Mathematics and Statistics at McGill University who helped me and encouraged me. Many thanks to my parents, brothers and sisters from whom I have been away during my study, for their support and encouragement.

Chapter 1

Introduction

1.1 Brain mapping

In recent years new technologies have been developed to produce informative images about the living human brain. Two of these techniques are positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). These two techniques have enabled neurologists to study the functional activation of the living human brain under different conditions. The data collected by these techniques are smooth images of the brain activity over the time of an experiment. By analyzing this type of data we can detect whether a region in the brain is activated or not.

The simplest way of doing this is to assume that the time course of the images at each point have a Gaussian distribution whose mean follows a linear model with regressors for the presence or absence of the different conditions applied during the course of the experiment (Friston et al., 1995, Worsley et al., 2002). The condition is then detected by a simple T or F test statistic. This is repeated at each point or voxel in the image, and the result is a 3D image of test statistics, $X(\mathbf{t})$, $\mathbf{t} \in S \subset \mathbb{R}^D$. Here $D = 3$ and the search region S is usually the whole brain. We expect a small number of isolated regions of S to be activated, producing high values of the test statistic image $X(\mathbf{t})$. These can then be detected by the excursion set of $X(\mathbf{t})$, defined as the set of points t where $X(\mathbf{t})$ exceeds a threshold x .

Figure 1.1 shows an application to an fMRI experiment in pain perception, fully described

in Worsley *et al.* (2002). During the course of the experiment, a subject was given an alternating 9s hot and 9s warm stimulus to the left calf, interspersed with 9s periods of rest, repeated 10 times. The T statistic $X(\mathbf{t})$ (110 degrees of freedom) for the contrast between the hot and warm stimulus should show those areas of the brain that are activated by the hot pain, compared to just the warm touch. Figure 1.1(b) shows the search region S (the part of the brain covered by the fMRI data), together with the excursion set above a threshold $x = 3.17$ chosen so that the P-value at any point is 0.001.

1.2 Detecting activation

1.2.1 Value of the random field

There are two common approaches to detecting the activated regions in such an image. The first is based on setting the threshold x so that the probability that X exceeds x anywhere in the unactivated parts of S is controlled to be say $\alpha = 0.05$. This is done conservatively by assuming that the unactivated parts cover the whole search region S . The threshold is then chosen so that

$$\mathbb{P}\{\sup_{\mathbf{t} \in S} X(\mathbf{t}) \geq x\} = \alpha.$$

Under the assumption that X is a smooth isotropic random field, good approximations are available for this based on the expected Euler characteristic of the excursion set

$$A = \{\mathbf{t} \in S : X(\mathbf{t}) \geq x\}$$

(Adler, 1981; Worsley, 1994, 1995). Figure 1.1(c) shows the excursion set above a threshold $x = 4.86$ chosen so that the P-value of the maximum in S is $\alpha = 0.05$.

1.2.2 Volume of clusters of the excursion set

The second method is based on the volume or Lebesgue measure of connected components or *clusters* of the excursion set (see Figure 1.1(d)). To do this, we first set the threshold x to a high value, typically chosen so that if there is no activation, $\mathbb{P}\{X(\mathbf{t}) \geq x\} = 0.001$, so that

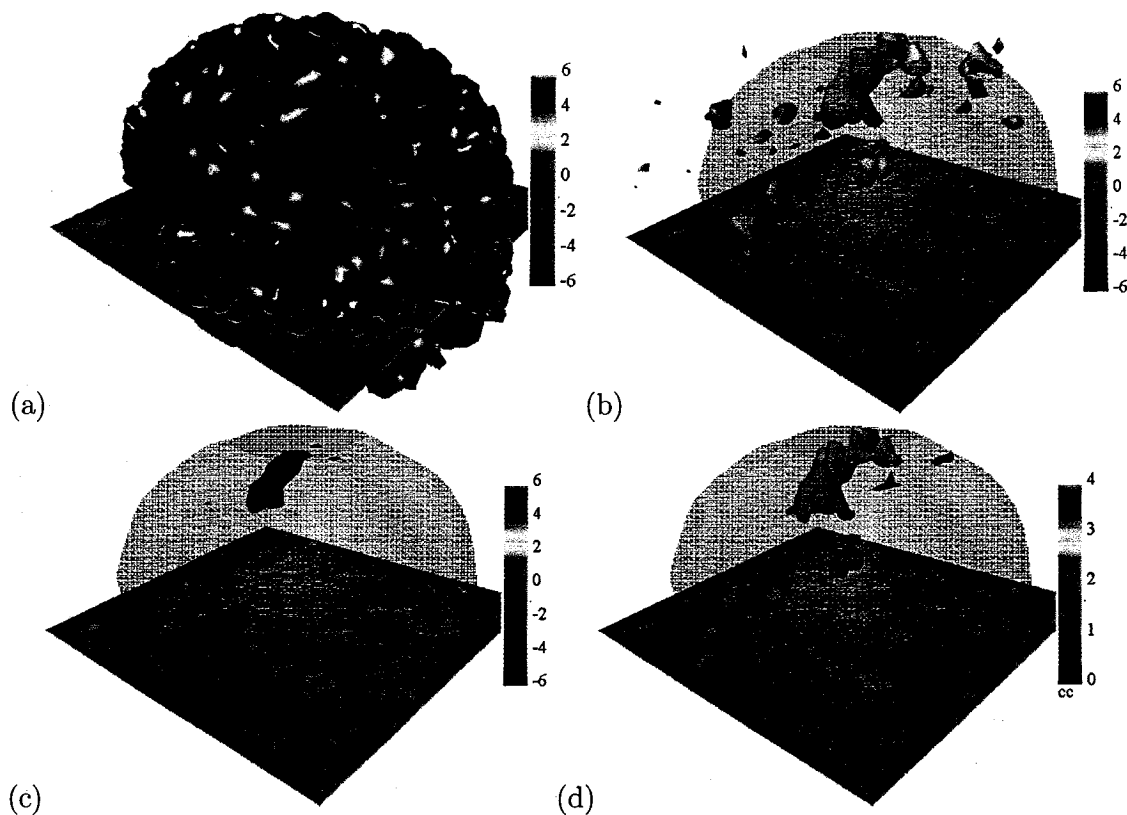


Figure 1.1: Application to pain perception. (a) The brain (back left facing the camera), together with a slice of the T statistic image $X(t)$ (110 df) for a difference between the hot and warm stimulus. (b) Search region S (transparent), and excursion set (blobs) above $x = 3.17$ chosen so that the P-value at any point is less than or equal to 0.001. (c) Excursion set above $x = 4.86$ chosen so that the P-value of the maximum of $X(t)$ inside S is at least 0.05. (d) The clusters of the excursion set above $x = 3.17$ whose volume exceeds $v = 0.61\text{cc}$, chosen so that the P-value of the maximum volume is 0.05, coloured by their volume (the large cluster has a volume of 14.15cc). Note that both methods (c) and (d) detect activation in the right primary somatosensory area, (white cluster in (d)), and the left and right thalamus (green and red clusters in (d)).

we expect 0.1% of the search region to be above threshold if there is no activation. Suppose there are N disjoint clusters C_n with volumes V_n , $n = 1, \dots, N$:

$$A = \cup_{n=1}^N C_n, \quad V_n = |C_n|, \quad n = 1, \dots, N,$$

where $|\cdot|$ denotes Lebesgue measure. Then the activated regions are those clusters of the excursion set with volumes V_n exceeding some threshold v . The threshold v is then chosen conservatively so that, if the whole search region S is unactivated, then

$$\mathbb{P}\{\max_{1 \leq n \leq N} V_n \geq v\} = \alpha.$$

Under the assumption that X is a smooth isotropic random field, good approximations are available for some common test statistics (Friston et al., 1994; Cao, 1999; Hayasaka et al., 2004). Figure 1.1(d) shows the clusters whose volumes exceed the threshold $v = 0.61\text{cc}$, chosen so that $\alpha = 0.05$.

These approximations assume that the search region S is sufficiently large that the clusters rarely intersect the boundary of S . A key step in deriving these approximations is that for high thresholds x the cluster volumes are approximately independent, so that a Bonferroni approximation is quite accurate:

$$\mathbb{P}\{\max_{1 \leq n \leq N} V_n \geq v\} \approx \mathbb{E}\{N\}\mathbb{P}\{V_1 \geq v\}.$$

There are very good approximations to the expected number of clusters $\mathbb{E}(N)$ from work on the expected Euler characteristic of the excursion set (Adler, 1981; Worsley, 1994, 1995). This means that the most important (and most challenging) problem is to find the distribution of the volume of a single randomly chosen cluster (see Figure 1.2(d)).

The motivation for the second method comes from the expectation that activation might be more diffuse and produce larger components of the excursion set. In contrast, the first method is based on the expectation that activation will be more focused and produce larger *values* of the test statistic image. The volume should therefore be more sensitive to activation that is spread over a large region, whereas the value of the test statistic image should be more sensitive to activation that is focused on small isolated regions.

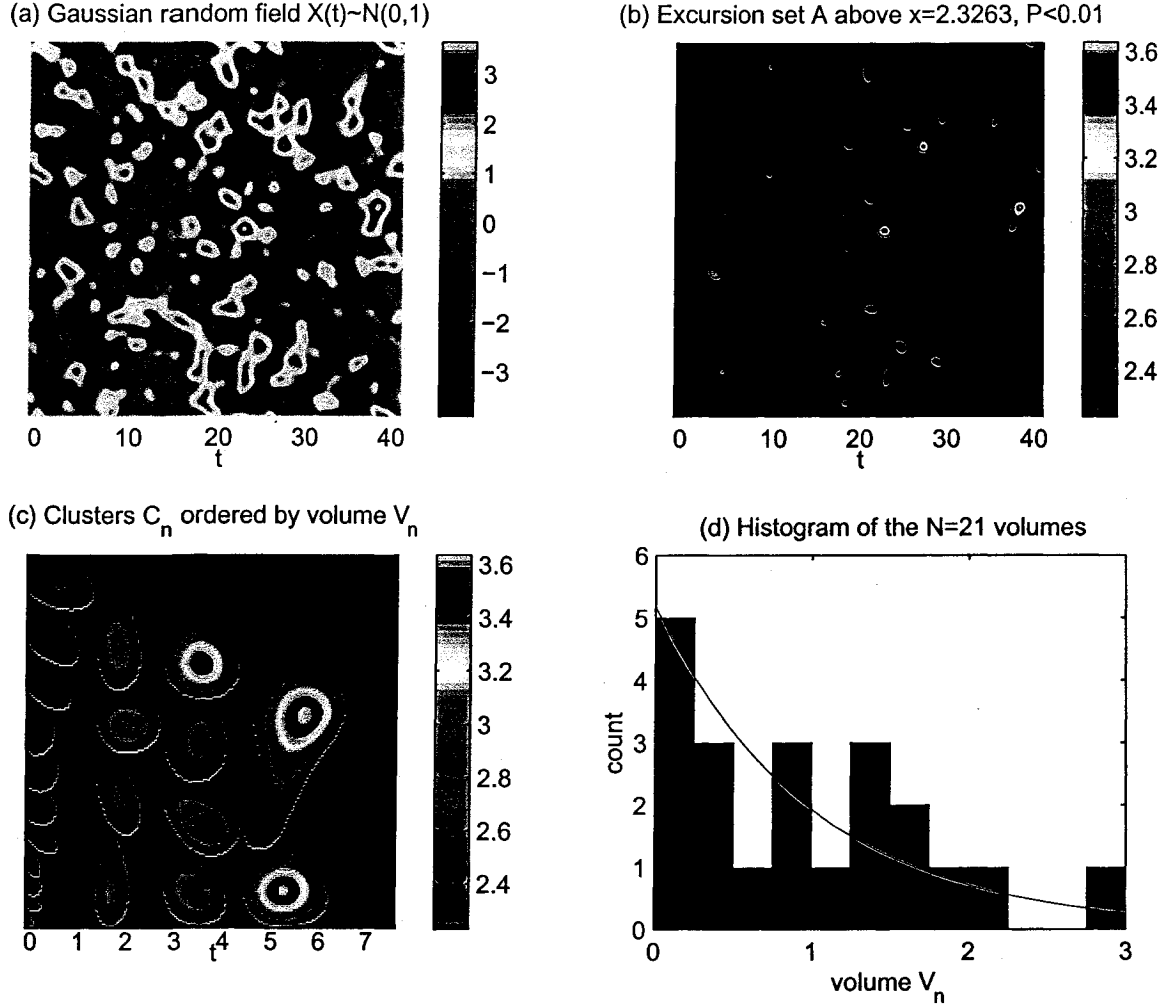


Figure 1.2: Example of clusters in $D = 2$ dimensions. (a) Gaussian random field, $X(t) \sim N(0,1)$ at each point, with $V\{\dot{X}\} = 1$ in each direction. (b) The excursion set A above threshold $x = 2.3263$, chosen so that $\mathbb{P}\{X(t) \geq x\} = 0.01$. (c) Clusters (connected components) C_n of the excursion set, ordered by their volume (area) V_n . To avoid the boundary, only those clusters whose centers are within 5% of the boundary in (b) are shown. (d) Histogram of the cluster volumes, together with the asymptotic density from Friston et al. (1994) (curved line).

1.3 Conjunctions

Often the experiment is repeated on several different subjects or for several different stimuli on the same subject, and the researcher is interested in the common points in the brain where ‘activation’ occurs in all test statistic images. The conjunction is thus defined as those points in the brain that show ‘activation’ in all subjects. As before, we are interested in parts of the conjunction which are noise, and those which show true activation in all test statistic images.

The simplest case is where we have two test statistic images $X_1(\mathbf{t})$ and $X_2(\mathbf{t})$ with excursion sets A_1 and A_2 above a common threshold x . This is illustrated in Figure 1.3 for two runs of the same subject performing the same pain perception experiment. The conjunction A_* is then

$$A_* = A_1 \cap A_2.$$

Another way of looking at this is that A_* is just the excursion set of the minimum $X_*(\mathbf{t})$ of the two random fields $X_1(\mathbf{t})$ and $X_2(\mathbf{t})$:

$$\begin{aligned} X_*(\mathbf{t}) &= \min\{X_1(\mathbf{t}), X_2(\mathbf{t})\}, \\ A_* &= \{\mathbf{t} \in S : X_*(\mathbf{t}) \geq x\} \end{aligned}$$

(see Figure 1.4). Making inference using conjunctions is therefore equivalent to making inference using the minimum of the two random fields. For example, if the component random fields are independent and identically distributed then

$$\mathbb{P}\{X_*(\mathbf{t}) \geq x\} = \mathbb{P}\{X_1(\mathbf{t}) \geq x\}^2.$$

Extensions to more than two random fields are obvious.

1.4 Conjunction cluster volume

To detect activation using the minimum random field, the first method, based on its value, has been solved by Worsley & Friston (2000) for an arbitrary number of independent random fields, and by Taylor (2001) for two correlated Gaussian random fields (see Figure 1.3(d)).

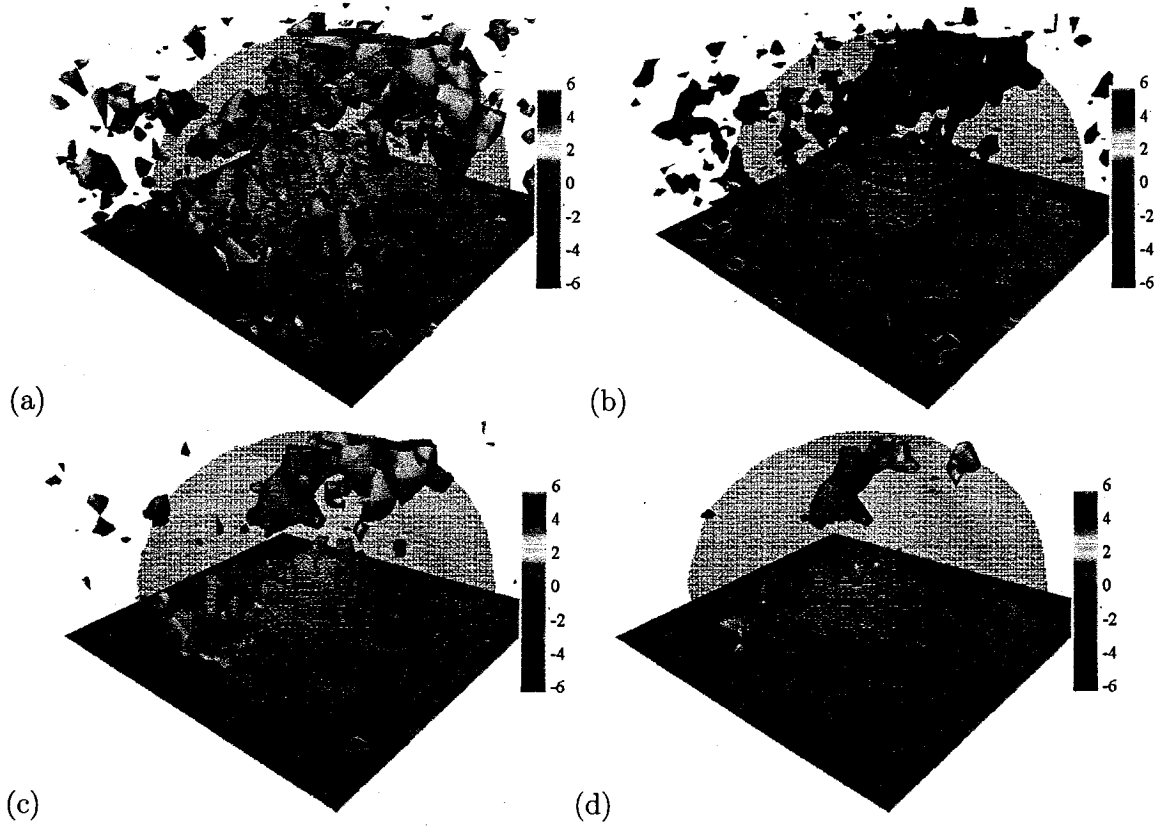


Figure 1.3: Application to conjunctions of pain perception. (a) The first run of the pain perception data (same as in Figure 1.1) threshold at $x = 1.88$. (b) The second run at the same threshold. (c) The conjunction, the intersection of the excursion sets in (a) and (b). The threshold was chosen so that the P-value of the conjunction at any point is 0.001. (d) The conjunction threshold at $x = 3.06$ chosen so that the P-value of the maximum of the conjunction is 0.05. The aim of this thesis is to find a threshold for the volume of the conjunction clusters in (c), analogous to that in Figure 1.1(d), illustrated in Figure 5.1.

However the second method, based on volume of clusters, is so far unsolved. This will be the subject of this thesis.

As we can see above, the key problem is to find the distribution of a single cluster of the excursion set. For the case of a single random field, Nosko (1969) made use of the fact that clusters are roughly circular in shape (see Figure 1.2(c)), so it was only necessary to find an approximate distribution for the radius. This in turn depends only on the square root of the height of the central peak above the threshold x , which itself can be found from the distribution of the maximum of the random field. The resulting theoretical cluster volume density is added to Figure 1.2(d).

However the case of two conjunctions is quite different. As we can see from Figure 1.5, clusters of the conjunction of two random fields are more elliptical in shape. We shall attack this problem in Chapter 3 by approximating the clusters as the intersection of two discs with random radii, themselves approximated by the Nosko method. Things become more complicated for the conjunction of three random fields (see Figure 1.5). Here the clusters are more triangular in shape, and sometimes concave. For more conjunctions the shapes of the clusters become much more erratic, and the volume distribution becomes more highly skewed. For the conjunction of 10 random fields, clusters have highly irregular shapes with a very large number of very small clusters.

Thus the aim of this thesis is to find an accurate approximation to the distribution of volumes of cluster conjunctions, as simulated in Figure 1.6. It seems that the Nosko method, or in fact any method that is based on modelling cluster shape, will be extremely difficult to apply. For these reasons, this thesis will be concerned only with the conjunction of two random fields (as in the top left panel of Figure 1.6), for which we will find reasonably accurate results.

1.5 Outline of the thesis

Let $X_1(t)$, $X_2(t)$, $t \in S$, be two independent, stationary, random fields. Define another random field $X_*(t)$ as follows: $X_*(t) = \min\{X_1(t), X_2(t)\}$, $t \in S$. Let A_* be the excursion

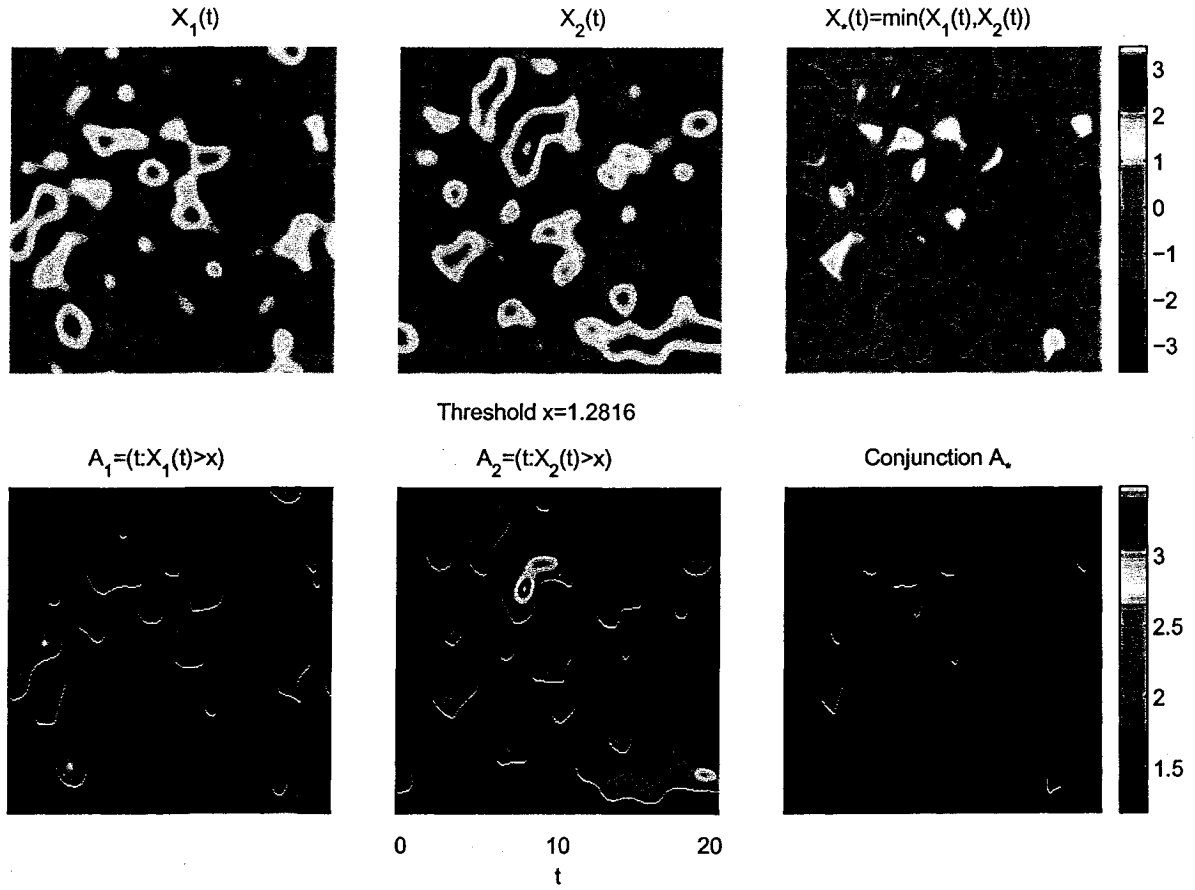


Figure 1.4: Example of the conjunction of two random fields $X_1(t)$ and $X_2(t)$ in $D = 2$ dimensions, $X_j(t) \sim N(0, 1)$ at each point, with $\mathbb{V}\{\dot{X}_j\} = 1$ in each direction, $j = 1, 2$. The conjunction is the intersection of the excursion sets of each field, or equivalently, the excursion set of $X_*(t) = \min\{X_1(t), X_2(t)\}$. The threshold $x = 1.2816$ is chosen so that $\mathbb{P}\{X_*(t) \geq x\} = 0.01$.

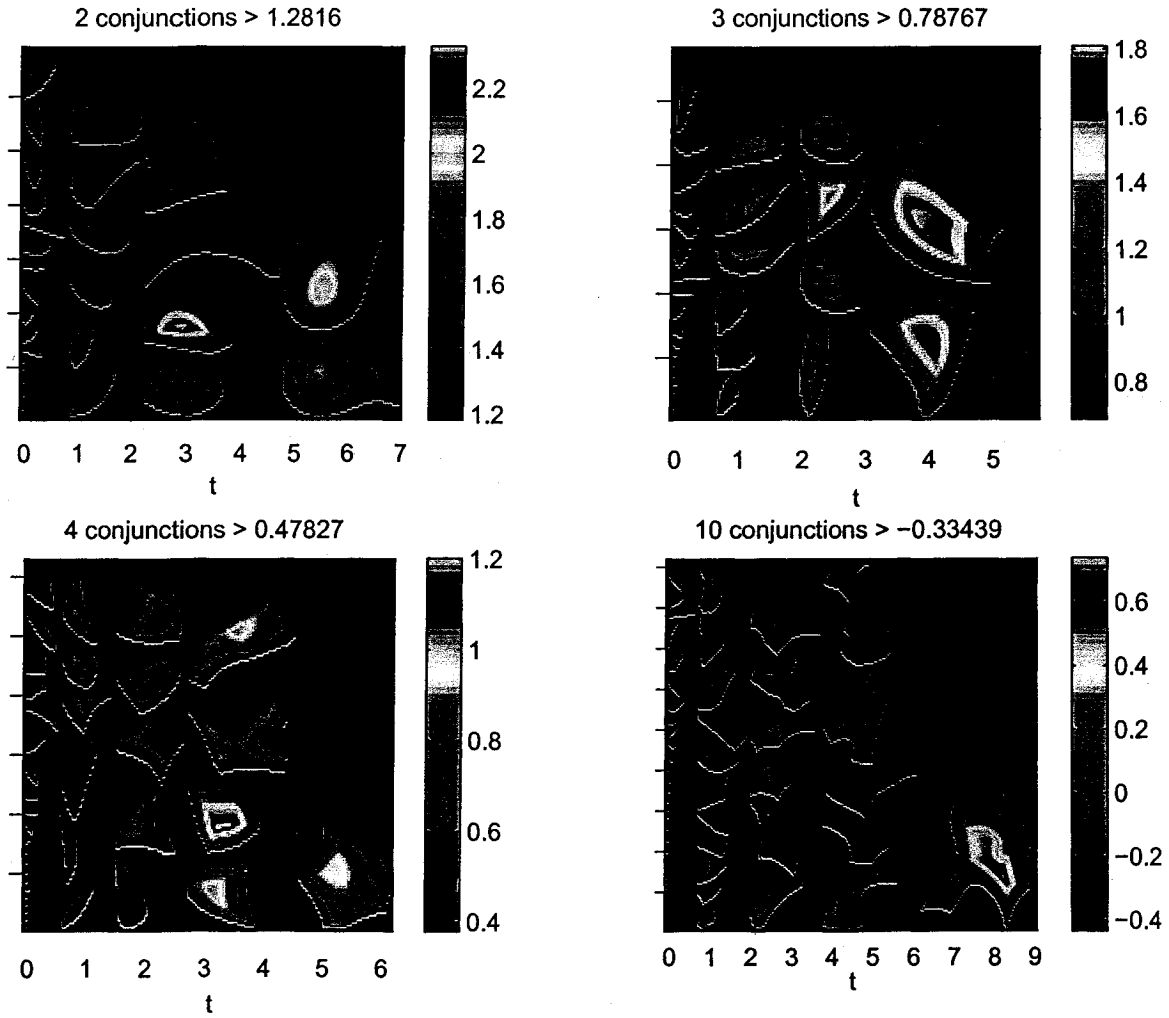


Figure 1.5: Example of conjunction clusters in $D = 2$ dimensions, ordered by cluster volume. The random fields have the same distributions as that in Figure 1.2. The threshold is chosen so that probability of a conjunction at a point is 0.01. Note that the clusters become much more erratic in shape as the number of conjunctions increases.

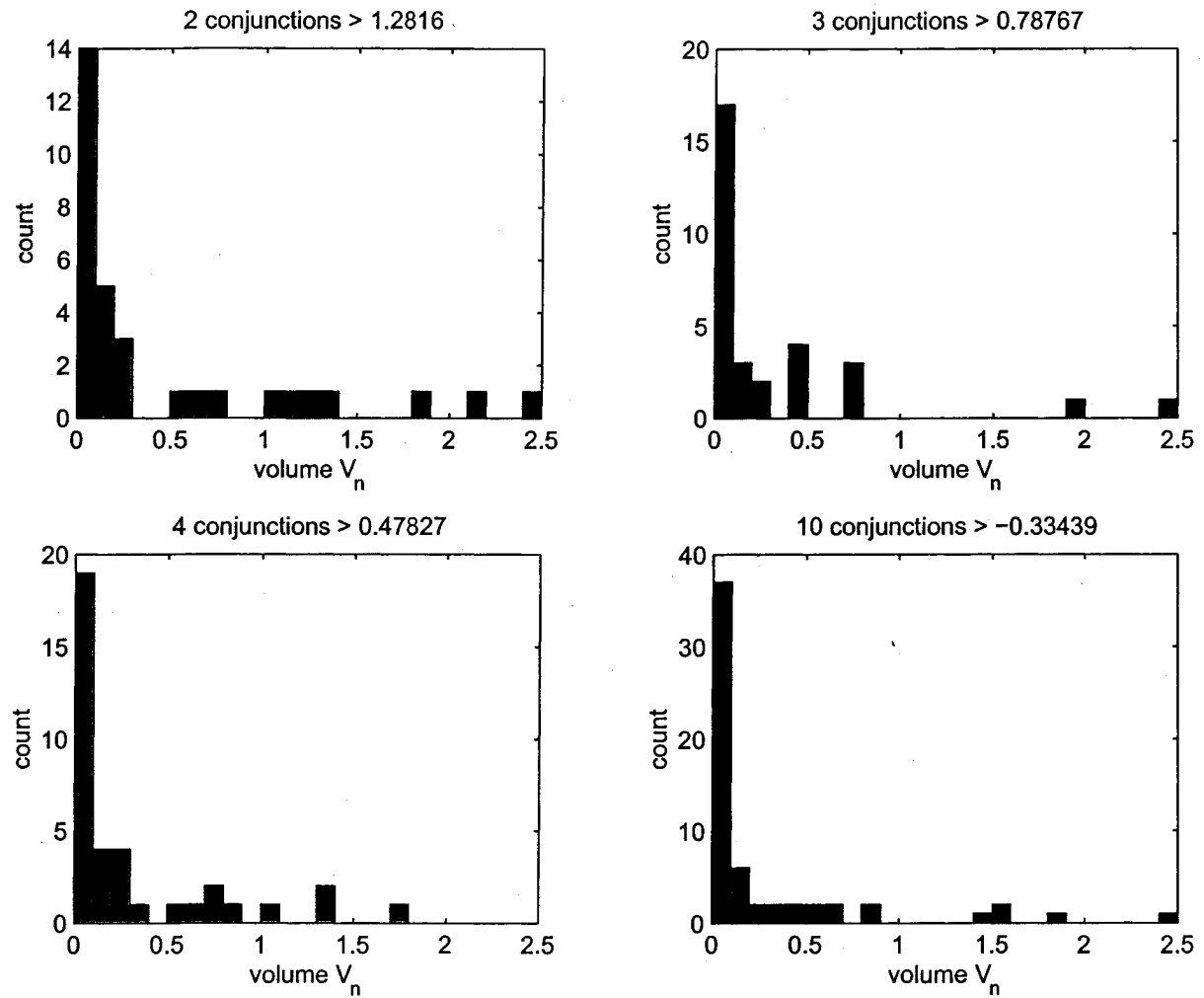


Figure 1.6: Histograms of the conjunction cluster volumes in Figure 1.5 . The aim of this thesis is to find a theoretical distribution for these histograms, but we shall only be successful in the first case of 2 conjunctions.

set of X_* when the level is x , i.e.,

$$A_* = \{t \in S : X_*(t) \geq x\}.$$

A_* will be composed of disjoint clusters $C_1, \dots, C_N$. Consider one of these clusters C_1 , say and let V_1 be its volume. Then, our main interest in Chapter 3 of this thesis will be in approximating the probability distribution of the random variable V_1 when the random fields X_1, X_2 are both Gaussian or both χ^2 random fields. Finally we will approximate the distribution of $\max_n V_n$, the largest cluster.

First, Theorems 2.4.1 and 2.4.2 are used to approximate the shape of the clusters of the excursion set of the Gaussian field. The volume of one cluster of the conjunction according to these theorems will be the volume of the common overlap between two balls with random radius and random centers. In Chapter 3 we will use the Fundamental Kinematic Formula to find an approximate mean value of the volume of one cluster of the excursion set of the conjunction. The joint distribution of the radii and the center is calculated in a closed form. This joint distribution is used to simulate random observation from the distribution of the cluster volume of the conjunction. The same work is also repeated for the χ^2 random field but with Theorems 2.4.3, 2.4.4 and 2.4.5. We also compared our results on a special case, the cosine Gaussian field.

Since the cluster volume has a complicated form it is not easy to find its probability distribution in closed form so in Chapter 4 we will do a simulation to check the validity of the theory developed in Chapter 3. We did the simulation only for two dimensional Gaussian fields since the complexity of the computation becomes high as the dimension increases.

In Chapter 5 we will apply the theory developed in Chapter 3 to real data taken from two fMRI images. In Chapter 6 we will present our conclusions.

Chapter 2

Random field theory

In this chapter we will give a brief introduction to random field theory that we need to solve the problem of conjunctions. This introduction will include some definitions as well as some important results in random field theory. Most of the material in this chapter are based on Adler (1981). We will also recal some mathematical tools from integral geometry which will be used in the sequel. We will assume that all probabilistic concepts from now until the end of this thesis are defined in a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

2.1 Random fields

We are interested in real valued random fields. The random field $X(\mathbf{t}), \mathbf{t} \in S \subset \mathbb{R}^D$ is a collection of random variables $X(\mathbf{t}), \mathbf{t} \in S$ together with a collection of measures or distribution functions of the form $F_{\mathbf{t}_1, \dots, \mathbf{t}_n}$ on $\mathcal{B}(\mathbb{R}^D)$, the Borel sigma field on $\mathbb{R}^D$, for $n = 1, 2, \dots, \mathbf{t}_i \in \mathbb{R}^D$ such that

$$F_{\mathbf{t}_1, \dots, \mathbf{t}_n}(B) = \mathbb{P}\{(X(\mathbf{t}_1), \dots, X(\mathbf{t}_n)) \in B\}$$

for every $B \in \mathcal{B}(\mathbb{R}^D)$. For a given $\omega \in \Omega$, $X(\mathbf{t}, \omega)$ is a deterministic real valued function on $\mathbb{R}^D$ which is a realization of the field $X(\mathbf{t})$. The set $\{(\mathbf{t}, X(\mathbf{t})) : \mathbf{t} \in \mathbb{R}^D\}$ is called the *sample function* or *sample path* of X .

The random field X is said to be *strictly homogenous* or *stationary* if for any k , any set of real numbers $x_1, \dots, x_k$ and any $(k + 1)$ points $\tau, \mathbf{t}_1, \dots, \mathbf{t}_k$ in $\mathbb{R}^D$ the following condition

on its finite-dimensional distribution holds

$$\mathbb{P}\{X(\mathbf{t}_1) \leq x_1, \dots, X(\mathbf{t}_k) \leq x_k\} = \mathbb{P}\{X(\mathbf{t}_1 + \tau) \leq x_1, \dots, X(\mathbf{t}_k + \tau) \leq x_k\}.$$

This means that the random field is invariant under translation of the parameter space.

For every random field $X(\mathbf{t})$ we can define two functions, the mean and the covariance functions. The mean function is

$$\mu(\mathbf{t}) = \mathbb{E}\{X(\mathbf{t})\}, \quad \mathbf{t} \in S,$$

and the covariance function is

$$\mathcal{R}(\mathbf{s}, \mathbf{t}) = \mathbb{E}\{(X(\mathbf{s}) - \mu(\mathbf{s}))(X(\mathbf{t}) - \mu(\mathbf{t}))\}, \quad \mathbf{s}, \mathbf{t} \in S.$$

A Gaussian field $X(\mathbf{t})$ with covariance function $\mathcal{R}$ is *isotropic* if its covariance function depends only on $\|\mathbf{t} - \mathbf{s}\|$, i.e. if $\mathcal{R}(\mathbf{s}, \mathbf{t}) = \mathcal{R}(\|\mathbf{t} - \mathbf{s}\|)$ where $\|\mathbf{t}\| = \sqrt{\mathbf{t}\mathbf{t}^T}$.

In this thesis we will consider two types of random fields: the Gaussian random field and the χ^2 random field. The χ^2 is derived from the Gaussian field. A random field $X(\mathbf{t}), \mathbf{t} \in S$, is said to be a Gaussian field over S if its finite dimensional distribution is multivariate Gaussian. To define the χ^2 random field, let $X_1(\mathbf{t}), \dots, X_\nu(\mathbf{t}), \mathbf{t} \in \mathbb{R}^D$, be independent zero mean, unit variance and stationary Gaussian random fields. Then Adler (1981), page 169, defines the χ^2 field as follows

$$U(\mathbf{t}) = \sum_{i=1}^{\nu} X_i(\mathbf{t})^2, \quad \mathbf{t} \in \mathbb{R}^D.$$

Note that for every $\mathbf{t} \in \mathbb{R}^D$, $U(\mathbf{t})$ is a χ^2 random variable with ν degrees of freedom.

In testing for brain functional and structural changes, the regions in the brain where the random field is above a high level are of main interest since these regions are related to high changes. The set of points in the brain related to high changes is estimated by the excursion set of the random field. The definition of the excursion set of a random field is given as follows:

Definition: Excursion set. Let $X(\mathbf{t}) : \mathbb{R}^D \rightarrow \mathbb{R}$ be a random field. For any fixed real number x and any subset S of $\mathbb{R}^D$ we define the excursion set of the field X above the level x in S to be the set

$$A = \{\mathbf{t} \in S : X(\mathbf{t}) \geq x\}.$$

The Lebesgue measure of the excursion set is

$$\mu_D(A) = \int_S 1_A dt.$$

By taking the expectation of the last equation and then changing the order of the integration we get

$$\mathbb{E}\{\mu_D(A)\} = \int_S \mathbb{P}\{X(t) \geq x\} dt.$$

If the field is homogeneous then the integration on the right-hand side is easy to integrate, and we have the following important formula

$$\mathbb{E}\{\mu_D(A)\} = \mu_D(S) \mathbb{P}\{X(0) \geq x\}.$$

2.2 Continuity and differentiability of random fields

In this section we give the definitions of the stochastic version of the real analysis concepts of the limits and derivatives of random functions.

A sequence of random variables $\{X_n\}$ is said to converge to another random variable X in the mean square (m.s.) sense if

$$\mathbb{E}\|X_n - X\|^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

We will denote this limit by $\text{l.i.m.}_{n \rightarrow \infty} X_n = X$. A field $X(t)$ is *continuous in m.s* if

$$\lim_{\|h\| \rightarrow 0} \mathbb{E}\{(X(t+h) - X(t))^2\} = 0.$$

In fact the field $X(t)$ will be continuous in m.s if and only if its autocorrelation function is continuous. Also if $X(t)$ is continuous in m.s. then its mean function is continuous. A random field $X(t)$ is said to have a m.s partial derivative in the i^{th} direction and is denoted by $\dot{X}_i(t)$ if

$$\text{l.i.m.}_{h \rightarrow 0} \frac{X(t + h e_i) - X(t)}{h} = \dot{X}_i(t).$$

where e_i is the i^{th} unit vector in the standard basis of $\mathbb{R}^D$. We will denote $\dot{X}(t)$ to be the vector of the m.s. first order partial derivatives for the field $X(t)$ and $\ddot{X}(t)$ to be the matrix of m.s. second order partial derivatives of $X(t)$.

The random field is *almost surely continuous* at $\mathbf{t}^*$ if for every sequence $\mathbf{t}_n$ for which $\|\mathbf{t}_n - \mathbf{t}^*\| \rightarrow 0$ as $n \rightarrow \infty$ we have $X(\mathbf{t}_n) \xrightarrow{a.s.} X(\mathbf{t}^*)$. X is almost surely continuous on $A \subseteq \mathbb{R}^D$ if it is almost surely continuous at every point in A . This type of continuity is called *sample function* or *sample path continuity*. Also *almost sure differentiability* can be defined in the same fashion.

The *moduli of continuity* of X and its first and second order partial derivatives are defined by

$$\begin{aligned}\xi(h) &= \sup_{\|\mathbf{s}-\mathbf{t}\|<h} |X(\mathbf{t}) - X(\mathbf{s})|, \\ \xi_j(h) &= \sup_{\|\mathbf{s}-\mathbf{t}\|<h} |\dot{X}_j(\mathbf{t}) - \dot{X}_j(\mathbf{s})|, \\ \xi_{ij}(h) &= \sup_{\|\mathbf{s}-\mathbf{t}\|<h} |\ddot{X}_{ij}(\mathbf{t}) - \ddot{X}_{ij}(\mathbf{s})|.\end{aligned}$$

We will assume that all Gaussian random fields used in this thesis in addition to the Gaussian random fields used to define the χ^2 random field will satisfy the following conditions. The random field has almost surely continuous partial derivatives up to second order with finite variances in an open neighborhood of S and the joint distribution of the random field and these partial derivatives is non-degenerate. Assume also that the moduli of continuity of $\ddot{X}_{ij}$ satisfies the following condition

$$\mathbb{P}\{\max_{i,j}\{\xi_i(h), \xi_{ij}(h)\} > \epsilon\} = o(h^D) \text{ as } h \downarrow 0.$$

When we are interested in the conjunction of the χ^2 random field we will assume that its moduli of continuity satisfy the same conditions.

Let $\Lambda = \mathbb{V}\{\dot{X}\}$ be the $D \times D$ variance-covariance matrix of the partial derivatives of $X(\mathbf{t})$ with elements $\lambda_{kl} = \text{Cov}(\dot{X}_k, \dot{X}_l)$, $k, l = 1, 2, \dots, D$. The following theorem is proved in Adler (1981), page 114:

Theorem 2.2.1. *Let $X(\mathbf{t})$ be a stationary Gaussian random field over $\mathbb{R}^D$. Then*

- (a) $\dot{X} \sim \text{Normal}_N(\mathbf{0}, \Lambda)$ and is independent of $X, \ddot{X}$,
- (b) conditional on X , $\ddot{X}|X \sim \text{Normal}(-X\Lambda, \mathbf{M}(\Lambda))$ and the elements of $\mathbf{M}(\Lambda)$ are such that $\text{Cov}(\ddot{X}_{ij}, \ddot{X}_{kl}|X) = \epsilon(i, j, k, l) - \lambda_{ij}\lambda_{kl}$ where the function $\epsilon(i, j, k, l)$ is symmetric in its arguments.

2.3 Horizontal window conditioning

To study the behavior of a random field near local maxima we need to condition on the event that the random field has a local maximum at some point in its parameter space. Since this event has zero probability we need to define the *horizontal window conditioning* (HW), denoted by $\parallel$. HW conditioning has been introduced by Kac and Slepian (1959). To define the HW conditioning let $X(\mathbf{t})$ be a random field on $\mathbb{R}^D$. Let us define the new field $X_x(\mathbf{t}) = X(\mathbf{t}) \parallel \mathcal{A}$, where $\mathcal{A}$ is the event that $X(\mathbf{t})$ has a local maximum with height x at $\mathbf{t} = \mathbf{0}$, by approximating the event $\mathcal{A}$ by a sequence of events $\mathcal{A}(h, h')$, where

$$\begin{aligned} \mathcal{A}(h, h') = & \{X(\mathbf{t}) \text{ has a local maximum of height } x \text{ in } (x, x+h) \\ & \text{at some point in the ball } \|\mathbf{t}\| < h'\}. \end{aligned}$$

The distribution of $X_x(\mathbf{t})$ is given by

$$\mathbb{P}\{X_x(\mathbf{t}) \in B\} = \lim_{h \rightarrow 0} \lim_{h' \rightarrow 0} \mathbb{P}\{X(\mathbf{t}) \in B \mid \mathcal{A}(h, h')\},$$

where B is a Borel set. If the field $X(\mathbf{t})$ is ergodic then we can write the right hand side of the last equation as a ratio of two expectations. For more information about how to do this see page 150 of Adler (1981).

2.4 Random field near local maxima

In this section we will report five theorems describing the behavior of a random field near local maxima. The first two are for the Gaussian random field and the others are for χ^2 .

The following two theorems will be very useful to solve the conjunction problem when the underlying field is Gaussian. These two theorems are due to Nosko (1969) and are reported by Adler (1981), Section 6.8. Let X be a real valued, stationary, zero mean, unit variance, ergodic, Gaussian random field satisfying the conditions of Section 6.6 of Adler (1981).

Theorem 2.4.1. *Approximation of a random field over a cluster: Given that the random field $X(\mathbf{t})$ takes the value x at $\mathbf{t} = \tau$, then with probability approaching one as*

$x \rightarrow \infty$, the field has the following representation over that component of the excursion set containing τ :

$$X(t) = x + \dot{X}(\tau)(t - \tau)^T - \frac{1}{2}x(t - \tau)\Lambda(t - \tau)^T + o(1/x), \quad (2.4.1)$$

where $\Lambda = \mathbb{V}\{\dot{X}\}$.

According to this theorem we can say that the random field $X(t)$ near high local maxima has a deterministic shape which is quadratic.

Theorem 2.4.2. *Approximate distribution of the height of local maxima of a stationary Gaussian random field: Given that the field $X(t)$ satisfying the conditions above has a local maximum at $t = 0$ with height exceeding x , the conditional distribution of $m_x = X(0) - x$ (i.e, the excess height above the level x) is given by*

$$\lim_{x \rightarrow \infty} \mathbb{P}\{xm_x > v | m_x > 0\} = \exp(-v). \quad (2.4.2)$$

The following three theorems will be used to solve the problem of conjunctions for the χ^2 random fields. These theorems are due to Cao (1999). We will assume that we are dealing with D -dimensional, real-valued, stationary, ergodic, Gaussian random fields X with zero mean, unit variance, and $\Lambda = \mathbb{V}\{\dot{X}\}$.

Theorem 2.4.3. *Given that a χ^2 field $U(t)$, $t \in \mathbb{R}^D$, has a local maximum at 0 with height x , then $x^{-1}\ddot{U}(t) \rightarrow -2\Lambda$ as $x \rightarrow \infty$.*

This means that the curvature of the χ^2 random field is deterministic near its local maxima.

Theorem 2.4.4. *Given that a χ^2 field $U(t)$, $t \in \mathbb{R}^D$ has a local maximum at 0 with height $U = U(0)$ exceeding x , then*

$$\lim_{x \rightarrow \infty} \mathbb{P}\{U > x + v | U > x\} = e^{-v/2}$$

for $v > 0$.

Theorem 2.4.5. *Assume that 0 is a local maximum of a χ^2 field $U(t)$ with height x . Then as $x \rightarrow \infty$, $U_x(t/\sqrt{x}) - x$ converges uniformly to the elliptic paraboloid $-t^T \Lambda t$ in the neighborhood of $t = 0$.*

This theorem also says that the χ^2 has a deterministic form near its local maxima which is an elliptic paraboloid.

2.5 Integral geometry

In this section we describe some tools and results from integral geometry which will be used in the next chapter. One of these tools is the Euler characteristic of a set. A very important result is the Fundamental Kinematic Formula. Let $\mathcal{K}^D$ denote the collection of all compact convex subsets of $\mathbb{R}^D$. A finite union of compact convex sets will be called a *polyconvex* set. Also we shall say a polyconvex set K in $\mathbb{R}^D$ is of dimension D if it is not contained in a finite union of hyperplanes of $\mathbb{R}^D$. The union and intersection of polyconvex sets is also polyconvex. In other words, the family of polyconvex sets in $\mathbb{R}^D$ is a distributive lattice and we will denote it by $PolyConv(D)$.

The *Euler characteristic* function (*EC* or χ) is an additive functional on $\mathcal{K}^D$ such that for any $K \in \mathcal{K}^D$, $K = \cup_{i=1}^m K_i$, K_i a compact convex set,

$$\chi(K) = \sum_i \chi(K_i) - \sum_{i < j} \chi(K_i \cap K_j) + \dots + (-1)^{m+1} \chi(K_1 \cap K_2 \cap K_3 \cap \dots \cap K_m).$$

and where for a compact convex set K

$$\chi(K) = \begin{cases} 1 & \text{if } K \neq \phi, \\ 0 & \text{if } K = \phi. \end{cases}$$

The Euler characteristic describes the set K in a purely topological way, without reference to any kind of metric. For $D=2$, χ equals the number of connected components minus the number of holes while for $D=3$ χ equals to the number of connected components minus the number of tunnels plus the number of cavities.

Since the excursion set above a high level decomposes into disjoint convex sets (Adler (1981) page 138), the Euler characteristic of this set will be a good approximation to the number of convex components.

Let C be a subset of $\mathbb{R}^D$. Let $a_i = 2\pi^{i/2}/\Gamma(i/2)$ be the surface area of a unit $(i-1)$ -dimensional sphere in $\mathbb{R}^i$. Let M be the inside curvature matrix of ∂C at a point $\mathbf{t}$ and

let $\text{detr}_i(M)$ be the sum of the determinants of all $i \times i$ principal minors of M for $i = 0, 1, 2, \dots, D-1$. We define $\mu_i(C)$, the i -dimensional *Minkowski functional* or *intrinsic volume* of C , to be

$$\mu_i(C) = \frac{1}{a_{D-i}} \int_{\partial C} \text{detr}_{D-1-i}(M) dt, \quad i = 0, 1, \dots, D-1,$$

and $\mu_D(C)$ to be the Lebesgue measure of C . In this thesis we will be interested in the case when C is the D -dimensional ball $B(r)$ with radius r centered at the origin. The value of μ_i for general D and $C = B(r)$ is

$$\mu_i(B(r)) = \binom{D}{i} \frac{\omega_D}{\omega_{D-i}} r^i, \quad (2.5.3)$$

where $\omega_i = \pi^{i/2}/\Gamma(i/2 + 1)$ is the volume of the i -dimensional unit ball.

The following theorem is the *Kinematic Fundamental Formula*, well known in integral geometry, and it is useful in many areas where we are interested in problems of rigid random motion of convex sets or bodies.

Theorem 2.5.1. *Kinematic Fundamental Formula:* *For all $A, K \in \text{PolyConv}(D)$ and for $0 \leq k \leq D$ we have*

$$\int_{E_D} \mu_k(A \cap gK) dg = \sum_{i=0}^{D-k} \begin{bmatrix} i+k \\ k \end{bmatrix} \begin{bmatrix} D \\ i \end{bmatrix}^{-1} \mu_{i+k}(A) \mu_{D-i}(K),$$

where $\mu_i(A)$ is the i^{th} Minkowski functional of the polyconvex set A and E_D is the set of all rigid motions in $\mathbb{R}^D$ and $g \in E_D$. The factor $\begin{bmatrix} i \\ j \end{bmatrix}$ is given by the following formula

$$\begin{bmatrix} i \\ j \end{bmatrix} = \binom{i}{j} \frac{\omega_i}{\omega_j \omega_{i-j}}$$

and ω_i is the volume of the i -dimensional unit ball.

As an application of this theorem, if A and K are convex, the ratio

$$\frac{\int_{E_D} \mu_k(A \cap gK) dg}{\int_{E_D} \mu_0(A \cap gK) dg}. \quad (2.5.4)$$

is the mean value of the k^{th} intrinsic volume of $A \cap gK$, taken over all gK in $\mathbb{R}^D$ congruent to K that meet A . The reason for this is that $\mu_0(A \cap gK)$ takes the value 1 if A and gK intersect, and zero otherwise.

To approximate the value of $\mathbb{P}\{\sup_{t \in S} X(t) \geq x\}$ for a random field $X(t)$, $t \in S$, we need to define the *EC intensities* of an isotropic random field $X(t)$. Let $\dot{X}_{|i}$ be the vector of the first i components of $\dot{X}$, and $\ddot{X}_{|i}$ be the $i \times i$ matrix of the first i rows and columns of $\ddot{X}$. Then the i -dimensional EC *intensity* of $X(t)$ is defined by

$$\rho_i(x) = \mathbb{E}\{\mathbf{1}_{\{X \geq x\}} \det(-\ddot{X}_{|i}) | \dot{X}_{|i} = 0\} p_{|i}(0),$$

where $p_{|i}(\cdot)$ is the density of $\dot{X}_{|i}$. Then Worsley (1995) shows that

$$\mathbb{P}\{\sup_{t \in S} X(t) \geq x\} \approx \mathbb{E}\{\chi(A)\} = \sum_{i=0}^D \mu_i(S) \rho_i(x),$$

where $A = \{t \in S : X(t) \geq x\}$.

We are interested in the values $\rho_i(x)$ for $i = 0, 1, 2, 3$ for two different types of random fields: Gaussian, and χ^2 with ν degrees of freedom. Since the random fields are isotropic, let $\Lambda = \lambda \mathbf{I}_D$, where $\mathbf{I}_D$ is the $D \times D$ identity matrix, be the roughness matrix of all the Gaussian random fields, so that λ is the roughness parameter. From Cao and Worsley (1999) we report the following. For the Gaussian field,

$$\begin{aligned} \rho_0(x) &= \int_x^\infty \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy \\ \rho_1(x) &= \sqrt{\lambda} \frac{e^{-x^2/2}}{2\pi} \\ \rho_2(x) &= \lambda \frac{x e^{-x^2/2}}{(2\pi)^{3/2}} \\ \rho_3(x) &= \lambda^{3/2} \frac{e^{-x^2/2} (x^2 - 1)}{(2\pi)^2}. \end{aligned}$$

and for the χ^2 random field with ν degrees of freedom

$$\begin{aligned}\rho_0(x) &= \int_x^\infty \frac{y^{\nu/2-1} e^{-y/2}}{2^{\nu/2} \Gamma(\nu/2)} dy \\ \rho_1(x) &= \sqrt{\lambda} \frac{x^{(\nu-1)/2} e^{-x/2}}{\sqrt{2\pi} 2^{(\nu-2)/2} \Gamma(\nu/2)} \\ \rho_2(x) &= \frac{\lambda}{2\pi} \frac{x^{\nu/2-1} e^{-x/2} (x - (\nu - 1))}{2^{\nu/2-1} \Gamma(\nu/2)} \\ \rho_3(x) &= \frac{\lambda^{3/2}}{(2\pi)^{3/2}} \frac{x^{(\nu-3)/2} e^{-x/2} (x^2 - (2\nu - 1)x + (\nu - 1)(\nu - 2))}{2^{\nu/2-1} \Gamma(\nu/2)}.\end{aligned}$$

2.6 EC densities of the conjunction

Worsley and Friston (2000) gave a method to approximate the mean value of the Euler characteristic of the excursion set of the field $X_*(t) = \min_{i=1}^n \{X_i(t)\}$ where $X_i(t)$'s are independent, isotropic, random fields. We will describe this method in this section. Let $b_i = \Gamma((i+1)/2)/\Gamma(1/2)$. Let ρ_{ik} be the EC intensity of $X_k(t)$ in $\mathfrak{R}^i$, $1 \leq k \leq D$ and $B \subseteq \mathfrak{R}^D$. Define the upper triangular Toeplitz matrix $\mathbf{M}_k$ and the vector $\mu(B)$ to be

$$\mathbf{M}_k = \begin{pmatrix} \rho_{0k}/b_0 & \rho_{1k}/b_1 & \dots & \rho_{Dk}/b_D \\ 0 & \rho_{0k}/b_0 & \dots & \rho_{(D-1)k}/b_{D-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho_{0k}/b_0 \end{pmatrix}, \quad \mu(B) = \begin{pmatrix} \mu_0(B)b_0 \\ \mu_1(B)b_1 \\ \vdots \\ \mu_D(B)b_D \end{pmatrix}.$$

Then the following relation holds

$$\mathbb{E}\{\mu(A_*)\} = \left(\prod_{k=1}^n \mathbf{M}_k \right) \mu(S),$$

where

$$A_* = \{t \in S : X_*(t) \geq x\}.$$

The last formula gives us the mean vector of Minkowski functionals of the excursion set of X_* . To find the approximate value of $\mathbb{P}\{\sup_{t \in S} X_*(t) \geq x\}$ we need only to find the mean value of the EC of the excursion set of the random field X_* which is the first component of the vector $\mathbb{E}\{\mu(A_*)\}$. So we have the following formula

$$\mathbb{E}\{\mu_0(A_*)\} = (1, 0, \dots, 0) \left(\prod_{k=1}^n \mathbf{M}_k \right) \mu(S). \quad (2.6.5)$$

Let N be the number of clusters in S above high level x . Since the number of clusters above high level x can be approximated by $\mu_0(A_*)$ we can use (2.6.5) to approximate $\mathbb{E}(N)$.

2.7 Poisson clumping heuristic

The *mosaic process* is a formalization of throwing sets down at random. Let $\mathcal{B}$ be a collection of sets in $\mathbb{R}^D$. Let $\mathcal{C}$ be a probability distribution over $\mathcal{B}$. Let $\mathbf{a}_i \in \mathbb{R}^D$ be the events of a Poisson process with rate ψ . We define the mosaic process as

$$A = \cup_i (\mathbf{a}_i \oplus B_i),$$

where $B_i \in \mathcal{B}$ are chosen randomly according to $\mathcal{C}$ and $\oplus$ denotes the Minkowski sum. The $\mathbf{a}_i$ are called the centers and $\mathbf{a}_i \oplus B_i$ are called the clumps.

According to Theorem 6.9.3 of Adler (1981), the local maxima of a Gaussian random field above sufficiently high level x occur randomly according to a Poisson process with a spatial rate γ_x . For high levels each cluster will contain one local maximum, therefore the random number N of clusters above level x in a set S tends towards a Poisson random variable with probability function

$$\mathbb{P}\{N = n\} \approx \frac{(\mu_D(S)\gamma_x)^n e^{-\mu_D(S)\gamma_x}}{n!}, \quad n = 0, 1, 2, \dots,$$

as $x \rightarrow \infty$, where γ_x is given by

$$\gamma_x = \frac{|\Lambda|^{1/2} x^{D-1} e^{-x^2/2}}{(2\pi)^{\frac{D+1}{2}}},$$

where Λ is the same as that defined in Theorem 2.2.1.

The clusters of the excursion set of a Gaussian random field for large x can be viewed as clumps that are centered at points of a Poisson point process so we will model the conjunction problem by a mosaic process. In the next chapter we will use the theory in Sections 2.6 and 2.7 to approximate the mean value of the volume of one cluster of the field X_* . Also this will enable us to find an approximation of the distribution of the largest volume of the clusters of the excursion set of X_* above high level x .

Chapter 3

Volumes of clusters of conjunctions

In this Chapter we will consider the problem of finding an approximation to the probability distribution of the volume of one cluster of the excursion set of the conjunction of two independent isotropic random fields satisfying suitable conditions. We will develop the theory only for two types of random fields: the Gaussian and the χ^2 random fields. These two fields are the only two fields that have a deterministic shape near local maxima. Other types of fields such as T or F fields have random shapes (Cao, 1999). For example, the excursion set of isotropic Gaussian and χ^2 fields near high local maxima are balls, whereas they are random ellipsoids for T and F fields. This will enable us to approximate the probability distribution of the maximum volume of the clusters.

In this Chapter the Gaussian random field will be considered and an approximation to the distribution of the volume of one cluster of the excursion set of the field $\min\{X_1(\mathbf{t}), X_2(\mathbf{t})\}$ will be given for any dimension. The mean value of the volume will be given in a closed form based on this approximation. Also calculations for a special case will be given.

3.1 Distribution of the volume of a cluster

Let $X_1(\mathbf{t})$ and $X_2(\mathbf{t})$, $\mathbf{t} \in S$, be two independent smooth stationary Gaussian random fields with $\mathbb{V}\{\dot{X}_1(\mathbf{t})\} = \mathbb{V}\{\dot{X}_2(\mathbf{t})\} = \lambda \mathbf{I}_D$ where $\mathbf{I}_D$ is the identity matrix of order D . Let $\mathbf{t}_1, \mathbf{t}_2$ be local maximizers of $X_1(\mathbf{t})$ and $X_2(\mathbf{t})$ respectively. According to Theorem 2.4.1, for high thresholds, the component of the excursion set that contains $\mathbf{t}_j$ is a D -dimensional ball

with center $\mathbf{t}_j$ and radius $R_j = \sqrt{2W_j/x\lambda}$, $j = 1, 2$. According to Theorem 2.4.2 $W_1 = X_1(\mathbf{t}_1) - x$, $W_2 = X_2(\mathbf{t}_2) - x$ are independent exponentially distributed random variables with mean $1/x$. Using the Jacobian method it is easy to show that the density of the radius R_j is

$$f_{R_j}(r_j) = x^2 \lambda r_j e^{-\frac{x^2}{2} \lambda r_j^2} \mathbf{1}_{\{r_j > 0\}}. \quad (3.1.1)$$

The conjunction of the two fields occurs when the two balls overlap. So let $H = \|\mathbf{t}_2 - \mathbf{t}_1\|$ be the distance between the centers of the two balls and let V be the volume of the overlap of the two balls. We are interested in the probability distribution of V given that there is an overlap, that is, the distribution of V given $\mathcal{G} = \{0 \leq H \leq R_1 + R_2\}$. Since the distributions of $\mathbf{t}_1$ and $\mathbf{t}_2$ are uniform over S , then the density function of H given $R_1, R_2, \mathcal{G}$ is

$$f_H(h|R_1 = r_1, R_2 = r_2, \mathcal{G}) = \frac{Dh^{D-1}}{(r_1 + r_2)^D} \mathbf{1}_{\{0 < h < r_1 + r_2\}}. \quad (3.1.2)$$

The joint *cdf* of (R_1, R_2) , given $\mathcal{G}$ can be found as follows:

$$\mathbb{P}\{\mathcal{G}|R_1 = v_1, R_2 = v_2\} = \frac{\omega_D(v_1 + v_2)^D}{\mu_D(S)},$$

$$\begin{aligned} \mathbb{P}\{R_1 \leq r_1, R_2 \leq r_2, \mathcal{G}\} &= \int_0^{r_1} \int_0^{r_2} \mathbb{P}\{\mathcal{G}|R_1 = v_1, R_2 = v_2\} f_{R_1}(v_1) f_{R_2}(v_2) dv_2 dv_1 \\ &= x^4 \lambda^2 \omega_D \mu_D(S)^{-1} \int_0^{r_1} \int_0^{r_2} (v_1 + v_2)^D v_1 v_2 e^{-\frac{x^2}{2} \lambda (v_1^2 + v_2^2)} dv_2 dv_1 \end{aligned}$$

and

$$\mathbb{P}\{\mathcal{G}\} = \kappa^{-1} x^4 \lambda^2 \omega_D \mu_D(S)^{-1},$$

where

$$\begin{aligned} \kappa^{-1} &= \int_0^\infty \int_0^\infty (v_1 + v_2)^D v_1 v_2 e^{-\frac{1}{2} x^2 \lambda (v_1^2 + v_2^2)} dv_2 dv_1 \\ &= \int_0^\infty \int_0^\infty \sum_{i=0}^D \binom{D}{i} v_1^{i+1} v_2^{D+1-i} e^{-\frac{1}{2} x^2 \lambda (v_1^2 + v_2^2)} dv_2 dv_1 \\ &= \sum_{i=0}^D \binom{D}{i} \int_0^\infty v_1^{i+1} e^{-\frac{1}{2} x^2 \lambda v_1^2} dv_1 \int_0^\infty v_1^{D+1-i} e^{-\frac{1}{2} x^2 \lambda v_2^2} dv_2 \\ &= 2^{D/2} (x^2 \lambda)^{-D/2-2} \sum_{i=0}^D \binom{D}{i} \Gamma(i/2 + 1) \Gamma((D-i)/2 + 1). \end{aligned}$$

The joint *cdf* of (R_1, R_2) given $\mathcal{G}$ is

$$\begin{aligned} F_{R_1, R_2}(r_1, r_2 | \mathcal{G}) &= \frac{\mathbb{P}\{R_1 \leq r_1, R_2 \leq r_2, \mathcal{G}\}}{\mathbb{P}\{\mathcal{G}\}} \\ &= \kappa \int_0^{r_1} \int_0^{r_2} (v_1 + v_2)^D v_1 v_2 e^{-\frac{\lambda}{2}(v_1^2 + v_2^2)} dv_2 dv_1 \end{aligned}$$

and the joint *pdf* of (R_1, R_2) given $\mathcal{G}$ is the derivative:

$$f_{R_1, R_2}(r_1, r_2 | \mathcal{G}) = \kappa r_1 r_2 (r_1 + r_2)^D e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{\{r_1 > 0, r_2 > 0\}}. \quad (3.1.3)$$

The joint density of (R_1, R_2, H) given $\mathcal{G}$ is the product of the two conditional densities $f_H(h | R_1 = r_1, R_2 = r_2, \mathcal{G})$ and $f_{R_1, R_2}(r_1, r_2 | \mathcal{G})$. So

$$\begin{aligned} f_{R_1, R_2, H}(r_1, r_2, h | \mathcal{G}) &= f_H(h | R_1 = r_1, R_2 = r_2, \mathcal{G}) f_{R_1, R_2}(r_1, r_2 | \mathcal{G}) \\ &= \frac{D h^{D-1}}{(r_1 + r_2)^D} \mathbf{1}_{[0, r_1 + r_2]}(h) \kappa r_1 r_2 (r_1 + r_2)^D e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{\{r_1 > 0, r_2 > 0\}} \\ &= \kappa D h^{D-1} \mathbf{1}_{[0, r_1 + r_2]}(h) r_1 r_2 e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{\{r_1 > 0, r_2 > 0\}}. \end{aligned} \quad (3.1.4)$$

We need the following three special cases later in simulation and in application, for $D=1$,

$$f_{R_1, R_2, H}(r_1, r_2, h | \mathcal{G}) = \frac{x^5 \lambda^{5/2}}{\sqrt{2\pi}} r_1 r_2 e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{[0, r_1 + r_2]}(h) \mathbf{1}_{\{r_1 > 0, r_2 > 0\}},$$

for $D=2$,

$$f_{R_1, R_2, H}(r_1, r_2, h | \mathcal{G}) = \frac{2h x^6 \lambda^3}{\pi + 4} r_1 r_2 e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{[0, r_1 + r_2]}(h) \mathbf{1}_{\{r_1 > 0, r_2 > 0\}},$$

and for $D=3$,

$$f_{R_1, R_2, H}(r_1, r_2, h | \mathcal{G}) = \frac{h^2 x^7 \lambda^{7/2}}{3\sqrt{2\pi}} r_1 r_2 e^{-\frac{\lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{[0, r_1 + r_2]}(h) \mathbf{1}_{\{r_1 > 0, r_2 > 0\}}. \quad (3.1.5)$$

To do more inference about the distribution of the volume of the clusters we need to know what is the intersection volume between two D -dimensional balls. We give it in the following theorem:

Theorem 3.1.1. Intersection volume of two D -dimensional balls: Consider two D -dimensional balls of radii R_1, R_2 separated by a distance H between their centers and let

$R_{(1)} \leq R_{(2)}$ be the ordered R_1 and R_2 . Let

$$\begin{aligned} x_2 &= (R_{(2)}^2 + H^2 - R_{(1)}^2)/2H, \\ x_1 &= H - x_2. \end{aligned}$$

Let g be a function defined by the following formula

$$\begin{aligned} g(D, r, a) &= \omega_{D-1} \int_a^r (r^2 - y^2)^{\frac{D-1}{2}} dy \\ &= \frac{\sqrt{\pi} r^D \Gamma(\frac{D+1}{2})}{D \Gamma(D/2)} - ar^{D-1} F\left(\frac{1}{2}, -\frac{D-1}{2}, \frac{3}{2}, \frac{a^2}{r^2}\right), \end{aligned}$$

where F is the hypergeometric function defined by

$$F(a, b, c, z) = \sum_{i=0}^{\infty} \frac{(a)_i (b)_i z^i}{(c)_i i!}$$

and $(a)_i = \Gamma(a+i)/\Gamma(a)$. The volume of the intersection of the two D -dimensional balls is

$$V = \begin{cases} \omega_D R_{(1)}^D & \text{if } 0 \leq H \leq R_{(2)} - R_{(1)}, \\ g(D, R_{(1)}, x_1) + g(D, R_{(2)}, x_2) & \text{if } R_{(2)} - R_{(1)} \leq H \leq R_{(1)} + R_{(2)}, \end{cases}$$

Proof: Let B_1, B_2 be two D -dimensional balls with radii $R_{(1)}, R_{(2)}$ respectively. Let H be the distance between their centers. If $0 \leq H \leq R_{(2)} - R_{(1)}$, then $B_1 \subseteq B_2$, so $V = \text{volume of } B_1 = \omega_D R_{(1)}^D$. If $H > R_{(1)} + R_{(2)}$ then $B_1 \cap B_2$ is empty. If $R_{(2)} - R_{(1)} \leq H \leq R_{(1)} + R_{(2)}$ then $B_1 \cap B_2$ is the union of two disjoint D -dimensional spherical caps one in B_1 and the other one in B_2 . So V is the sum of their volumes. To find V we need to find the volume of a spherical cap. The volume of a spherical cap of height h in a D -dimensional ball of radius r can be obtained by integrating a $(D-1)$ -dimensional ball with radius $\sqrt{r^2 - x^2}$ from $r-h$ to r , i.e.,

$$\omega_{D-1} \int_{r-h}^r (r^2 - x^2)^{\frac{D-1}{2}} dx,$$

which is equal to

$$\frac{\sqrt{\pi} r^D \Gamma(\frac{D+1}{2})}{D \Gamma(D/2)} - ar^{D-1} F\left(\frac{1}{2}, -\frac{D-1}{2}, \frac{3}{2}, \frac{a^2}{r^2}\right),$$

by using the trigonometric substitution $x = r \cos(\theta)$. We need to find the heights of the two caps that form $B_1 \cap B_2$. The spheres of the balls B_1 and B_2 intersect in a circle with radius a , say. Let x_1, x_2 be the distances between the centers of the balls and the center of this circle. Then we have

$$(H - x_2)^2 + a^2 = R_{(1)}^2, \quad (3.1.6)$$

$$x_1 + x_2 = H, \quad (3.1.7)$$

$$x_2^2 + a^2 = R_{(2)}^2. \quad (3.1.8)$$

Let h_1 and h_2 denote the heights of the two caps of B_1 and B_2 respectively. From equations (3.1.6), (3.1.7) and (3.1.8) we get

$$x_2 = (R_{(2)}^2 + H^2 - R_{(1)}^2)/2H,$$

$$x_1 = H - x_2,$$

$$h_1 = R_{(1)} - x_1,$$

$$h_2 = R_{(2)} - x_2.$$

Then

$$\begin{aligned} V &= \omega_{D-1} \int_{x_1}^{R_{(1)}} (R_{(1)}^2 - x^2)^{\frac{D-1}{2}} dx + \omega_{D-1} \int_{x_2}^{R_{(2)}} (R_{(2)}^2 - x^2)^{\frac{D-1}{2}} dx \\ &= g(D, R_{(1)}, x_1) + g(D, R_{(2)}, x_2). \end{aligned}$$

This completes the proof. For $D=1, 2$ and 3 there is a simple formula for V . These simple formulas will be used later in simulations and applications, so we give them in the following three corollaries:

Corollary 3.1.2. *for $D=1$,*

$$V = (R_{(1)} + R_{(2)} - H) \mathbf{1}_{[R_{(2)}-R_{(1)}, R_{(1)}+R_{(2)}]}(H) + 2R_{(1)} \mathbf{1}_{[0, R_{(2)}-R_{(1)}]}(H).$$

Corollary 3.1.3. *for $D=2$,*

$$V = \begin{cases} \sum_{j=1}^2 \frac{R_{(j)}^2}{2} \{2\theta_j - \sin(2\theta_j)\} & \text{if } R_{(2)} - R_{(1)} \leq H \leq R_{(1)} + R_{(2)}, \\ \pi R_{(1)}^2 & \text{if } 0 \leq H \leq R_{(2)} - R_{(1)}, \end{cases}$$

where $\theta_1 = \cos^{-1}((R_{(1)}^2 + H^2 - R_{(2)}^2)/(2HR_{(1)}))$, $\theta_2 = \cos^{-1}((R_{(2)}^2 + H^2 - R_{(1)}^2)/(2HR_{(2)}))$.

Corollary 3.1.4. *for $D=3$,*

$$V = \begin{cases} \frac{4\pi}{3}R_{(1)}^3 & \text{if } 0 \leq H \leq R_{(2)} - R_{(1)}, \\ \frac{2\pi}{3}(R_{(1)}^3 + R_{(2)}^3) - \pi(R_{(1)}^2 h_1 + R_{(2)}^2 h_2 - \frac{1}{3}(h_1^3 + h_2^3)) & \text{if } R_{(2)} - R_{(1)} \leq H \leq R_{(1)} + R_{(2)}, \end{cases}$$

where

$$h_1 = (H^2 + R_{(1)}^2 - R_{(2)}^2)/(2H), \quad h_2 = H - h_1.$$

3.2 $\mathbb{E}\{V\}$ using the balls model

The distribution of V is not in closed form. However we can find a closed form for the mean of the distribution, which we shall use later to adjust the approximation using the Poisson clumping heuristic in Section 2.7. Using (3.1.4) and the Kinematic Fundamental Formula we can find the mean of the distribution of V given $\mathcal{G}$ as follows.

In the Kinematic Fundamental Formula, let A be a fixed D -dimensional ball with radius R_1 and K be a D -dimensional ball with radius R_2 which moves around A uniformly. Then the conjunction occurs when the two balls overlap. The overlap of the two balls is a cluster related to the conjunction. The intrinsic volumes $\mu_i(B_r)$ for $i = 0, 1, \dots, D$ of a ball B_r with radius r is given by $\mu_i(B_r) = \binom{D}{i} \omega_D r^i / \omega_{D-i}$. Substituting this information in (2.5.4) and using the Kinematic Fundamental Formula to simplify the integrals we get the expected volume of the overlap of the two balls conditional that there is an overlap and conditional on R_1, R_2 . Let V be the volume of one cluster of the excursion set of $\min\{X_1(\mathbf{t}), X_2(\mathbf{t})\}$.

The ratio of the two integrals in 2.5.4 is

$$\begin{aligned}
 \mathbb{E}\{V(H)|V(H) > 0, R_1, R_2, \mathcal{G}\} &= \frac{\begin{bmatrix} D \\ D \end{bmatrix} \begin{bmatrix} D \\ 0 \end{bmatrix}^{-1} \mu_D(A) \mu_D(K)}{\sum_{i=0}^D \begin{bmatrix} i \\ 0 \end{bmatrix} \begin{bmatrix} D \\ i \end{bmatrix}^{-1} \mu_i(A) \mu_{D-i}(K)} \\
 &= \omega_D \frac{R_1^D R_2^D}{\sum_{i=0}^D \binom{D}{i} R_1^i R_2^{D-i}} \\
 &= \omega_D \frac{R_1^D R_2^D}{(R_1 + R_2)^D}.
 \end{aligned}$$

Using the total probability law $\mathbb{E}\{V|\mathcal{G}\} = \mathbb{E}\{\mathbb{E}\{V(H)|R_1, R_2, \mathcal{G}\}\}$ and (3.1.4), the mean of V is

$$\begin{aligned}
 \mathbb{E}_{Ball}\{V\} = \mathbb{E}\{V|\mathcal{G}\} &= \omega_D \int_0^\infty \int_0^\infty \frac{r_1^D r_2^D}{(r_1 + r_2)^D} f_{R_1, R_2}(r_1, r_2|\mathcal{G}) dr_1 dr_2 \\
 &= \kappa \omega_D \int_0^\infty \int_0^\infty r_1^{D+1} r_2^{D+1} e^{-\frac{x^2 \lambda}{2}(r_1^2 + r_2^2)} dr_1 dr_2 \\
 &= \frac{\kappa \omega_D 2^D \Gamma(D/2 + 1)^2}{(x^2 \lambda)^{D+2}}.
 \end{aligned} \tag{3.2.9}$$

3.3 $\mathbb{E}\{V\}$ using the Poisson clumping heuristic

In this section we will describe how to approximate the mean value of the volume of a cluster of the excursion set of a stationary field $X(\mathbf{t})$, $\mathbf{t} \in S \subset \mathbb{R}^D$ in a different way. Let A be the excursion set of X above a level x . Then we can write

$$\mu_D(A) = \int_S \mathbf{1}_A d\mathbf{t}.$$

Taking the expectation for both sides and considering stationarity of X we get

$$\begin{aligned}
 \mathbb{E}\{\mu_D(A)\} &= \int_S \mathbb{E}\{\mathbf{1}_A\} d\mathbf{t} \\
 &= \int_S \mathbb{P}\{X(\mathbf{t}) \geq x\} d\mathbf{t} \\
 &= \mu_D(S) \mathbb{P}\{X(\mathbf{0}) > x\}.
 \end{aligned} \tag{3.3.10}$$

Since the excursion set above a high level is composed of disjoint components we can write it as follows

$$A = C_1 \cup C_2 \cup \dots \cup C_N,$$

where N is the number of these clusters. Then if $V_j = \mu_D(C_j)$,

$$\mu_D(A) = V_1 + V_2 + \dots + V_N,$$

and so

$$\mathbb{E}\{\mu_D(A)\} = \mathbb{E}\{\mathbb{E}\{V_1 + V_2 + \dots + V_N|N\}\} = \mathbb{E}\{N\mathbb{E}\{V\}\}.$$

where V denotes the volume of one cluster of the excursion set. If S is large, the number of clusters is independent of their sizes, and we have the following formula

$$\mathbb{E}\{\mu_D(A)\} = \mathbb{E}\{N\}\mathbb{E}\{V\}$$

and so combining this with (3.3.10),

$$\mathbb{E}\{V\} = \mu_D(S)\rho_0(x)/\mathbb{E}\{N\}, \quad (3.3.11)$$

where $\rho_0(x) = \mathbb{P}\{X(0) > x\}$. $\mathbb{E}\{N\}$ can be approximated by the mean value of the Euler characteristic function using (2.6.5). Since S is large, the $d = D$ term is the most important, so

$$\mathbb{E}\{N\} \approx \mu_D(S)\rho_D(x).$$

When $X(t) = X_*(t)$ is a conjunction of two isotropic random fields, each with EC densities ρ_i , then the zero and D -dimensional EC densities of the conjunction are

$$\begin{aligned} \rho_{*0}(x) &= \rho_0^2, \\ \rho_{*D}(x) &= b_D \sum_{i=0}^D \frac{\rho_i(x)\rho_{D-i}(x)}{b_i b_{D-i}}. \end{aligned}$$

This can be obtained from (2.6.5) by writing $\mathbb{E}\{\mu(A_*)\}$ as a linear combination of the Minkowski functionals of S . This can be done by simplifying the right hand side of (2.6.5) and then picking the coefficients of $\mu_0(S)$ and $\mu_D(S)$. Hence an approximation to $\mathbb{E}\{V\}$ is

$$\mathbb{E}_{PCH}\{V\} = \frac{\rho_{*0}}{\rho_{*D}}.$$

In Figure 3.1 (a)-(c) we note that for large values of x the mean values of V obtained by both balls model and Poisson clumping heuristic are close to each other. Also the value of $\mathbb{E}_{Balls}\{V\}$ becomes closer to $\mathbb{E}_{PCH}\{V\}$ as the dimension D increases, which is a good

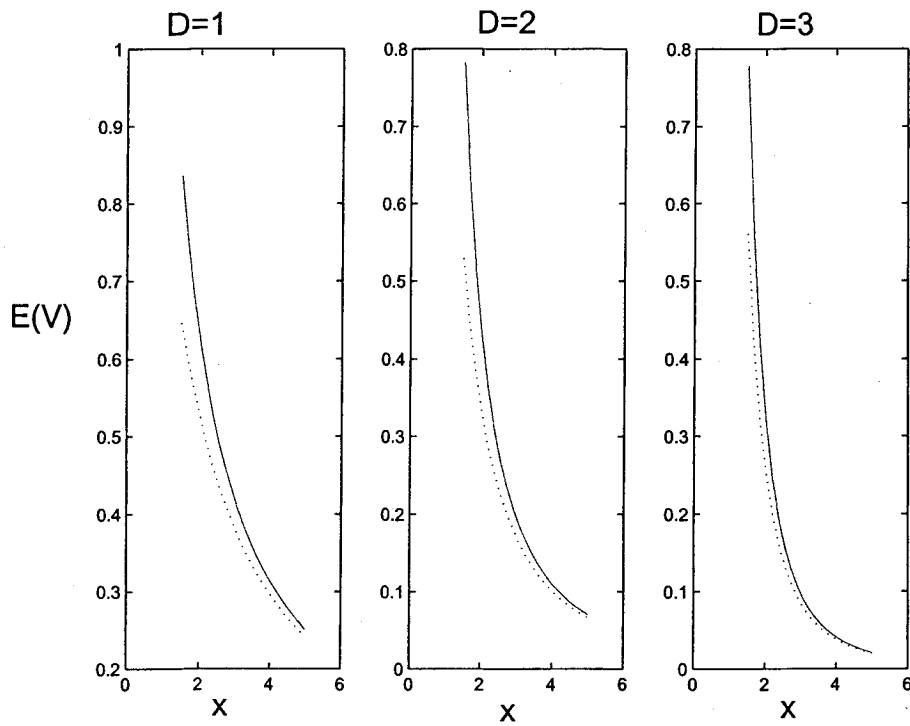


Figure 3.1: Two different approximations to the expected volume of a single cluster: $E_{Balls}\{V\}$ (dotted) and $E_{PCH}\{V\}$ (solid) plotted against threshold x .

advantage of the balls model because we are interested in application for higher dimensions. This gives an indication that the approximation of the distribution of V from the balls model is working well for large thresholds and high dimensions.

3.4 Distribution of the maximum cluster volume

In this section we will describe how to approximate the distribution of the maximum volume of the clusters of the conjunction using the Poisson clumping heuristic approach given by Aldous (1989) and used by Friston *et al.* (1994) to find the distribution of the maximum volume for the excursion set of a single Gaussian random field. Let $C_1, C_2, \dots, C_N$ be the clusters of the excursion set A_* of the random field X_* . Let V_i denote the volume of cluster i . Let V_{max} denote the maximum of the $V_1, V_2, \dots, V_N$. If N is the number of clusters of X_* above x , then N has approximately a Poisson distribution with mean $\mathbb{E}\{N\}$, i.e.,

$$\mathbb{P}\{N = n\} \approx \frac{(\mathbb{E}\{N\})^n}{n!} e^{-\mathbb{E}\{N\}}, \text{ for } n = 0, 1, 2, \dots$$

The *cdf* of V_{max} is

$$\begin{aligned} \mathbb{P}\{V_{max} \leq v | N \geq 1\} &= \sum_{n=1}^{\infty} \mathbb{P}\{V_i \leq v, 1 \leq i \leq n, N = n | N \geq 1\} \\ &= \sum_{n=1}^{\infty} \mathbb{P}\{N = n | N \geq 1\} \mathbb{P}\{V_i \leq v, 1 \leq i \leq n | N = n\} \\ &\approx \frac{1}{\mathbb{P}\{N \geq 1\}} \sum_{n=1}^{\infty} \frac{(\mathbb{E}\{N\})^n}{n!} e^{-\mathbb{E}\{N\}} \mathbb{P}\{V \leq v\}^n \\ &= \frac{\exp(-\mathbb{E}\{N\}) \mathbb{P}\{V \geq v\} - \exp(-\mathbb{E}\{N\})}{1 - \exp(-\mathbb{E}\{N\})}, \end{aligned} \quad (3.4.12)$$

where V denotes the volume of any one cluster. Note that for large $\mathbb{E}\{N\}$ we have

$$\mathbb{P}\{V_{max} \leq v | N \geq 1\} \approx \exp\{-\mathbb{E}\{N\} \mathbb{P}\{V \geq v\}\}.$$

The last formula means that we can write the distribution of V_{max} in terms of $\mathbb{E}\{N\}$ and the distribution of V . In Section 3.1 we found the joint distribution of $(R_1, R_2, H) | \mathcal{G}$ which can be used to approximate the distribution of V , and $\mathbb{E}\{N\}$ can be approximated according by the expected EC using (2.6.5). So we can now get the approximate distribution of V_{max} .

3.5 Correcting the mean of the distribution of V

Friston et al. (1994) and Cao (1999) found that the approximate distribution of V could be considerably improved by re-scaling V so that its expectation agreed with that given by (3.3.11) using the Poisson clumping heuristic. In other words, we suppose that $\mathbb{E}\{V\}$ given by the Poisson clumping heuristic (3.3.11) is much more accurate than that given by the balls model in (3.2.9). In one particular case, discussed in the next Section 3.6, $\mathbb{E}_{PCH}\{V\}$ is exact. We will correct the distribution of V given $\mathcal{G}$ in the same way, to give

$$\tilde{V} = \frac{\mathbb{E}_{PCH}\{V\}}{\mathbb{E}_{Ball}\{V\}} V,$$

so that $\mathbb{E}\{\tilde{V}\} = \mathbb{E}_{PCH}\{V\}$. We will show that using the distribution of $\tilde{V}$ instead of V , we get a better approximation to the distribution of the volume of a single cluster.

3.6 Example of a one-dimensional stationary Gaussian random field

In this section we will consider an example of a one-dimensional stationary Gaussian random field. In this case we have exact results which can be compared to our approximate results. This example is the cosine random field which satisfies the regularity conditions mentioned in Chapter 3 of Adler (1981). Let Z_1, Z_2, Z_3, Z_4 be independent standard normal random variables. We can define the following two cosine random fields over $S = [0, 2\pi]$.

$$\begin{aligned} X_1(t) &= Z_1 \sin(t) + Z_2 \cos(t), \\ X_2(t) &= Z_3 \sin(t) + Z_4 \cos(t). \end{aligned}$$

Figure 3.2 show a realization of X_1 and X_2 where the conjunction of them has one cluster. The derivatives are

$$\begin{aligned} \dot{X}_1(t) &= Z_1 \cos(t) - Z_2 \sin(t), \\ \dot{X}_2(t) &= Z_3 \cos(t) - Z_4 \sin(t). \end{aligned}$$

Note that the derivative fields $\dot{X}_1(t)$ and $\dot{X}_2(t)$ are independent $Normal(0, 1)$, so $\lambda = 1$. Let

$$\begin{aligned} Y_1 &= \sup_S X_1(t) = \sqrt{Z_1^2 + Z_2^2}, \\ Y_2 &= \sup_S X_2(t) = \sqrt{Z_3^2 + Z_4^2}. \end{aligned}$$

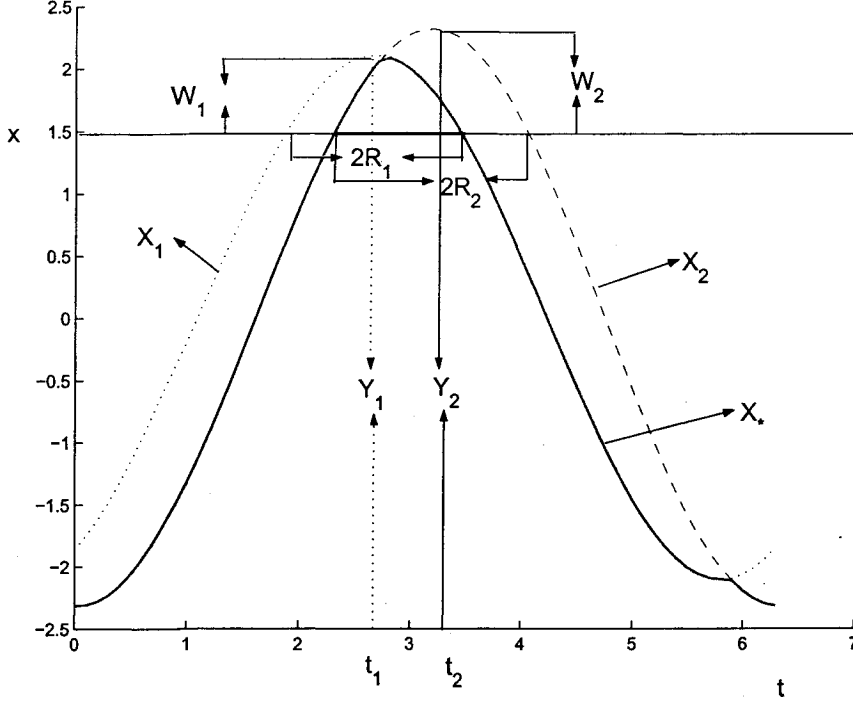


Figure 3.2: An example of the one dimensional conjunction of two cosine Gaussian Random fields. The threshold is $x = 1.5$.

Then we can write

$$X_1(t) = Y_1 \cos(t - \theta_1),$$

$$X_2(t) = Y_2 \cos(t - \theta_2),$$

where θ_1, θ_2 are independent $U(0, 2\pi)$ and Y_1, Y_2 are independent with the following density,

$$f(y) = y \exp(-\frac{1}{2}y^2) \mathbf{1}_{\{y>0\}}.$$

This follows from the fact that $Z_1^2 + Z_2^2$ is distributed as a χ^2 random variable with 2 degrees of freedom. We need to find the distribution of Y_j given that $Y_j > x$ where x is a large threshold. Now

$$\mathbb{P}\{Y_j > y \mid Y_j > x\} = \exp\left(\frac{x^2}{2} - \frac{y^2}{2}\right) \mathbf{1}_{\{y>x\}}. \quad (3.6.13)$$

If we regard the random fields X_j as periodic on $[0, 2\pi]$, so that we neglect the boundary, then the excursion set has one single cluster uniformly centered on $[0, 2\pi]$ whose radius is $\cos^{-1}(x/Y_j)$. If $x > 0$ then this cluster size never exceeds half S , so the conjunction is always empty or a single cluster but never more than one cluster. This fact, considered with the fact

that S is ‘periodic’, means that all the approximations used to derive the Poisson clumping heuristic are exact, and so

$$\mathbb{E}\{V\} = \mathbb{E}_{PCH}\{V\} = \sqrt{\frac{\pi}{2}} \frac{1 - \Phi(x)}{\phi(x)}$$

On the other hand,

$$\mathbb{E}\{V\} \approx \mathbb{E}_{Ball}\{V\} = \frac{1}{x} \sqrt{\frac{\pi}{2}}.$$

3.7 The χ^2 random field

Following the same techniques used for the Gaussian field but working with the Theorems 2.4.3, 2.4.4 and 2.4.5 we have that the radii of the clusters are asymptotically distributed as

$$R_1 = \sqrt{2\tilde{W}_1/x\tilde{\lambda}}, \quad R_2 = \sqrt{\tilde{W}_2/x\tilde{\lambda}},$$

where $\tilde{\lambda} = \lambda/x$ and $\tilde{W}_1, \tilde{W}_2$ are independent and exponentially distributed random variables with mean $1/x$. This is the same as for the Gaussian field in Section 3.1, so we can use the same theory of Section 3.1 but with λ replaced by $\tilde{\lambda}$.

Chapter 4

Simulation

Before we apply the theory that we have developed in Chapter 3 to real brain images we need to check if that developed theory works well or not, that is, whether the approximations we have found for the distribution of V and V_{max} are close to the true distribution. Simulation is a good way to check the validity of that theory. For higher dimensions the simulations will take a long time to get a large enough sample so we will restrict our simulation to the case where the dimension is two and the random field is Gaussian.

4.1 Simulation of V from the balls model

By Corollary 3.1.3 V is

$$V = \begin{cases} \sum_{j=1}^2 \frac{R_{(j)}^2}{2} \{2\theta_j - \sin(2\theta_j)\} & \text{if } R_{(2)} - R_{(1)} \leq H \leq R_{(1)} + R_{(2)}, \\ \pi R_{(1)}^2 & \text{if } 0 \leq H \leq R_{(2)} - R_{(1)}, \end{cases} \quad (4.1.1)$$

where $\theta_1 = \cos^{-1}((R_{(1)}^2 + H^2 - R_{(2)}^2)/2HR_{(1)})$, $\theta_2 = \cos^{-1}((R_{(2)}^2 + H^2 - R_{(1)}^2)/2HR_{(2)})$. Any inference about the distribution of V can be drawn based on a large sample from the density of the random vector (R_1, R_2, H) given $\mathcal{G}$. If we use the probability integral transform to draw this sample we will face the following two problems in simulation: the Newton-Raphson method may not converge and the cost in calculations. So we have to use another method of simulation. The *envelope accept-reject method* is a general method and works for a large family of distributions. The following theorem is from Christian and Casella (1999):

Theorem 4.1.1. *If there exists a density g_2 , a function g_1 and a constant M such that*

$$g_1(x) \leq f(x) \leq M g_2(x)$$

then the algorithm

1. *generate $X \sim g_2(x)$, $U \sim \mathcal{U}(0, 1)$.*
2. *accept X if $U \leq g_1(X)/M g_2(X)$.*
3. *otherwise, accept X if $U \leq f(X)/M g_2(X)$.*

produces random variables that are distributed according to f .

We can use the last theorem to simulate from the density (3.1.3)

$$f_{R_1, R_2}(r_1, r_2 | \mathcal{G}) = \frac{x^6 \lambda^3}{4 + \pi} r_1 r_2 (r_1 + r_2)^2 e^{-\frac{x^2 \lambda}{2}(r_1^2 + r_2^2)} \mathbf{1}_{\{r_1 > 0, r_2 > 0\}}$$

by finding a suitable g_1 and g_2 . To find g_2 we need only to dominate the term $(r_1 + r_2)$ by writing it as $(r_1 + r_2) = (r_1, 1) \cdot (1, r_2)$ and then applying the Schwartz inequality. The g_1 bound can be obtained by applying the inequality $2\sqrt{r_1 r_2} \leq (r_1 + r_2)$. This method will make simulation more easy since we will get g_1 and g_2 as the product of two independent densities up to normalizing constants.

Assume that R_1 and R_2 are simulated from the density (3.1.3). Then we simulate H from the density (3.1.2) by inverting its *cdf* as follows

1. Generate q from $\mathcal{U}(0, 1)$.
2. Let $H = q^{1/2}(R_1 + R_2)$.

Then (R_1, R_2, H) is distributed according to $f_{R_1, R_2, H}(r_1, r_2, h | \mathcal{G})$.

Another way of simulating observations from the density (3.1.3) is by writing it as a finite mixture distribution of $D + 1$ components which are easy to simulate. This can be done by expanding the term $(r_1 + r_2)^D$ using the Binomial theorem. Let

$$\mathcal{I}_i = \int_0^\infty r_1^{i+1} e^{-\frac{1}{2}x^2 \lambda r_1^2} dr_1 = \frac{2^{i/2} \Gamma(i/2 + 1)}{(x^2 \lambda)^{i/2+1}} \quad \text{for } i = 0, 1, \dots, D.$$

Then

$$\begin{aligned} \kappa^{-1} &= \int_0^\infty \int_0^\infty (r_1 + r_2)^D r_1 r_2 e^{-\frac{1}{2}x^2 \lambda (r_1^2 + r_2^2)} dr_2 dr_1 \\ &= \sum_{i=0}^D \binom{D}{i} \mathcal{I}_i \mathcal{I}_{D-i}. \end{aligned}$$

We can rewrite the density as follows

$$\begin{aligned} f_{R_1, R_2}(r_1, r_2 | \mathcal{G}) &= \kappa \sum_{i=0}^D \binom{D}{i} \mathcal{I}_i \mathcal{I}_{D-i} r_1^{i+1} e^{-\frac{1}{2}x^2 \lambda r_2^2} r_2^{D+1-i} e^{-\frac{1}{2}x^2 \lambda r_2^2} \mathcal{I}_i^{-1} \mathcal{I}_{D-i}^{-1} \\ &= \sum_{i=0}^D p_i f_i(r_1) f_{D-i}(r_2), \end{aligned}$$

where

$$\begin{aligned} p_i &= \frac{\binom{D}{i} \mathcal{I}_i \mathcal{I}_{D-i}}{\sum_{i=0}^D \binom{D}{i} \mathcal{I}_i \mathcal{I}_{D-i}}, \\ f_i(r_1) &= \mathcal{I}_i^{-1} r_1^{i+1} e^{-\frac{1}{2}x^2 \lambda r_1^2}, \\ f_{D-i}(r_2) &= \mathcal{I}_{D-i}^{-1} r_2^{D-i+1} e^{-\frac{1}{2}x^2 \lambda r_2^2}. \end{aligned}$$

The two densities $f_i(r_1)$ and $f_{D-i}(r_2)$ are easy to simulate. The following algorithm is to simulate from this finite mixture

1. Simulate i from $0, 1, \dots, D$ with probabilities $p_0, p_1, \dots, p_D$ respectively.
2. Simulate R_1 from $f_i(r_1)$ and R_2 from $f_{D-i}(r_2)$.
3. Simulate V using (4.1.1).

4.2 Simulation of V and V_{max} from random fields

An efficient method for simulating stationary Gaussian random field is to smooth white noise using the fast fourier transform (FFT). This method is based on the following theorem.

Theorem 4.2.1. *A strictly stationary continuous Gaussian random field with zero mean, variance σ^2 and auto-correlation function $R(\mathbf{t}) = \exp(-\mathbf{t}^T \Sigma^{-1} \mathbf{t}/4)$ can be obtained by convolving a white noise random field of variance $\sigma^2 2^D \pi^{D/2} \sqrt{|\Sigma|}$ with a Gaussian kernel of covariance Σ ,*

$$f(\mathbf{x}) = \exp(-\mathbf{x}^T \Sigma^{-1} \mathbf{x}/2) / \sqrt{(2\pi)^D |\Sigma|}.$$

This theorem means that a standard Gaussian random fields can be generated by smoothing a white noise field with a Gaussian kernel with covariance matrix Σ . The covariance matrix Λ of the partial derivatives of the Gaussian field and Σ are related by the formula

$$\Lambda = (2\Sigma)^{-1}$$

see Holmes (1994). In medical imaging, filter widths are commonly expressed in terms of Full Width at Half Maximum (FWHM) rather than Σ . For a one dimensional filter, the FWHM is the width of the filter at half its maximum. For a Gaussian shaped filter with variance σ^2 , $\text{FWHM} = \sigma\sqrt{8\ln 2}$.

This was used to simulate a sample of 5000 realizations of the sample paths of the random field of the conjunction of two independent isotropic Gaussian random fields in a rectangular region S . At each realization we threshold the sample path by a large threshold x and then find the area of each cluster of the excursion set of the conjunction. To do this we used the MATLAB function `bwlabel` which takes a binary images (value 0 where $X(\mathbf{t}) \leq x$, 1 where $X(\mathbf{t}) > x$) and assigns integer labels to each pixel so that all pixels in a cluster have the same label (label 0 is assigned to pixels outside the clusters).

We want to use the simulation to check two things: our approximation to the distribution of V , the volume of one cluster, and our approximation to the distribution of $V_{\max}$, the volume of the largest cluster in a finite region S . Before doing this, we must deal with the boundary of S .

A cluster that touches the boundary will be reduced in size, which would give a biased distribution for V , so it is important to remove the boundary effect if we want to simulate V accurately. We cannot simply ignore any cluster that touches the boundary, since big clusters are more likely to touch the boundary than small ones, which would again bias the distribution of V . Instead we took advantage of the fact that our simulations are periodic on S because we used the fast Fourier transform. To remove the effect of the boundary of S we joined the clusters that touch the opposite sides of the rectangular region S as one cluster. To do this we wrote our own MATLAB function `bwlabel2`, which produces the same results as `bwlabel`, but clusters that touch the opposite boundaries are connected as the same cluster. The `bwlabel2` is better to verify the distribution of V while `bwlabel` is closer to reality.

Another problem is the discrete sampling of the random field at pixels. The smoothness of the field, measured by FWHM, should be high relative to the pixel size (which is 1) to ensure adequate coverage of the clusters. However in real applications the FWHM is not large relative to the pixel size, and we measure cluster volume by the number of pixels in the

cluster multiplied by the size of one pixel. To investigate the effect of this on the accuracy of our approximations, we tried varying the FWHM. To see the effect of the smoothness on the approximation we used FWHM=2.5, 5 and 10.

Also in real applications images are not periodic and boundary effects have an important effect on the distribution of V_{max} . To assess this, we also tried varying the shape of S while maintaining its area constant. In these simulations we used `bwlabel` so that clusters that touch the boundary were reduced in size, and we focussed on the distribution of V_{max} rather than V .

We chose a threshold x such that $P\{X_*(t) \geq x\} = \alpha$. We will considered three thresholds corresponding to $\alpha = 0.01, 0.001$ and 0.0001 - the middle value is the default for statistical packages such as SPM and FMRISTAT.

Also we simulated 5000 samples with $\lambda = 1$ from the density (3.1.5). The simulated data are then described by their empirical distributions as in Figures 4.1-4.3

We note from Figures 4.1 (a)-(b) that the corrected distribution of V , corrected by adjusting the mean, gives better results than the uncorrected one. The P -values obtained by the approximate distribution of V from the balls model are over estimates of the true P -values obtained from the simulation which means that the P -values obtained by the balls model is conservative. Also as we expect, the approximation becomes better as x gets larger. In Figure 4.2 (a)-(i) the approximate distribution of V_{max} is bad when FWHM=2.5, and becomes better when FWHM=5 and x is large. When FWHM=10 the approximation is the best for the largest two thresholds.

The boundary of S for the images in Figure 4.3 (a)-(c) is 256 and the boundary of S for the images in Figure 4.3 (d)-(f) is 320 while the boundary of S for the images in figure 4.3 (g)-(i) is 544. So we note that the boundary does not have a big effect on the approximation to the distribution of V_{max} so we can ignore it. Also we note that the approximation is bad when $x=1.8575$ since in this case the excursion set is more likely to cut the boundary. Finally, for $x \geq 1.8575$ we can use the theory safely to test wether a given cluster is significant or not.

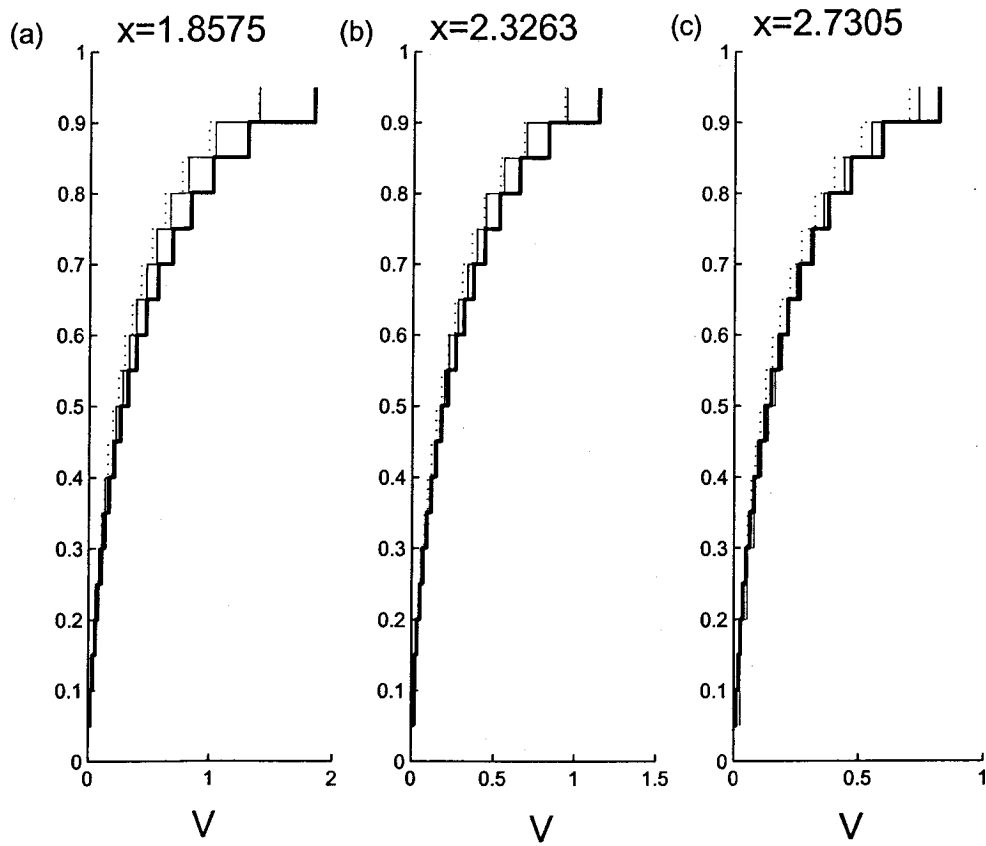


Figure 4.1: The bold line is the uncorrected *cdf* of V from the balls model and the thin line is its correction for the mean. The dotted line is the empirical *cdfs* of V from simulation using `bwlabel2`. The image size is 128×128 and $\text{FWHM}=10$.

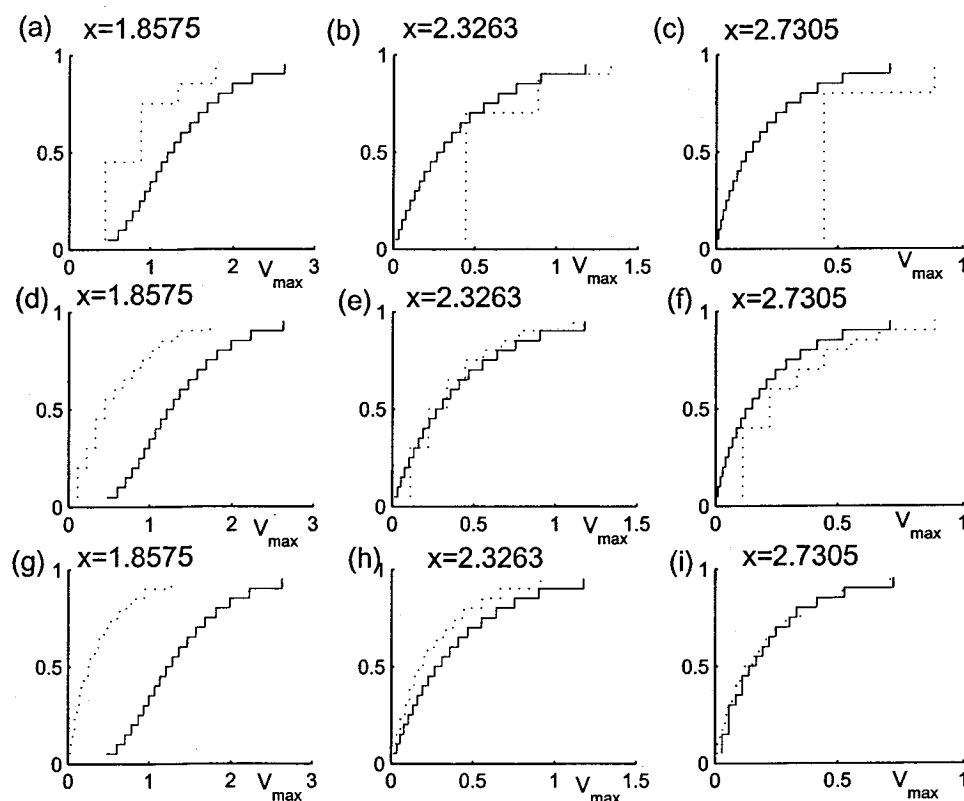


Figure 4.2: The solid line is the corrected theoretical *cdf* of V_{max} and the dotted is the true *cdf* of V_{max} from 64×64 simulated images. FWHM=2.5 in (a)-(c), FWHM=5 in (d)-(f) and FWHM=10 in (g)-(i). The boundary effect was removed using `bwlabel2`.

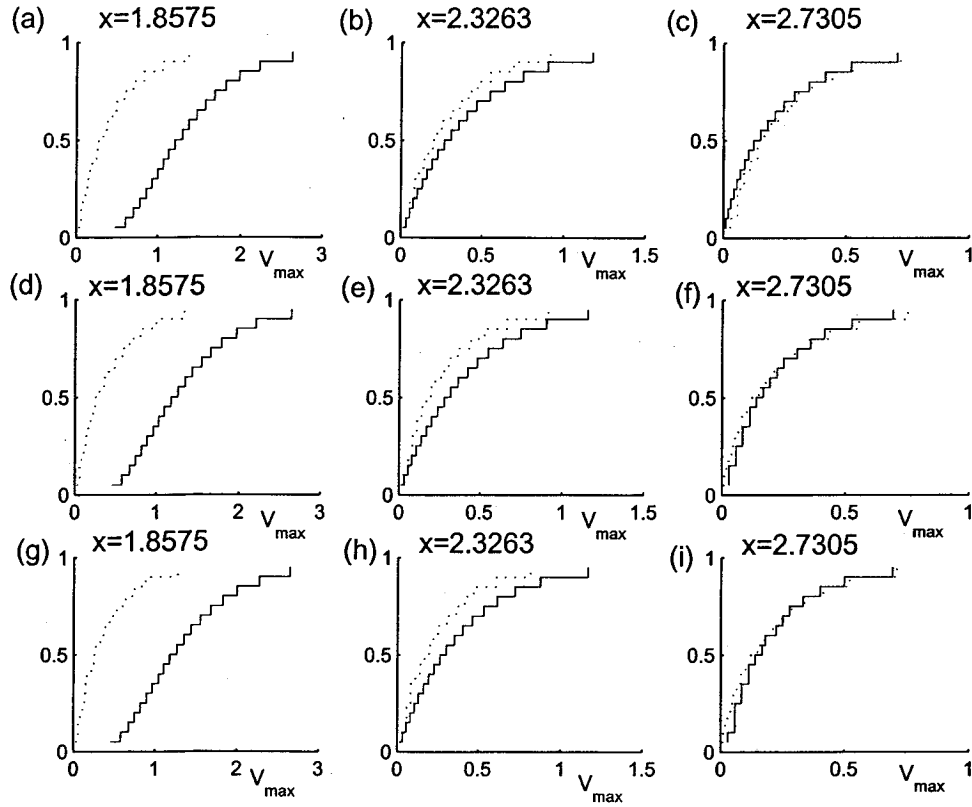


Figure 4.3: The solid line is the corrected theoretical *cdf* of V_{max} and the dotted is the true *cdf* of V_{max} from simulated rectangular images with FWHM=10. The image size is 64×64 in (a)-(c), 32×128 in (d)-(f) and 16×256 in (g)-(i). The boundary effect was kept using bwlabel.

Chapter 5

Application

We now consider real brain images and apply the theory of Chapter 3 to them. We have two brain images (Figures 1.3(a,b)) which are following the T random field with 110 degrees of freedom. Since the number of degrees of freedom is large we can approximate the data as Gaussian random fields. This approximation was improved by applying the transformation $\Phi^{-1}(F(\cdot))$ to the data, where F is the *cdf* of the T distribution with 110 degrees of freedom, and Φ is the *cdf* of the standard Normal distribution. The smoothness was taken as $FWHM=8\text{mm}$ (Worsley *et al.*, 2002), so this gives $\lambda = 4 \log(2)/FWHM^2 = 0.0433\text{mm}^{-2}$. The volume of the search region S is $970,000\text{mm}^3$. If the parameter space of the random fields is re-scaled so that $\lambda = 1$, then the re-scaled volume of S becomes $\mu_3(S) = 970000 \times 0.0433^{3/2} = 8746.3$, which is now unitless.

The two images are threshold at $x = 1.8575$ (1.88 for the untransformed T statistic images), which corresponds to $\alpha = 0.001$ for the conjunction. The mean number of conjunction clusters above this threshold, if there is no activation, is $\mathbb{E}\{N\} = 26.96$. MATLAB was used to find and locate the clusters of the conjunction. The observed number of clusters is $N = 60$.

To find the P -values of the observed volumes of the clusters of the conjunctions, 10,000 observations from the joint density of (R_1, R_2, H) given $\mathcal{G}$ were generated according to 3.1.5 and the volume V is calculated according to Corollary 3.1.4. The expected volume of a single cluster, found using the Poisson clumping heuristic, is $\mathbb{E}\{V\} = 0.3244(\text{unitless})$, or

36.0mm³. Table 5.1 lists the volumes of all 60 clusters and their approximate P -values.

The clusters with P -value less than 0.05 are shown in Figure 5.1. The largest cluster covers the right primary somatosensory area. This is to be expected, since the left leg received the hot and warm stimuli. The next two largest clusters cover the left and right thalamus. These regions are thought to be involved in the perception of pain, as opposed to just the sensation of touch which activates the primary somatosensory area.

Frequency	Volume, mm ³	Volume ($\lambda = 1$)	Approximate P -value
20	38.5	0.35	0.274
13	76.9	0.69	0.154
2	115.4	1.04	0.097
5	153.8	1.39	0.066
1	192.3	1.73	0.047
2	230.7	2.08	0.034
1	346.1	3.12	0.015
2	384.5	3.47	0.012
2	423.0	3.81	0.010
1	461.4	4.16	0.008
1	499.9	4.51	0.006
1	576.8	5.20	0.004
1	615.2	5.55	0.004
3	692.1	6.24	0.002
1	884.4	7.97	0.001
1	1345.8	12.14	0
1	4114.4	37.10	0
1	6075.4	54.78	0
1	22379.2	201.79	0

Table 5.1: Approximate P -values of the volumes of all 60 clusters of the conjunctions of the two T -statistic images. Because the volume is measured by the number of voxels times the volume of a single voxel, some clusters have equal size, so the first column (Frequency) counts the number of such clusters. The cluster volume is measured in mm³ and on the unitless scale of $\lambda = 1$. Only those clusters with P -values less than 0.05 (rows 5-19) are shown in Figure 5.1.

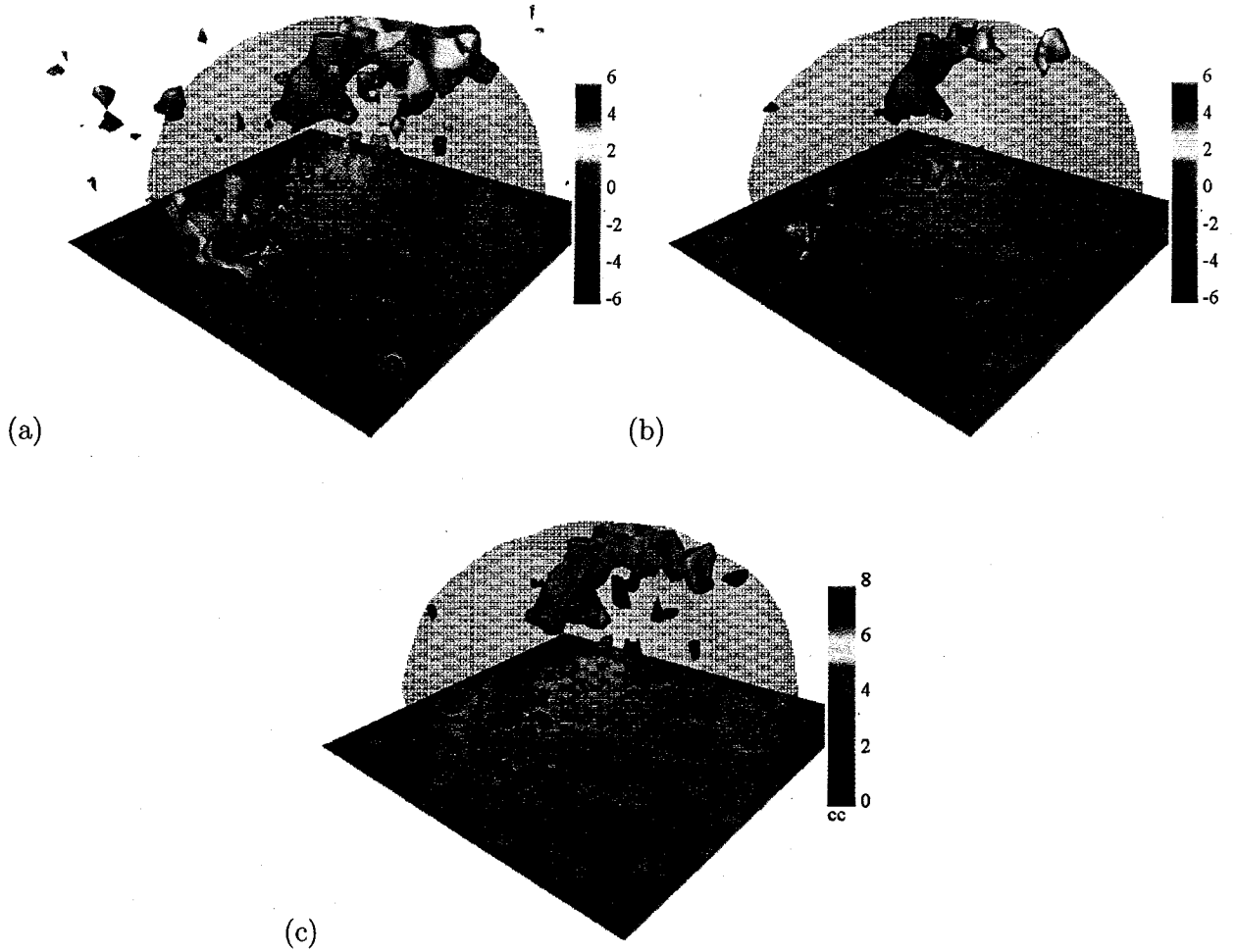


Figure 5.1: Application to conjunctions of pain perception. (a) The conjunction, the intersection of the excursion sets in Figures 1.3(a) and (b), the same as in Figure 1.3(c). The threshold $x = 1.88$ was chosen so that the P-value of the conjunction at any point is 0.001. (b) The conjunction threshold at $x = 3.06$ chosen so that the P-value of the maximum of the conjunction is 0.05, the same as in Figure 1.3(d). (c) The clusters of the excursion set above $x = 1.88$ whose volume exceeds $v = 0.186\text{cc}$, chosen so that the P-value of the maximum volume is 0.05, colored by their volume (the large cluster has a volume of 22.38cc). Note that both methods (b) and (c) detect activation in the right primary somatosensory area, (white cluster in (c)), and the left and right thalamus (green and orange clusters in (c)). More regions are detected in the conjunction of two runs, than in one single run (Figure 1.1), because the amount of data has doubled.

Chapter 6

Conclusion

For two random fields $X_1(t)$ and $X_2(t)$, $t \in S$, we can define a new random field $X_*(t) = \min\{X_1(t), X_2(t)\}$. A conjunction of $X_1(t)$ and $X_2(t)$, above x occurs when the event $\{t \in S : X_*(t) \geq x\}$ occurs. For smooth and stationary Gaussian or χ^2 random fields this event has a simple form which is a union of convex sets. In this thesis, I have attacked the problem of finding an approximation to the distribution of the volume of one of these convex sets. I used Theorems 2.4.1, 2.4.2, 2.4.4 and 2.4.5 to find the shapes of one clusters of X_1 when the threshold is high. If the fields are isotropic then these clusters are disjoint balls with random radii and random centers uniformly distributed on S . The Gaussian and χ^2 random fields have a deterministic curvatures which allows us to find the distribution of the radii. The conjunction of the two fields occurs when two such balls overlap. Then the distribution of the volume of the overlap was found.

I have followed the same method above to attack the problem for more than two conjunctions and I have found that there is a difficulty to get an answer. This is because the condition $\mathcal{G}$ which represents the occurrence of the conjunctions has no simple theoretical representation. Moreover, it is difficult or impossible to express the volume of the overlap of more than two balls in a simple closed form. For these reasons we can say that the balls model of the clusters is not sufficient to solve the problem for more than two conjunctions. Since the cluster volume of two conjunctions V is a function of the radii and the distance between two balls, it is difficult to find a closed form for the approximate distribution of V

from the balls model. The best we could do was to find the joint distribution of the two radii and the distance between the two balls in closed form. Then any information about the approximate distribution of V can be obtained from this joint distribution.

In Figure 1.5 we note that the clusters become more erratic in shape as the number of conjunctions increases. Also we note that as the number of conjunctions increases a very small threshold is needed for the conjunction to be likely to occur. As the threshold decreases the clusters of the component Gaussian random fields will become less like balls and more like the complement of balls, then the conjunction is the intersection of the complements of balls. So in future work it seems to be possible to attack this problem when the number of conjunctions n is very large and the threshold x is very small.

Also the same approach in Chapter 3 can be followed to solve the conjunction of two random fields of different type, i.e. one is Gaussian and the other one is χ^2 . This problem is simple and easy to do, but we need to look for an application in reality.

For T and F random fields it is difficult to solve the conjunction problem since the curvatures of these fields are random. But it is easy to attack this problem in the one dimensional case when we are interested in applications of these fields to one dimensional real data.

Bibliography

- [1] Adler, R.J. (1981). *The Geometry of Random Fields*. Wiley, New York.
- [2] Aldous, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic*. Springer-Verlag, New York, 1989.
- [3] Cao, J. (1999). The size of the connected components of excursion sets of χ^2 , t and F fields. *Advances in Applied Probability*, **31**, 579-595.
- [4] Cao, J., and Worsley, K.J. (1999a). The geometry of correlation fields, with an application to functional connectivity of the brain. *Annals of Applied Probability*, **9**, 1021-1057.
- [5] Christian, P.R., and Casella, C. (1999). *Monte Carlo Statistical Methods*. Springer-Verlag, New York, Inc.
- [6] Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J-B., Frith, C.D., and Frackowiak, R.S.J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, **2**, 189-210.
- [7] Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., and Evans, A.C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, **1**, 214-220.
- [8] Hayasaka, S., Luan-Phan, K., Liberzon, I., Worsley, K.J., and Nichols, T.E. (2004). Non-Stationary cluster size inference with random field and permutation methods. *NeuroImage*, in press.

- [9] Holmes, A.P. (1994). *Statistical Issues In Functional Brain Mapping*. PhD thesis, Department of Statistics, University of Glasgow.
- [10] Kac, M., and Slepian, D. (1959). Large excursions of Gaussian processes. *Annals of Mathematical Statistics*, **30**, 1215-1228.
- [11] Nosko, V.P. (1969). Local structure of Gaussian random fields in the vicinity of high-level light sources. *Soviet Mathematics Doklady*, **10**, 1481-1484.
- [12] Taylor, J. E. (2001). Gaussian volumes of tubes, Euler characteristic densities and correlated conjunctions. *Proceedings of Singular models and geometric methods in statistics. Institute of Statistical Mathematics*, 7-33.
- [13] Worsley, K.J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2 , f and t fields. *Advances in Applied Probability*, **26**, 13-42.
- [14] Worsley, K.J. (1995). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Annals of Statistics*, **23**, 640-669.
- [15] Worsley, K.J., and Friston, K.J. (2000). A test for a conjunction. *Statistics and Probability Letters*, **47**, 135-140.
- [16] Worsley, K.J., Liao, C., Aston, J.A.D., Petre, V., Duncan, G.H., Morales, F., and Evans, A.C. (2002). A general statistical analysis for fMRI data. *NeuroImage*, **15**, 1-15.