THE SIZE OF THE CONNECTED COMPONENTS OF EXCURSION SETS OF χ^2, t AND F FIELDS

J. Cao,* *Bell Laboratories*

Abstract

The distribution of the size of one connected component and the largest connected component of the excursion set is derived for stationary χ^2, t and F fields, in the limit of high or low thresholds. This extends previous result for stationary Gaussian fields (Nosko, 1969; Adler, 1981) and for χ^2 fields in one and two dimensions (Aronowich and Adler, 1986, 1988). An application of this is to detect regional changes in positron emission tomography images of blood flow in human brain, using the size of the largest connected component of the excursion set as a test statistic.

EXCURSION SET; RANDOM FIELDS; IMAGE ANALYSIS.

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60G15

* Postal address: Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Room 2C-260, Murray Hill, NJ 07974-2070.

1 Introduction

In typical activation studies of brain function with positron emission tomography (PET), a 3-D image representing the difference of the brain activity under baseline and stimulation conditions is obtained. The purpose of these studies is to see which areas of the brain show an increase in brain activity, or "activation", due to the stimulation condition. Usually, the activation is so slight that the experiment must be repeated on several subjects and the differenced images are averaged to improve the signal-to-noise ratio.

If the variance of the differenced image is stationary, then the averaged image is standardized by using the pooled sample variance over all the values in the image to give a Gaussian statistic of unit variance at each location of the brain. This can be modelled as a stationary Gaussian random field with zero mean (Worsley *et al.*, 1992) under the hypothesis of no activation. However, if the stationarity assumption is questionable and indeed the variance could be arbitrary, then the averaged image should be standardized pointwise using the sample variance to give a t statistic at each location. This can also be modelled as a stationary t field (Worsley *et al.* 1992; Worsley 1994) with zero mean under the assumption of no change.

Usually, the global maximum of the random field (Worsley *et al.* 1992; Worsley 1994) is chosen as the test statistic for significant activations. However, some workers have suggested an alternative, the size $S_{\max}$ of the largest connected component of the excursion set above a relatively high threshold, based on the assumption that areas of true neural activity will tend to stimulate changes over contiguous voxels (see Figure 1). For this test statistic, Friston *et al.* (1994) derived an approximation for its null distribution based on the work of Nosko (1969) and Adler (1981). Although this result is only valid for Gaussian fields, Friston *et al.* (1994) suggested that it can be extended to t fields with high degrees of freedom by "Gaussianization". That is, transforming the t field at each point to a Gaussian random variable. If $T(\mathbf{t})$ is a t random field with m degrees of freedom, then the "Gaussianized" t field is defined as $W(\mathbf{t}) = \Phi^{-1}(\Psi(T(\mathbf{t}))$, where Φ and Ψ are the cumulative density functions for normal and Student-t random variable with m degrees of freedom, respectively. Although the resulting "Gaussianized" random field is not a Gaussian random field, this method appears to be reasonably accurate for t fields with large degrees of freedom. One of the aims of this paper is to provide results for the spatial extent test statistic that are valid not only for t fields with high degrees of freedom but also for t fields with low degrees of freedom.

This paper aims to study the distribution of $S_{\max}$ for a smooth stationary χ^2, t or F field at high or low level excursions, which was used to detect signal changes for an image modelled as a χ^2, t or F field. The essential tool used for this study is the Poisson clumping heuristic which was explained in detail in Aldous (1989). The essence of this approach, applied to our case, is the following. First, the problem of $S_{\max}$ can be translated to the problem of the size S of one connected component of the excursion set. This is due to the fact that the connected components can be viewed as clumps that centered at points of a

Poisson process. Second, the clump size of the individual components can be characterized by approximating the underlying random process locally by a simpler, known process for which explicit calculations are possible.

Results presented in this paper extend previous work for a Gaussian field (Nosko, 1969; Adler, 1981) and for a χ^2 field in one and two dimensions (Aronowich and Adler, 1986, 1988). The generalization to χ^2, t and F fields in any dimensions is made possible by results of Worsley (1994) where representations of the derivatives of the fields in terms of independent random variables were obtained. The overall plan is the following. In Section 2, we study the curvature and height of a χ^2, t or F field at high local maxima or low local minima. In Section 3, we investigate the structure of these random fields at local extrema. Using these results, in Section 4, we obtain the limiting distribution of the size of individual and the largest connected component of the excursion set. Finally in Section 5, we apply these results to detect changes in PET images of regional blood flow from a study on pain perception.

2 Curvature and height of a χ^2, t or F field at local extrema

We shall need the following notation throughout this paper. If Y is a function defined on $\mathbf{t} \in \mathrm{I\!R}^N$, then we shall denote the first-order derivative of Y with respect to $\mathbf{t}$ by $\dot{\mathbf{Y}}$ and second-order derivative by $\ddot{\mathbf{Y}}$. For any symmetric matrix $\mathbf{D}$, let $\mathbf{D}^- = \mathbf{D}$ if $\mathbf{D}$ is negative definite and zero otherwise, and let $\mathbf{D}^+ = \mathbf{D}$ if $\mathbf{D}$ is positive definite and zero otherwise. We shall use the notation $\mathrm{Normal}_d(\mu, \mathbf{\Sigma})$ to represent the multivariate normal distribution on $\mathrm{I\!R}^d$ with mean μ and variance $\mathbf{\Sigma}$, χ^2_ν to represent the χ^2 distribution with ν degrees of freedom, t_m to represent the Student-t distribution with m degrees of freedom, $F_{n,m}$ to represent the F distribution with n degrees of freedom in numerator and m degrees of freedom in denominator, $\mathrm{Exponential}(\mu)$ to represent the exponential distribution with expectation μ, $\mathrm{Beta}(\alpha, \beta)$ to represent the beta distribution with parameters α and β, and $\mathrm{Wishart}_d(\mathbf{\Sigma}, \nu)$ to represent the Wishart distribution of a $d \times d$ matrix with expectation $\nu\mathbf{\Sigma}$ and ν degrees of freedom. Finally, we shall let $\stackrel{\mathcal{L}}{=}$ represent equality in law between two random variables and $\stackrel{\mathcal{L}}{\longrightarrow}$ represent convergence in law.

Let $X(\mathbf{t}), X_1(\mathbf{t}), \ldots, X_n(\mathbf{t}), Y_1(\mathbf{t}), \ldots, Y_m(\mathbf{t}), \mathbf{t} \in \mathrm{I\!R}^N$ be independent, identically distributed, stationary, real-valued, zero mean, unit variance Gaussian random fields. Let $\mathbf{\Lambda} = \mathrm{Var}(\dot{\mathbf{X}}(\mathbf{t}))$ be the variance of their first derivatives. Adler (1981) and Worsley (1994) gave the following definitions for a χ^2 field $U(\mathbf{t})$ with n degrees of freedom, a t field $T(\mathbf{t})$ with m degrees of freedom, and an F field $F(\mathbf{t})$ with n numerator and m denominator degrees of freedom:

$$U(\mathbf{t}) = \sum_{i=1}^{n} X_i(\mathbf{t})^2, \tag{2.1}$$

$$T(\mathbf{t}) = X(\mathbf{t}) \Big/ \left\{ \sum_{i=1}^{m} Y_i(\mathbf{t})^2/m \right\}^{\frac{1}{2}}, \tag{2.2}$$

$$F(\mathbf{t}) = \left\{ \sum_{i=1}^{n} X_i(\mathbf{t})^2/n \right\} \Big/ \left\{ \sum_{i=1}^{m} Y_i(\mathbf{t})^2/m \right\}. \tag{2.3}$$

For reasons explained in Section 4.1 and 5.1 of Worsley (1994), we shall require that $m \geq N$ for the t field and $m + n > N$ for the F field to ensure the fields are well defined.

Throughout the paper we shall assume that any component Gaussian random field $Y(\mathbf{t})$ in the above definitions are ergodic and satisfying the condition:

$$\max_{i,j} \mathrm{E}\{|Y^{(ij)}(\mathbf{t}) - Y^{(ij)}(\mathbf{0})|^2\} \leq c\|\mathbf{t}\|^2, \tag{2.4}$$

for some finite $c > 0$ and all small enough $\mathbf{t}$. Here we use the notation $Y^{(ij)}$ to represent the second derivative of Y respect to the ith and jth component of $\mathbf{t}$. This plus the stationarity of the component Gaussian field ensure the stationarity and ergodicity of the χ^2, t and F fields thus defined. This also provides, for these fields, adequate regularity conditions required for the use of the Slepian model process and HW conditioning later on. For instance, one consequence of these regularity conditions is that the random field has continuous first- and second-order derivatives with probability one.

2.1 Curvature

To simplify notation, from now on for a function $Y(\mathbf{t})$, we shall drop the dependence on $\mathbf{t}$ if $\mathbf{t} = \mathbf{0}$ and write e.g. $Y = Y(\mathbf{0})$. We shall use $\mathrm{tr}(\mathbf{B})$ to represent the trace of a matrix $\mathbf{B}$. Unless otherwise stated, we shall use $\phi(\cdot)$ to indicate a probability density function.

For a χ^2, t or F field $Z(\mathbf{t})$, we shall determine, in this subsection, the limiting conditional distribution of the curvature $\ddot{\mathbf{Z}}$ at $\mathbf{0}$, given that $\mathbf{0}$ is a local extremum with height z. This is equivalent to conditioning on the event that $Z = z$, $\dot{\mathbf{Z}} = \mathbf{0}$, and $\ddot{\mathbf{Z}} < \mathbf{0}$ (local maximum) or $\ddot{\mathbf{Z}} > \mathbf{0}$ (local minimum). Similar problems have been studied for a Gaussian random field by Nosko (1972) and for a χ^2 random field by Aronowich and Adler (1986, 1988) in one and two dimensions. The appropriate framework for such a study is via the notion of *horizontal window* (HW) conditioning introduced by Kac and Slepian (1959), and the so-called *Slepian model process*. We shall assume that the reader is familiar with these ideas, and with the basic motivation and methodology behind it. A full survey of the model process approach is given in Lindgren (1984), and a background sufficient for our purposes can be found in Section 10.3 of Leadbetter *et al.* (1983). Using the notation $\|$ adopted by Aronowich and Adler (1986, 1988) to denote conditioning in the HW (or ergodic) sense and using $|$ to denote conditioning in the usual sense, we are now in a position to give the following theorems. We shall omit proofs for χ^2 and F fields and prove the result only for a t field.

Theorem 2.1 *Given that a χ^2 field $U(\mathbf{t})$ has a local extremum at $\mathbf{0}$ with height u,* *(a) if* $\mathbf{t}=\mathbf{0}$ *is a local maximum, then* $u^{-1}\ddot{\mathbf{U}} \xrightarrow{\mathcal{L}} -2\mathbf{\Lambda}$ *as* $u\to\infty$.

(b) if $\mathbf{t}=\mathbf{0}$ *is a local minimum, then* $\ddot{\mathbf{U}} \xrightarrow{\mathcal{L}} 2\mathbf{B}$ *as* $u\to 0$, *where* $\mathbf{B}\sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n+1)$.

Theorem 2.2 *Given that a* t *field* $T(\mathbf{t})$ *has a local maximum at* $\mathbf{0}$ *with height* t *at* $\mathbf{t}=\mathbf{0}$, *then as* $t\to\infty$,

$$-t^{-1}(1+t^2/m)^{-1}\ddot{\mathbf{T}} \xrightarrow{\mathcal{L}} A^{-1}\mathbf{B},$$

where $A\sim\chi^2_{m+1-N}$ *and* $\mathbf{B}\sim\mathrm{Wishart}_N(\mathbf{\Lambda}, m+1)$ *independently.*

Proof. By Lemma 5.1 of Worsley (1994), we can express the first and second derivatives of $T=T(\mathbf{t})$ in terms of independent random variables as follows:

$$\begin{aligned}
\dot{\mathbf{T}} &\stackrel{\mathcal{L}}{=} m^{\frac{1}{2}}(1+T^2/m)S^{-\frac{1}{2}}\mathbf{z}_1,\\
\ddot{\mathbf{T}} &\stackrel{\mathcal{L}}{=} m^{\frac{1}{2}}(1+T^2/m)S^{-1}\{-m^{-\frac{1}{2}}T(\mathbf{Q}-2\mathbf{z}_1\mathbf{z}'_1)-\mathbf{z}_1\mathbf{z}'_2-\mathbf{z}_2\mathbf{z}'_1+S^{\frac{1}{2}}\mathbf{H}\},
\end{aligned}\tag{2.5}$$

where $T\sim t_m$, $S\sim\chi^2_{m+1}$, $\mathbf{z}_1,\mathbf{z}_2\sim\mathrm{Normal}_N(\mathbf{0},\mathbf{\Lambda})$, $\mathbf{Q}\sim\mathrm{Wishart}_N(\mathbf{\Lambda},m-1)$ and $\mathbf{H}\sim\mathrm{Normal}_{N\times N}(\mathbf{0},\mathbf{M}(\mathbf{\Lambda}))$, all independently. Note that $\mathbf{0}$ is a local maximum of $T(\mathbf{t})$ of height t, then we can write

$$\ddot{\mathbf{T}} \stackrel{\mathcal{L}}{=} -t(1+t^2/m)S^{-1}\mathbf{Q}+m^{\frac{1}{2}}(1+t^2/m)S^{-\frac{1}{2}}\mathbf{H}.$$

Following the same arguments as in the proof of Theorem 6.3.1 of Adler (1981), page 134, we can see that as $t\to\infty$, $\ddot{\mathbf{T}}$ converges to the negative definite matrix $-t(1+t^2/m)S^{-1}\mathbf{Q}$. Under the regularity conditions, the joint HW conditional distribution of S and $\mathbf{Q}$, given that $T(\mathbf{t})$ has a local maximum at $\mathbf{0}$ with height t, is:

$$\begin{aligned}
\phi(s,\mathbf{q}\parallel T=t,\dot{\mathbf{T}}=\mathbf{0},\ddot{\mathbf{T}}<\mathbf{0}) &= \mu^{-1}\mathrm{E}(|\det(\ddot{\mathbf{T}}^-)|\,\Big|\, T=t,\dot{\mathbf{T}}=\mathbf{0},S=s,\mathbf{Q}=\mathbf{q})\phi(t,\mathbf{0},s,\mathbf{q})\\
&\to \mu^{-1}t^N(1+t^2/m)^N s^{-N}\det(\mathbf{q})\phi(t,\mathbf{0},s,\mathbf{q}) \text{ as } t\to\infty,
\end{aligned}$$

where $\phi(t,\dot{\mathbf{t}},s,\mathbf{q})$ is the joint density of $T,\dot{\mathbf{T}},S$ and $\mathbf{Q}$. By (2.5),

$$\phi(t,\mathbf{0},s,\mathbf{q})=\phi(t)\phi(s)\phi(\mathbf{q})\phi(\mathbf{0}|t,s),$$

and $\dot{\mathbf{T}}|T,S\sim\mathrm{Normal}_N(\mathbf{0},mS^{-1}(1+T^2/m)^2\mathbf{\Lambda})$ which implies

$$\phi(\mathbf{0}|t,s)=\{2\pi m(1+t^2/m)^2s^{-1}\}^{-\frac{N}{2}}\det(\mathbf{\Lambda})^{-\frac{1}{2}}\propto s^{\frac{N}{2}}.$$

Hence as $t\to\infty$,

$$\begin{aligned}
\phi(s,\mathbf{q}\parallel T=t,\dot{\mathbf{T}}=\mathbf{0},\ddot{\mathbf{T}}<\mathbf{0}) &\propto s^{-N}\det(\mathbf{q})\phi(s)\phi(\mathbf{q})s^{\frac{N}{2}}\\
\propto\ & s^{\frac{m+1-N}{2}-1}\exp(-s/2)\det(\mathbf{q})^{\frac{m+1-N-1}{2}}\exp(\mathrm{tr}(-\mathbf{\Lambda}^{-1}\mathbf{q}/2)),
\end{aligned}$$

which proves the theorem.

Theorem 2.3 *Given that an F field $F(\mathbf{t})$ has a local extremum at $\mathbf{0}$ with height f,*

(a) if $\mathbf{t}=\mathbf{0}$ is a local maximum, then $f^{-2}\ddot{\mathbf{F}} \xrightarrow{\mathcal{L}} (-2n/m)A^{-1}\mathbf{B}$ as $f\to\infty$, where $A\sim\chi^2_{m+n-N}$ and $\mathbf{B}\sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m+1)$ independently.

(b) if $\mathbf{t}=\mathbf{0}$ is a local minimum, then $\ddot{\mathbf{F}} \xrightarrow{\mathcal{L}} (2m/n)A^{-1}\mathbf{B}$ as $f\to 0$, where $A\sim\chi^2_{m+n-N}$ and $\mathbf{B}\sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n+1)$ independently.

2.2 Height

Let $M_z^+(Z,C)$ be the number of local maxima of the random field $Z(\mathbf{t})$ above z inside a compact subset C, and $M_z^-(Z,C)$ be the number of local minima below z inside C. As a simplification, sometimes we shall omit the superscript and use $M_z(Z,C)$ to indicate either $M_z^+(Z,C)$ or $M_z^-(Z,C)$. Worsley (1994) has derived the asymptotic expectations of $M_z(Z,C)$ for a χ^2, t or F field (Theorem 3.3, 5.3, 4.3). These results, aside from being of intrinsic interest, can also be used to give the limiting conditional distributions of the height of these random fields at local extrema.

Theorem 2.4

(a) Given that a χ^2 field $U(\mathbf{t})$, has a local maximum(or minimum) at $\mathbf{t}=\mathbf{0}$ with a height U exceeding (or below) u,

$$\lim_{u\to\infty} \mathrm{P}\{U>u+v \parallel U>u; \text{maximum}\} = \exp(-v/2) \quad \textit{for } v>0,$$

$$\lim_{u\to 0} \mathrm{P}\{U<u(1-v) \parallel U<u; \text{minimum}\} = (1-v)^{\frac{1}{2}(n-N)}$$

for $0<v<1$ and $n>N$.

(b) Given that a t field $T(\mathbf{t})$, $m>N$, has a local maximum at $\mathbf{t}=\mathbf{0}$ with a height T exceeding t, for $0<v<1$,

$$\lim_{t\to\infty} \mathrm{P}\left\{T>(1-v)^{-\frac{1}{2}}t \parallel T>t; \text{maximum}\right\} = (1-v)^{\frac{1}{2}(m-N)}.$$

(c) Given that an F field $F(\mathbf{t})$, $m>N$, has a local maximum at $\mathbf{t}=\mathbf{0}$ with a height F exceeding f, for $0<v<1$,

$$\lim_{f\to\infty} \mathrm{P}\left\{F>(1-v)^{-1}f \parallel F>f; \text{maximum}\right\} = (1-v)^{\frac{1}{2}(m-N)}.$$

Proof. We shall prove the result (b) for a t field as an illustration. Since $M_t^+(T,C)$ is derived from an ergodic point process,

$$\mathrm{P}\left\{T>(1-v)^{-\frac{1}{2}}t \parallel T>t; \text{maximum}\right\} = \frac{\mathrm{E}\left\{M^+_{(1-v)^{-\frac{1}{2}}t}(T,C)\right\}}{\mathrm{E}\{M_t^+(T,C)\}}.$$

Express $\mathrm{E}\{M_t^+(T,C)\}$ and $\mathrm{E}\{M^+_{(1-v)^{-\frac{1}{2}}t}(T,C)\}$ by Theorem 3.3 of Worsley (1994) and the result follows.

3 Structure of the random fields near local extrema

Having obtained the asymptotic curvature and height of a χ^2, t or F field $Z(\mathbf{t})$ at a local extremum, we shall now turn to a related study, the behaviour of $Z(\mathbf{t})$ in the vicinity of the local extremum of height z. That is to say, we shall investigate the *conditional field*

$$Z_z(\mathbf{t}) = Z(\mathbf{t}) \parallel \mathcal{A}, \tag{3.1}$$

where $\mathcal{A}$ denotes the event that $Z(\mathbf{t})$ has a local extremum with height z at $\mathbf{t} = \mathbf{0}$. The right tool is again the ergodic-sense Slepian model based on HW conditioning. This means we shall approach the conditioning event $\mathcal{A}$ by the series of events:

$$\mathcal{A}(h, h') = \{Z(\mathbf{t}) \text{ has an extremum of height in } (z, z+h) \text{ at a point in the ball } \|\mathbf{t}\| < h'\},$$

and represent the distribution of the conditional field $Z_z(\mathbf{t})$ by:

$$\mathrm{P}\{Z_z(\mathbf{t}) \in \mathcal{B}\} = \lim_{h \downarrow 0} \lim_{h' \downarrow 0} \mathrm{P}\{Z(\mathbf{t}) \in \mathcal{B} \mid \mathcal{A}(h, h')\},$$

where $\mathcal{B}$ is any set in the σ-field. Provided that $Z(\mathbf{t})$ is ergodic, the above probability can be written as a ratio of expectations. Lindgren (1972) and Adler (1981) gave the following lemma about the finite-dimensional distributions for the conditional field $Z_z(\mathbf{t})$.

Lemma 3.1 *Under the appropriate regularity conditions, the HW conditional distribution of* $\mathbf{v} = (Z(\mathbf{t}_1), \ldots, Z(\mathbf{t}_n))$*, given that* $Z(\mathbf{t})$ *has a local extremum at* $\mathbf{0}$ *with height* z*, has the density*

$$\beta_z(\mathbf{v}) = \frac{\int_D |\det \ddot{\mathbf{z}}| \phi(z, \mathbf{0}, \ddot{\mathbf{z}}, \mathbf{v}) d\ddot{\mathbf{z}}}{\int_D |\det \ddot{\mathbf{z}}| \phi(z, \mathbf{0}, \ddot{\mathbf{z}}) d\ddot{\mathbf{z}}}, \tag{3.2}$$

where D *is the region* $\{\ddot{\mathbf{z}} < \mathbf{0}\}$ *or* $\{\ddot{\mathbf{z}} > \mathbf{0}\}$*, depending on whether conditioning is on a maximum or minimum,* $\phi(z, \dot{\mathbf{z}}, \ddot{\mathbf{z}}, \mathbf{v})$ *is the joint density of* $Z, \dot{\mathbf{Z}}, \ddot{\mathbf{Z}}$ *and* $\mathbf{v}$*, and* $\phi(z, \dot{\mathbf{z}}, \ddot{\mathbf{z}})$ *is the joint density of* $Z, \dot{\mathbf{Z}}$ *and* $\ddot{\mathbf{Z}}$*. Furthermore, (3.2) is equivalent to*

$$\beta_z(\mathbf{v}) = \int_D \psi_z(\ddot{\mathbf{z}}) \phi(\mathbf{v} | z, \mathbf{0}, \ddot{\mathbf{z}}) d\ddot{\mathbf{z}}, \tag{3.3}$$

where

$$\psi_z(\ddot{\mathbf{z}}) = \frac{|\det \ddot{\mathbf{z}}| \phi(\ddot{\mathbf{z}} | z, \mathbf{0})}{\int_D |\det \ddot{\mathbf{z}}| \phi(\ddot{\mathbf{z}} | z, \mathbf{0}) d\ddot{\mathbf{z}}} \quad \text{if } \ddot{\mathbf{z}} \in D \tag{3.4}$$

and 0 *otherwise, is the HW conditional density of* $\ddot{\mathbf{Z}}$ *given that* $Z(\mathbf{t})$ *has an extremum at* $\mathbf{0}$ *with height* z*, and* $\phi(\mathbf{v} | z, \dot{\mathbf{z}}, \ddot{\mathbf{z}})$ *is the density of* $\mathbf{v}$ *conditional on* Z*,* $\dot{\mathbf{Z}}$ *and* $\ddot{\mathbf{Z}}$*.*

The re-expressing of $\beta_z(\mathbf{v})$ given by (3.3) and (3.4) provides some insight into the underlying problem. It is now possible to divide our job of finding $\beta_z(\mathbf{v})$ for a χ^2, t or F field into two parts. One is to find the form of $\psi_z(\ddot{\mathbf{z}})$, the HW conditional density of $\ddot{\mathbf{Z}}$, which is already attacked in Section 3, the other is to find the form of $\phi(\mathbf{v} | z, \mathbf{0}, \ddot{\mathbf{z}})$ which we shall now pursue. In general for non-Gaussian fields, the form of $\phi(\mathbf{v}|z, \mathbf{0}, \ddot{\mathbf{z}})$ is rather complex. Here we shall adopt the same approach as in Aronowich and Adler (1986, 1988) to deal with this difficulty. That is, we shall condition on the Gaussian components and then integrate them out. For a stationary Gaussian field $X(\mathbf{t})$ with mean 0 and variance 1, the distribution of $X(\mathbf{t})$ conditional on $X = x, \dot{\mathbf{X}} = \mathbf{0}, \ddot{\mathbf{X}} = \ddot{\mathbf{x}}$ at $\mathbf{0}$ can be found in Lindgren (1972) or Section 6.7 of Adler (1981). They have essentially shown the following.

Lemma 3.2 *Conditional on $X, \dot{\mathbf{X}}$ and $\ddot{\mathbf{X}}$ at $\mathbf{0}$, $X(\mathbf{t})$ has the same distribution as:*

$$X + \dot{\mathbf{X}}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\ddot{\mathbf{X}}\mathbf{t} + Mo(\|\mathbf{t}\|^2) + K(\mathbf{t}), \tag{3.5}$$

where $M = \max(|X|, \|\dot{\mathbf{X}}\|, \|\ddot{\mathbf{X}}\|)$ and $K(\mathbf{t})$ is normal with mean 0 and covariance $o(\|\mathbf{t}\|^4)$.

It is straightforward to conclude from the above lemma that the Gaussian random field $X(\mathbf{t})$ conditional on $X, \dot{\mathbf{X}}, \ddot{\mathbf{X}}$ at $\mathbf{0}$ can be represented by $X + \dot{\mathbf{X}}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\ddot{\mathbf{X}}\mathbf{t} + Mo(\|\mathbf{t}\|^2) + K(\mathbf{t})$, where $K(\mathbf{t})$ is a Gaussian random field with mean 0 and $K(\mathbf{t}) = o(\|\mathbf{t}\|^2)$ a.s., as easily follows from the fact that $K(\mathbf{t})$ is twice continuously differentiable.

We shall say that a sequence of random fields Y_u converges to a random field $Y^\star$ *uniformly* as $u \to a$ if, for each $h > 0$ and $\|\mathbf{t}\| < h$, we have

$$\mathrm{P}\left\{\lim_{u \to a} \sup_{\|\mathbf{t}\| < h} |Y_u(\mathbf{t}) - Y^\star(\mathbf{t})| = 0\right\} = 1.$$

With all these preparations, we are now ready to give the following theorems about the structure of a χ^2, t or F field near a local extremum. In particular, Theorem 3.3 extends previous results of Aronowich and Adler (1986, 1988) for a χ^2 field in one or two dimensions to any dimensions. Again we shall only prove the result for a t field as an illustration.

Theorem 3.3 *Assume $\mathbf{0}$ is a local extremum of a χ^2 field $U(\mathbf{t})$ with height u,*

(a) if $\mathbf{0}$ is a local maximum and $U_u(\mathbf{t})$ is the conditional field defined in (3.1), then as $u \to \infty$, $U_u(\mathbf{t}/\sqrt{u}) - u$ converges uniformly to the elliptic paraboloid $-\mathbf{t}'\mathbf{\Lambda}\mathbf{t}$;

(b) if $\mathbf{0}$ is a local minimum and $\bar{U}_u$ is the conditional field thus defined, then as $u \to 0$, $\bar{U}_u(\sqrt{u}\mathbf{t})/u$ converges uniformly to a random elliptic paraboloid $1 + \mathbf{t}'\mathbf{B}\mathbf{t}$, where $\mathbf{B} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n+1)$.

Theorem 3.4 *If $\mathbf{0}$ is a local maximum of a t field $T(\mathbf{t})$ with height t and $T_t(\mathbf{t})$ is the conditional t field defined in (3.1), then as $t \to \infty$,*

$$t^{-1}T_t(\mathbf{t}/t) \text{ converges uniformly to } \sqrt{\frac{mA}{mA + \mathbf{t}'\mathbf{B}\mathbf{t}}},$$

where $A \sim \chi^2_{m+1-N}$ and $\mathbf{B} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m+1)$ independently.

Proof. Suppose $X(\mathbf{t}), Y_i(\mathbf{t}), i = 1, \ldots, m$, are the component Gaussian fields in the definition of the t field $T(\mathbf{t})$ (2.2). Replacing each component Gaussian field by its representation in Lemma 3.2, we can write $X(\mathbf{t})^2$ and $V(\mathbf{t}) = \sum_{i=1}^m Y_i(\mathbf{t})^2$ as:

$$X(\mathbf{t})^2 = \left\{ X + \mathbf{t}'\dot{\mathbf{X}} + \frac{1}{2}\mathbf{t}'\ddot{\mathbf{X}}\mathbf{t} + Mo(\|\mathbf{t}\|^2) + K(\mathbf{t}) \right\}^2, \quad \text{where } K(\mathbf{t}) = o(\|\mathbf{t}\|^2),$$
$$V(\mathbf{t}) = \sum_{i=1}^m \left\{ Y_i + \mathbf{t}'\dot{\mathbf{Y}}_i + \frac{1}{2}\mathbf{t}'\ddot{\mathbf{Y}}_i\mathbf{t} + M_i o(\|\mathbf{t}\|^2) + K_i(\mathbf{t}) \right\}^2, \quad \text{where } K_i(\mathbf{t}) = o(\|\mathbf{t}\|^2).$$

Therefore, as $t \to \infty$,

$$X(\mathbf{t}/t)^2 = X^2 + \frac{2}{t}\mathbf{t}'X\dot{\mathbf{X}} + \frac{1}{t^2}\mathbf{t}'(X\ddot{\mathbf{X}} + \dot{\mathbf{X}}\dot{\mathbf{X}}')\mathbf{t} + \mathrm{error}_1 = \mathrm{main}_1 + \mathrm{error}_1 \tag{3.6}$$
$$t^2 V(\mathbf{t}/t) = t^2 V + t\mathbf{t}'\dot{\mathbf{V}} + \frac{1}{2}\mathbf{t}'\ddot{\mathbf{V}}\mathbf{t} + \mathrm{error}_2 = \mathrm{main}_2 + \mathrm{error}_2. \tag{3.7}$$

Define $R = T/\sqrt{m}$. By using the same notation as in (2.5) and following the proof of Lemma 5.1 of Worsley (1994), we have:

$$X^2 = \frac{SR^2}{1+R^2}, \quad X\dot{\mathbf{X}} = \frac{S^{\frac{1}{2}}R}{1+R^2}(\mathbf{z}_1 + R\mathbf{z}_2),$$
$$X\ddot{\mathbf{X}} + \dot{\mathbf{X}}'\dot{\mathbf{X}} = (1+R^2)^{-1}(\mathbf{z}_1 + R\mathbf{z}_2)(\mathbf{z}_1 + R\mathbf{z}_2)' - (1+R^2)^{-1}SR^2\mathbf{\Lambda} + (1+R^2)^{-\frac{1}{2}}S^{\frac{1}{2}}R\mathbf{H}_x,$$
$$V = \frac{S}{1+R^2}, \quad \dot{\mathbf{V}} = \frac{2S^{\frac{1}{2}}}{1+R^2}(-R\mathbf{z}_1 + \mathbf{z}_2),$$
$$\ddot{\mathbf{V}} = 2\left\{ \mathbf{Q} + (1+R^2)^{-1}(-R\mathbf{z}_1 + \mathbf{z}_2)(-R\mathbf{z}_1 + \mathbf{z}_2)' - (1+R^2)^{-1}S\mathbf{\Lambda} + (1+R^2)^{-\frac{1}{2}}S^{\frac{1}{2}}\mathbf{H}_v \right\}.$$

Substituting the above expressions into main_1 and main_2 in (3.6) and (3.7), it can be easily shown that as $t \to \infty$ and hence $R = t/\sqrt{m} \to \infty$,

$$\mathrm{main}_1 \to S, \quad \mathrm{main}_2 \to mS + \mathbf{t}'\mathbf{Q}\mathbf{t},$$

uniformly on $\|\mathbf{t}\| < h$. Furthermore, when $t \to \infty$, both of the error terms in (3.6) and (3.7) go to 0 uniformly on $\|\mathbf{t}\| < h$ by observing that the HW conditional distributions of the component Gaussian fields and their derivatives have the property that

$$X \sim O(1), \quad \dot{\mathbf{X}} \sim O(1), \quad \ddot{\mathbf{X}} \sim O(1),$$
$$Y_i \sim O(t^{-1}), \quad \dot{\mathbf{Y}}_i \sim O(1), \quad \ddot{\mathbf{Y}}_i \sim O(1).$$

Now the theorem follows from the fact that the HW conditional distribution of S and $\mathbf{Q}$ are distributed as χ^2_{m+1-N} and $\mathrm{Wishart}_N(\mathbf{\Lambda}, m+1)$ independently when $t \to \infty$ as proved in Theorem 2.2.

Theorem 3.5 *Assume* $\mathbf{0}$ *is a local extremum of an* F *field* $F(\mathbf{t})$ *with height* f.

(a) *If* $\mathbf{0}$ *is a local maximum and* $F_f(\mathbf{t})$ *is the conditional field defined in (3.1), then as* $f \to \infty$,

$$\frac{F_f(\mathbf{t}/\sqrt{f})}{f} \text{ converges uniformly to } \frac{mA}{mA + n\mathbf{t}'\mathbf{B}\mathbf{t}},$$

where $A \sim \chi^2_{m+n-N}$ *and* $\mathbf{B} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m+1)$ *independently.*

(b) *If* $\mathbf{0}$ *is a local minimum and* $\bar{F}_f(\mathbf{t})$ *is the conditional field defined in (3.1), then as* $f \to 0$,

$$\frac{\bar{F}_f(\sqrt{f}\mathbf{t})}{f} \text{ converges uniformly to } 1 + \frac{m}{n}\mathbf{t}'A^{-1}\mathbf{B}\mathbf{t},$$

where $A \sim \chi^2_{m+n-N}$ *and* $\mathbf{B} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n+1)$ *independently.*

It is interesting to see the difference of the behaviour of these non-Gaussian fields at local extrema from that of the Gaussian random field. At high level excursions, a χ^2 field shows a *deterministic* quadratic form, like that of a Gaussian field. However, at low level excursions, it has a *random* quadratic form. Unlike both a Gaussian and a χ^2 field, at high level excursions, it is the inverse of an F field that has a quadratic form. Similar behaviour of the square of a t field can be expected since it can be viewed as a special case of the F field with one numerator degree of freedom.

4 The size distribution of the connected components of the excursion set

4.1 The size of one connected component

By representing the field locally at a high or low level extremum by a simple process, the distribution of the size S of the connected component of the excursion set that contains the extremum can now be easily characterized for a χ^2, t and F field.

Let S_z^+ (or S_z^-) be the size of one connected component of the excursion set of $Z(\mathbf{t})$ above (or below) z. As a simplification, sometimes we drop superscript and use S_z to denote either S_z^+ or S_z^-. Let $a_N = \pi^{\frac{N}{2}}/\Gamma(\frac{N}{2}+1)$ be the Lebesgue measure of the N-dimensional unit ball. We have the following theorems about the limiting distributions of S_z for a χ^2, t or F field. These extend the previous result for Gaussian random fields by Nosko (1969), which states that to the power of $\frac{2}{N}$, S has an exponential distribution. We again prove the result only for a t field.

Theorem 4.1 *For a* χ^2 *field* $U(\mathbf{t})$,

(a) *when* $u \to \infty$, $u^{\frac{N}{2}} S_u^+ \xrightarrow{\mathcal{L}} a_N \det(\mathbf{\Lambda})^{-\frac{1}{2}} E^{\frac{N}{2}}$, *where* $E \sim \mathrm{Exponential}(2)$;

(b) *if* $n > N$, *then when* $u \to 0$, $u^{-\frac{N}{2}} S_u^- \xrightarrow{\mathcal{L}} a_N \det(\mathbf{\Lambda})^{-\frac{1}{2}} D^{\frac{N}{2}} \det(\mathbf{B}^\star)^{-\frac{1}{2}}$, *where* $D \sim \mathrm{Beta}(1, \frac{n-N}{2})$ *and* $\mathbf{B}^\star \sim \mathrm{Wishart}_N(\mathbf{I}, n+1)$ *independently.*

Theorem 4.2 *For a t field $T(\mathbf{t})$, $m > N$,*

$$t^N S_t^+ \xrightarrow{\mathcal{L}} a_N \det(\mathbf{\Lambda})^{-\frac{1}{2}} m^{\frac{N}{2}} D^{\frac{N}{2}} A^{\frac{N}{2}} \det(\mathbf{B}^\star)^{-\frac{1}{2}}$$

as $t \to \infty$, where $D \sim \mathrm{Beta}(1, \frac{m-N}{2})$, $\mathbf{B}^\star \sim \mathrm{Wishart}_N(\mathbf{I}, m+1)$ and $A \sim \chi^2_{m+1-N}$, all independently.

Proof. If $\mathbf{0}$ is a local maximum above t, then by Theorem 3.4, conditional on $T = T(\mathbf{0})$, with probability approaching one as $t \to \infty$, $T(\mathbf{t})$ has the following representation over the connected component B_t of the excursion set containing $\mathbf{0}$:

$$T(\mathbf{t})^2 = \frac{mAT^2}{mA + T^2\mathbf{t}'\mathbf{B}\mathbf{t}} + o(t^2). \tag{4.1}$$

Hence we can represent B_t as

$$B_t = \left\{\mathbf{t} : \frac{mAT^2}{mA + T^2\mathbf{t}'\mathbf{B}\mathbf{t}} + o(t^2) \geq t^2\right\} = \left\{\mathbf{t} : mA\frac{T^2 - t^2}{T^2} + o(1) \geq t^2\mathbf{t}'\mathbf{B}\mathbf{t}\right\}.$$

Note that the HW conditional distribution of $(T^2 - t^2)/T^2$ is

$$\mathrm{P}\left\{\frac{T^2 - t^2}{T^2} > v \parallel T > t; \text{maximum}\right\} = \mathrm{P}\left\{T > \frac{t}{\sqrt{1-v}} \parallel T > t; \text{maximum}\right\},$$

which has the limiting distribution given in Theorem 2.4 (b) independent of A and $\mathbf{B}$, then the theorem follows immediately by using Skorohod Theorem and the fact that Lebesgue measure is continuous with respect to a.s. convergence.

Theorem 4.3 *For an F field $F(\mathbf{t})$, $m > N$,*

$$f^{\frac{N}{2}} S_f^+ \xrightarrow{\mathcal{L}} a_N \det(\mathbf{\Lambda})^{-\frac{1}{2}} n^{-\frac{N}{2}} m^{\frac{N}{2}} D^{\frac{N}{2}} A^{\frac{N}{2}} \det(\mathbf{B}^\star)^{-\frac{1}{2}}$$

as $f \to \infty$, where $D \sim \mathrm{Beta}(1, \frac{m-N}{2})$, $\mathbf{B}^\star \sim \mathrm{Wishart}_N(\mathbf{I}, m+1)$ and $A \sim \chi^2_{m+n-N}$, all independently.

We can see some similarities between the limiting distribution of S_u^- for the χ^2 field and S_t^+, S_f^+ for the t or F field. This can be explained by the fact that when the t or F field goes to infinity, the χ^2 field $\sum Y_i^2(\mathbf{t})$, which appears in the denominator in the definition of the t or F field, goes to zero.

4.2 The size of the largest connected component

Let $A_z(Z, C)$ be the excursion set of the random field $Z(\mathbf{t})$ above or below level z inside a compact set C. Let $S_i, i = 1, \ldots, L$ be the size of the connected components in $A_z(Z, C)$, where L is the number of such connected components. We shall now derive an approximation of the distribution of the size $S_{\max}$ of the largest connected component of the excursion set using the Poisson clumping heuristic (Aldous, 1989). For high threshold z, the centers of these connected components in the excursion set can be regarded as a multidimensional point process with no 'memory'. For smooth stationary Gaussian fields, Adler (1981) gave a detailed treatment of this heuristic in Section 6.9 and more reference shall follow from there. Hence in the limit, L should follow a Poisson distribution and $S_i, i = 1, \ldots, L$ are independently identically distributed. This is to say that,

$$\text{(4.2)} \qquad \mathrm{P}(L = l) \approx \frac{\mathrm{E}(L)^l}{l!} e^{-\mathrm{E}(L)},$$

and

$$\text{(4.3)} \qquad \begin{aligned} \mathrm{P}(S_{\max} \leq s \mid L \geq 1) &= \sum_{l=1}^{\infty} \mathrm{P}(S_i \leq s, 1 \leq i \leq l, L = l \mid L \geq 1) \\ &\approx \frac{1}{\mathrm{P}(L \geq 1)} \sum_{l=1}^{\infty} \frac{\mathrm{E}(L)^l}{l!} e^{-\mathrm{E}(L)} \mathrm{P}(S \leq s)^l \\ &= \frac{\exp\{-\mathrm{E}(L)\mathrm{P}(S \geq s)\} - \exp\{-\mathrm{E}(L)\}}{1 - \exp\{-\mathrm{E}(L)\}}. \end{aligned}$$

Expression (4.3) relates the distribution of $S_{\max}$ to the distribution of S, whose limiting distributions has been derived for χ^2, t and F fields in the previous section. In the extreme case when $\mathrm{E}(L)$ approaches 0, $\mathrm{P}(S_{\max} \leq x) \to \mathrm{P}(S \leq x)$ and hence $S_{\max}$ has the same distribution as S.

Although there are no available closed form expressions for $\mathrm{E}(L)$ except in the one dimensional case, we can still approximate $\mathrm{E}(L)$ accurately by using the following argument from Hasofer (1978) and Adler (1981). Let $\chi(A_z(Z, C))$ be the Euler characteristic of the excursion set $A_z(Z, C)$, then $\mathrm{E}(L)$, $\mathrm{E}(\chi(A_z(Z, C)))$ and $\mathrm{E}(M_z(Z, C))$ are asymptotically the same at high (or low) levels. That is,

$$\text{(4.4)} \qquad \mathrm{E}(L) \approx \mathrm{E}(\chi(A_z(Z, C))) \approx \mathrm{E}(M_z(Z, C)).$$

In practice, we favor $\mathrm{E}(\chi(A_z(Z, C)))$ as an approximation to $\mathrm{E}(L)$ since an exact formula for $\mathrm{E}(\chi(A_z(Z, C)))$ usually can be obtained, whereas only asymptotic results are available for $\mathrm{E}(M_z(Z, C))$. The expression for $\mathrm{E}(\chi(A_z(Z, C)))$ for the Gaussian field can be found in Adler (1981) and for χ^2, t and F fields in Worsley (1994).

We shall now mention the approach adopted by Friston *et al.* (1994) to approximate the distribution of $S_{\max}$ for non-Gaussian random fields, especially for the t field. If $T(\mathbf{t})$ is a t random field with m degrees of freedom, then $T(\mathbf{t})$ is first Gaussianized. The *Gaussianized t* field is defined as $W(\mathbf{t}) = \Phi^{-1}(\Psi(T(\mathbf{t})))$, where Φ and Ψ are the cumulative density functions for normal and t distribution t_m, respectively. Obviously, the excursion set of $T(\mathbf{t})$ above level t corresponds to the excursion set of the Gaussianized field $W(\mathbf{t})$ above level w, where $w = \Phi^{-1}(\Psi(t))$. By treating the Gaussianized field $W(\mathbf{t})$ as a Gaussian field, Friston *et al.* (1994) suggested approximating the distribution of $S_{\max}$ using the result available for the

Gaussian field. To improve the accuracy of the approximation, Worsley *et al.* (1992) also suggested a correction factor to account for the fact that the resulting Gaussianized field is slightly rougher than the original t field. However, the Gaussianized field is *not* a Gaussian field since its finite-dimensional distribution does not have a multi-dimensional Gaussian form. This is a serious problem especially for a t field with low degrees of freedom and hence this approach should only be used when the degrees of freedom of the t field is high enough.

As in Friston *et al.* (1994), an improved approximation of S and $S_{\max}$ can be obtained using what Aldous (1989) referred as *the fundamental identity* to correct for the mean of S:

$$\mathrm{E}(L)\mathrm{E}(S) = \mathrm{E}(|A_z(Z, C)|) = |C|F_z(-z), \tag{4.5}$$

where $F_z(\cdot)$ is the cumulative density function. [1]

4.3 Validation

We present some $N = 2$ dimensional simulations to validate the distributions of L, S and $S_{\max}$ for χ^2, t and F fields obtained in previous sections. We first generated the component Gaussian fields by convolving Gaussian white noise with a Gaussian kernel and then generated the χ^2, t or F field according to their definitions. In each realization, an excursion set was obtained by thresholding the χ^2, t or F field at a prescribed level. Then the number of connected components in the excursion set, L, the size of each connected component, S, and the size of the largest connected component, $S_{\max}$ were recorded. The number of realizations was chosen to be 5000. Simulations used random images with 256×256 pixels and a Gaussian kernel with FWHM=9.2. FWHM, or *Full Width at Half Maximum*, the width of the kernel at half its maximum height, is commonly used in imaging to measure the spread of a kernel. For a Gaussian kernel, FWHM is $\sqrt{8 \log 2}\times$ standard deviation. Using this data we compared the empirical estimates and theoretical approximations for $\mathrm{P}(L = l), \mathrm{P}(S = s)$ and $\mathrm{P}(S_{\max} > s)$ obtained in previous sections. Note all the theoretical results are in the limit sense.

Figure 2 shows the simulation result for a t field with $m = 7$ and threshold $t = 4.75$. Figure 3-4 is for a χ^2 field with $n = 7$ and threshold $u = 22.04, 0.7945$ respectively, and Figure 5 is for an F field with $n = 4$ numerator and $m = 5$ denominator degrees of freedom, and threshold $f = 21.05$. The thresholds were chosen so that the tail probabilities of the corresponding distributions are 0.001 for the t field and 0.025 for the rest. In all the figures, the solid lines are empirical estimates and the dotted lines are theoretical results. The agreement is evident. For the t field in Figure 2, the dashed lines represent the distributions from the Gaussianization approach of Friston *et al.* (1994). The theoretical P-value is very close to the simulated P-value, but the Gaussianized P-value is too conservative, especially when the degrees of freedom of the t field (7 in this case) is small. For example, the P-value

[1] A mean correction is always recommended for any practical use of the asymptotic formulae derived in the paper. Magnitude of the correction could vary significantly from case to case.

for $S_{\max} = 50$ was estimated as 0.054 from simulation. The theoretical P-value from (4.3) is 0.065, whereas that derived from the Gaussianized field is 0.216. Even larger discrepancies are expected for the Gaussianization approach for a t field in higher dimensions.

As mentioned in Friston *et al.* (1994), the small differences between the empirical and theoretical distributions may derive from a number of sources: (i) many of our approximations are only asymptotically true, in the limit of high or low thresholds; (ii) the effect of sampling the random field on a lattice of voxels; (iii) the edge effects of our simulated processes.

5 Application to positron emission tomography images

The PET technique uses positron-emitting labelled carriers, created in an on-site cyclotron, to produce an image of brain activity such as glucose utilization, oxygen consumption or blood flow. An increasingly important application of PET involves the measurement of regional CBF (cerebral blood flow) during the performance of various sensorimotor or cognitive tasks. Regional changes in CBF between two tasks reflect the mobilization of functional units specific to the new task. The purpose of these studies is to identify the regions activated by the stimulus. Blood flow images are obtained both under a baseline and a stimulus condition, and the two images are subtracted. Those regions that have been activated should show a change in blood flow. In many cognitive studies, the activation is so slight that the experiment must be repeated on several subjects, and the subtracted images are averaged to improve the signal-to-noise ratio. Typically, the number of subjects in the experiment is small. A detailed description of such an experiment can be found in Worsley *et al.* (1992) and Worsley (1994).

Suppose there are p subjects and that for the ith subject, $X_i(\mathbf{t}), i = 1, \ldots, p$ is the value of the differenced image at point $\mathbf{t} \in C$, where $C \subset \mathbb{R}^3$ is the brain. Let $\mu(\mathbf{t}) = \mathrm{E}(X_i(\mathbf{t}))$ be the mean change in blood flow, or mean activation, between the two conditions, and let $\sigma(\mathbf{t})^2 = \mathrm{Var}(X_i(\mathbf{t})), i = 1, \ldots, p$. It was assumed that the standardized images, $(X_i(\mathbf{t}) - \mu(\mathbf{t}))/\sigma(\mathbf{t}), i = 1, \ldots, p$, are i.i.d. stationary Gaussian fields with zero mean and unit variance. To test the null hypothesis of no activation, $\mu(\mathbf{t}) = 0$, the sample mean $\bar{X}(\mathbf{t})$ and sample variance $V(\mathbf{t})$ of the differenced images over all subjects were calculated:

$$\bar{X}(\mathbf{t}) = \sum_{i=1}^{p} X_i(\mathbf{t})/p \quad \text{and} \quad V(\mathbf{t}) = \sum_{i=1}^{p} (X_i(\mathbf{t}) - \bar{X}(\mathbf{t}))^2/(p-1). \tag{5.1}$$

If the underlying variance was stationary, then as proposed by Worsley *et al.* (1992), a Gaussian random field can be used to test for regional changes using the pooled sample variance. However, if the stationarity assumption is in doubt and the variance could be arbitrary, Worsley (1994) proposed the t random field $T(\mathbf{t})$ given by:

$$T(\mathbf{t}) = p^{\frac{1}{2}} \bar{X}(\mathbf{t})/V^{\frac{1}{2}}(\mathbf{t}), \tag{5.2}$$

to test for high regional changes instead.

Typically, in these studies, the global maximum of the random field is chosen as the test statistic. However, the size of the largest connected component of the excursion set above a high threshold was suggested as an alternative, based on the assumption that areas of true neural activity will tend to stimulate blood flow changes over contiguous voxels. If the underlying variance is stationary so that a Gaussian field is used, then we can apply the results in Friston *et al.* (1994) to assess the significance of the activated regions by their spatial extents. However, if the variance is not stationary and a t random field is used instead, then results obtained in previous sections are used to assess the significance instead.

Experiments were carried out to investigate the neural activity in response to a painful heat stimulus (Talbot *et al.*, 1991). $p = 9$ subjects underwent a procedure wherein a thermistor was applied to the forearm at both warm and painful hot states. The null hypothesis is that there is no difference in brain activity between the two states.

The stationarity of the variance was examined by the sample variance image in (5.1), which under the stationary assumption, should be a χ^2 random field with $p - 1$ degrees of freedom. Evidence against the stationarity was found using the global maximum as the test statistic, which was reported in Worsley (1994). Hence, the t field calculated by (5.2) was used to test for regions of activation.

To increase the power of detecting possible changes, we smoothed the individual differenced images $X_i(\mathbf{t})$ for the 9 subjects to have FWHM = 11.6mm. The t field was calculated using the smoothed $X_i(\mathbf{t})$ and then thresholded at level 5 to search for possible activations. Such an excursion set from the treatment experiment is shown in Figure 1.

Throughout the analysis 1000cm^3 was used as the volume of the whole brain. To facilitate comparison, we applied both our results and the Gaussianization approach adopted by Friston *et al.* (1994) to assess the significance of the observed signals. The threshold for the t field, chosen as 5, corresponds to a threshold of 3.28 for the Gaussianized t field. This gives 3.38 as approximation of the expected number of connected regions $\mathrm{E}(L)$, by treating the Gaussianized t field as a Gaussian field, as opposed to 10.89 using the Euler characteristic of the t field. Four areas shown in Figure 1 were found to be significant at level 0.1 in the treatment experiment, while none were found in the control experiment as expected. These four activated regions, numbered 1 to 4 by their sizes, are listed in Table 1 along with their approximate P-values from the two approaches.

From the table, we can see that for the same activated area, our result gives a smaller P-value than the Gaussianization approach. Hence an improvement in the power of detection can be expected. For example, at the 0.05 level, cluster 3 is detected by our method but not by the Gaussianized method. It is also worth noting that if the global maximum of the t field, $T_{\max}$, was used as the test statistic, then as reported in Worsley (1994), no significant activations were detected. This suggests that the global maximum as a test statistic is not powerful enough in this case.

Table 1: Approximate P-values of observed cluster sizes in the excursion set of the t field from the experiment for pain perception.

cluster	size	P-value	
		t field	Gaussianized t field
		threshold 5, E(L)=10.89	threshold 3.28, E(L)=3.38
1	4.74 cm^3	0.000	0.000
2	1.21 cm^3	0.005	0.029
3	0.88 cm^3	0.015	0.071
4	0.44 cm^3	0.098	0.258

Acknowledgement

I would like to thank K. J. Worsley for raising the question which initiated the research and for many discussions which advanced my understanding considerably, A.C. Evans of the Brain Imaging Centre of the Montreal Neurological Institute for his continuous encouragement, and Dr. C. Bushnell and Dr. G. Duncan for permission to use their data.

References

Adler, R.J. (1981). *The Geometry of Random Fields.* Wiley, New York.

Aldous, D. (1989). *Probability approximations via the Poisson Clumping Heuristic.* Springer-Verlag, New York.

Aronowich, M. and Adler, R.J. (1986). Extrema and level crossings of χ^2 processes. *Advances in Applied Probability*, **18**, 901-920.

Aronowich, M. and Adler, R.J. (1988). Sample path behaviour of χ^2 surfaces at extrema. *Advances in Applied Probability*, **20**, 719-738.

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.

Hasofer, A.M. (1978). Upcrossings of random fields. *Supplement to Advances in Applied Probability,* **10**, 14-21.

Kac, M. and Slepian, D. (1959). Large excursions of Gaussian processes. *Annals of Mathematical Statistics*, **30**, 1215-1228.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremal and Related Properties of Stationary Processes.* Springer-Verlag, New York.

LINDGREN, G. (1972). Local maxima of Gaussian fields. *Arkiv for Mathematik*, **10**, 195-218.

LINDGREN, G. (1984). Use and structure of Slepian model processes for prediction and detection in crossing and extreme value theory. In *Statistical Extremes and Applications*, Ed. J. Tiago de Oliveira. Reidel, Dordrecht.

NOSKO, V.P. (1969). Local structure of Gaussian random fields in the vicinity of high-level light sources. *Soviet Mathematics Doklady*, **10**, 1481-1484.

TALBOT, J.D., MARRETT, S., EVANS, A.C., MEYER, E., BUSHNELL, M.C. AND DUNCAN, G.H. (1991). Multiple representations of pain in human cerebral cortex. *Science*, **251**, 1355-1358.

WORSLEY, K.J., EVANS, A.C., MARRETT, S. AND NEELIN, P. (1992). A three dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, **12**, 900-918.

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.