# Local maxima and the expected Euler characteristic of excursion sets of $\chi^2$, $F$ and $t$ fields

K.J. Worsley

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6

**Key words**: Local maxima; Euler characteristic; Gaussian random fields; image analysis.

SUMMARY

The maximum of a Gaussian random field was used by Worsley et al. (1992) to test for activation at an unknown point in positron emission tomography images of blood flow in the human brain. The Euler characteristic of excursion sets was used as an estimator of the number of regions of activation. The expected Euler characteristic of excursion sets of stationary Gaussian random fields has been derived by Adler and Hasofer (1976) and Adler (1981). In this paper we extend the results of Adler (1981) to $\chi^2$, $F$ and $t$ fields. The theory is applied to some three dimensional images of cerebral blood flow from a study on pain perception.

---

This work was supported by the Natural Sciences and Engineering Research Council of Canada, and the Fonds pour la Formation des Chercheurs et l'Aide à la Recherche de Québec. The author would like to acknowledge the encouragement and assistance of A.C. Evans, S. Marrett and P. Neelin of the Brain Imaging Centre of the Montreal Neurological Institute. The author would like to thank Dr. C. Bushnell and Dr. G. Duncan for permission to use their data.

# 1 Introduction

Many studies of brain function with positron emission tomography (PET) involve the interpretation of subtracted PET images, usually the difference between two three-dimensional images of cerebral blood flow under baseline and stimulation conditions. The purpose of these studies is to see which areas of the brain show an increase in blood flow, or 'activation', due to the stimulation condition. The experiment is repeated on several subjects, and the subtracted images are averaged to improve the signal to noise ratio. The averaged image is standardized to have unit variance and then searched for local maxima, which might indicate points in the brain that are activated by the stimulus. The main statistical problem has been to assess the significance of these local maxima.

Worsley et al. (1992) have shown that the averaged image can be modelled as a Gaussian random field with a covariance function depending on the known resolution of the PET camera. They have used the Euler characteristic of excursion sets of this Gaussian random field as an estimator of the number of regions of activation. The excursion set inside a fixed set $C$ is just the set of points where the field exceeds a fixed threshold value. Adler (1981) defines the DT (differential topology) characteristic of the excursion set in such a way that it equals the Euler characteristic when the excursion set does not touch the boundary of $C$. For excursion sets above high threshold values the DT characteristic approximates the number of local maxima above the threshold. The importance of the DT characteristic, as opposed to the number of local maxima, is that it is more amenable to statistical analysis. There is no known result for the expected number of local maxima above a threshold, but Adler and Hasofer (1976) and Adler (1981) derived a simple expression for the expected DT characteristic of excursion sets of stationary Gaussian random fields. When no activation is present in the PET image, Worsley et al. (1992) show that the observed DT characteristic is close to the expected DT characteristic.

Similar problems arise in astrophysics. Hamilton, Gott and Weinberg (1986) and more recently Beaky, Scherrer, and Villumsen (1992) have applied methods similar to those discussed in this paper to study the density of matter in the universe. Gott, Park, Juskiewicz, Bies, Bennett, Bouchet and Stebbins (1986) have used similar tools to study the fluctuations in the cosmic microwave background which were recently discovered by Smoot et al. (1992).

Throughout their analyses, Worsley et al. (1992) have assumed that the variance of the image is stationary, thus enabling them to pool the sample variance over all values in the image. There has been some doubt about this assumption, and some workers have suggested using a local standard deviation rather than a pooled standard deviation to normalise the image, thus producing an image of $t$-statistics.

The purpose of this paper is to extend the work of Adler (1981) to derive the expected Euler characteristic of excursion sets of such non-Gaussian random fields. In section two we shall review the work of Adler and Hasofer (1976) and Adler (1981). In section three we shall extend a result of Adler (1981) for a $\chi^2$ field in two dimensions to higher dimensions.

In section four we shall consider the $F$ field and we shall use this to derive results for a $t$ field in section 5. Finally in section six we shall apply this work to some PET images from a study in pain perception.

# 2 Local maxima and the DT characteristic

## 2.1 The DT Characteristic

Let $Z = Z(\mathbf{t})$, $\mathbf{t} = (t_1, \ldots, t_N)' \in \mathbb{R}^N$, be a homogeneous real-valued random field and let $C$ be a compact subset of $\mathbb{R}^N$. The excursion set of $Z(\mathbf{t})$ inside $C$ above the level $z$ is defined as $A_z(Z, C) = \{\mathbf{t} \in C : Z(\mathbf{t}) \geq z\}$. Throughout this paper we shall denote derivatives with respect to $t_k$ by a superscript $^{(k)}$ and second order derivatives with respect to $t_k$ and $t_l$ by superscript $^{(kl)}$. Thus we denote $Z^{(k)} = Z^{(k)}(\mathbf{t}) = \partial Z/\partial t_k$ and $Z^{(kl)} = Z^{(kl)}(\mathbf{t}) = \partial^2 Z/\partial t_k \partial t_l$, $k, l = 1, \ldots, N$. Let $\mathbf{D}_{N-1} = \mathbf{D}_{N-1}(\mathbf{t})$ be the $(N-1) \times (N-1)$ matrix of second order partial derivatives of $Z(\mathbf{t})$, with $(k, l)$ element $Z^{(kl)}$, $k, l = 1, \ldots, N-1$. Under suitable regularity conditions on $Z(\mathbf{t})$, Adler (1981), page 90, defines the DT (differential topology) characteristic of $A = A_z(Z, C)$ as

$$\chi(A) = (-1)^{N-1} \sum_{j=0}^{N-1} (-1)^j \chi_j(A),$$

where $\chi_j(A)$ is the number of points $\mathbf{t} \in C$ satisfying the conditions: (a) $Z(\mathbf{t}) = z$, (b) $Z^{(1)}(\mathbf{t}) = 0, \ldots, Z^{(N-1)}(\mathbf{t}) = 0$, (c) $Z^{(N)}(\mathbf{t}) > 0$, and (d) the number of negative eigenvalues of $\mathbf{D}_{N-1}(\mathbf{t})$ is exactly $j$.

It can be shown that provided the excursion set does not touch the boundary of the region $C$ then $\chi(A)$ is the Euler, or Euler-Poincaré characteristic of the excursion set, and $\chi(A)$ is thus invariant under rotations of the coordinate system. Roughly speaking, it counts the number of connected components of the excursion set, minus the number of 'holes'. An illustration of the DT characteristic of the excursion set of an artificial two-dimensional image is shown in Figure 1(a,b). As the threshold level $z$ increases Adler (1981) shows that the holes tend to disappear and that we are left with isolated regions each of which contains just one local maximum (Figure 1(c)). Thus for large $z$ the presence of holes is a rare occurrence and the DT characteristic approximates the number of local maxima of $Z(\mathbf{t})$ above $z$ inside $C$, denoted by $M_z^+(Z, C)$. For even larger $z$ near the global maximum of $Z(\mathbf{t})$ inside $C$, denoted by $Z_{\max}$, the DT characteristic takes the value 0 if $Z_{\max} < z$ and 1 if $Z_{\max} \geq z$ (Figure 1(d)). Hasofer (1978) uses this approach to show that

$$\mathrm{P}(Z_{max} \geq z) = \mathrm{P}(M_z^+(Z, C) \geq 1) \leq \mathrm{E}(M_z^+(Z, C)) \approx \mathrm{E}(\chi(A))$$

as $\mathrm{P}(M_z^+(Z, C) > 1) \to 0$ for $x \to \infty$, and so the expected DT characteristic approximates the exceedence probability of $Z_{\max}$.

## 2.2 Expectations

The importance of $\chi(A)$, as opposed to $M_z^+(Z,C)$, is that despite its complex definition it is more amenable to statistical analysis. Only asymptotic results are known for the expectation of $M_z^+(Z,C)$, but for the DT characteristic $\chi(A)$ Adler and Hasofer (1976) and Adler (1981) obtained an expression for the expectation of $\chi(A)$ in terms of the second-order derivatives of $Z$.

We shall need the following notation. For any scalar $d$, let $d^+ = d$ if $d > 0$ and zero otherwise. For any symmetric matrix $\mathbf{D}$, let $\mathbf{D}^- = \mathbf{D}$ if $\mathbf{D}$ is negative definite and zero otherwise. Let $\mathbf{D}_N = \mathbf{D}_N(\mathbf{t})$ be the $N \times N$ matrix of all second order partial derivatives of $Z(\mathbf{t})$, with $(k,l)$ element $Z^{(kl)}$, $k,l = 1,\ldots,N$. Let $\lambda(C)$ be the Lebesgue measure of $C$. Finally, we define the *moduli of continuity* of $Z^{(k)}$ and $Z^{(kl)}$ inside $C$ to be:

$$\omega_k(h) = \sup_{\|\mathbf{t}-\mathbf{s}\|<h} \left| Z^{(k)}(\mathbf{t}) - Z^{(k)}(\mathbf{s}) \right|, \quad \omega_{kl}(h) = \sup_{\|\mathbf{t}-\mathbf{s}\|<h} \left| Z^{(kl)}(\mathbf{t}) - Z^{(kl)}(\mathbf{s}) \right|,$$

where the supremum is taken over all $\mathbf{t}, \mathbf{s} \in C$, $k,l = 1,\ldots,N$. Then Theorem 5.2.1, page 105, and Theorem 6.1.1, page 123, of Adler(1981) give the following results, which we shall state here, under slightly different conditions, for future reference:

**Theorem 2.1** *Assume (i) that for any $\epsilon > 0$*

$$P\left(\max_{k,l}\{\omega_k(h), \omega_{kl}(h)\} > \epsilon\right) = o(h^N) \quad \text{as } h \downarrow 0.$$

*If (ii) all the second order partial derivatives $Z^{(kl)}$ have finite variances conditional on $Z, Z^{(1)}, \ldots, Z^{(N)}$, and the density $\theta_N(z, z_1, \ldots, z_N)$ of $Z, Z^{(1)}, \ldots, Z^{(N)}$ is bounded above, then the expectation of $M_z^+(Z,C)$ is*

$$\mathrm{E}(M_z^+(Z,C)) = \lambda(C) \int_z^\infty \mathrm{E}\{-\det(\mathbf{D}_N^-) | Z = y, Z^{(1)} = 0, \ldots, Z^{(N)} = 0\} \theta_N(y, 0, \ldots, 0)\, dy.$$

*If (iii) the second order partial derivatives $\{Z^{(kl)},\ 1 \le k \le N,\ 1 \le l \le N-1\}$, and $Z^{(N)}$ have finite variances conditional on $Z, Z^{(1)}, \ldots, Z^{(N-1)}$, and the density $\theta_{N-1}(z, z_1, \ldots, z_{N-1})$ of $Z, Z^{(1)}, \ldots, Z^{(N-1)}$ is bounded above, and provided $C$ is a convex subset of $\mathbb{R}^N$, then the expectation of $\chi(A_z(Z,C))$ is*

$$\mathrm{E}(\chi(A_z(Z,C))) = \lambda(C)(-1)^{N-1}\mathrm{E}\{Z^{(N)+}\det(\mathbf{D}_{N-1}) | Z = z, Z^{(1)} = 0, \ldots, Z^{(N-1)} = 0\}$$

$$\theta_{N-1}(z, 0, \ldots, 0).$$

*Proof.* The conditions (ii) and (iii) given here replace those given by Lemma 5.2.1 of Adler(1981), page 98, which require finite variances of $Z^{(k)}$ conditional on $Z$ and $Z^{(kl)}$, and bounds on the density of $Z^{(kl)}$ conditional on $Z$. It is straightforward to re-work the proof of Lemma 5.2.1 by taking expectations over $Z^{(kl)}$ conditional on $Z$ and $Z^{(k)}$, before applying the dominated convergence theorem and Fatou's Lemma to bound the integral. □

## 2.3 Comments

There is no simple relationship between the DT characteristic of an excursion set $A_z(Z,C)$ and the DT characteristic of its 'complement' $A_{-z}(-Z,C)$, even though their union is $C$ and their intersection is the contour $Z(\mathbf{t}) = z$. However if the field is homogeneous then there is a simple relationship between the expectations of the DT characteristics. We shall need this result in section five:

**Corollary 2.2** *If $Z(\mathbf{t})$ satisfies the conditions of Theorem 2.1 then*

$$\mathrm{E}(\chi(A_{-z}(-Z,C))) = (-1)^{N-1}\mathrm{E}(\chi(A_z(Z,C)))$$

*Proof.* Using the same notation as in Theorem 2.1 we have

$$\mathrm{E}(\chi(A_{-z}(-Z,C))) = \lambda(C)(-1)^{N-1}\mathrm{E}\{(-Z^{(N)})^{+}\det(-\mathbf{D}_{N-1})|Z = z, Z^{(1)} = 0, \ldots, Z^{(N-1)} = 0\}$$

$$\theta_{N-1}(z, 0, \ldots, 0).$$

Since $Z(\mathbf{t})$ is homogeneous then $-Z^{(N)} = \partial Z/\partial(-t_N)$ has the same distribution as $Z^{(N)}$. Since $\det(-\mathbf{D}_{N-1}) = (-1)^{N-1}\det(\mathbf{D}_{N-1})$, the result follows. $\square$

The main importance of Theorem 2.1 is that it allows us to calculate the mean number of local maxima and the mean DT characteristic of excursion sets from the joint density of the field $Z(\mathbf{t})$ and its partial derivatives up to second order. Adler and Hasofer (1976) applied this result to a homogeneous Gaussian random field. Worsley et al. (1993) have extended this to a non-homogeneous Gaussian random field in three dimensions. Unfortunately the awkwardness of $\mathbf{D}_N^-$ prevents us from obtaining an exact expression for the mean number of local maxima above $z$, except in some special cases, but Adler(1981), Theorem 6.3.1, page 133, gives an asymptotic result for large values of $z$. On the other hand, Adler (1981), Theorem 5.3.1, page 111, derived an exact expression for the mean DT characteristic for a Gaussian random field with zero mean and unit variance.

Before embarking on the extension of these results to $\chi^2$, $F$ and $t$ fields, it may be worth mentioning the main steps. Results are first obtained for the $F$ field, and results for the $\chi^2$ and $t$ fields are derived as special cases. The method of proof hinges on a representation of the $F$ field and its derivatives up to second order in terms of independent random variables. We first derive such a representation for the Gaussian field (Lemma 3.1), then the $\chi^2$ field (Lemma 3.2), and finally the $F$ field (Lemma 4.1); a similar result is given for the $t$ field without proof (Lemma 5.1). The sufficient regularity of the $F$ field then follows (Lemma 4.2). We are then ready to apply Theorem 2.1 to find the mean number of local maxima greater than a high threshold level for the $F$ field (Theorem 4.1); similar results for the $\chi^2$ and $t$ fields then follow (Theorems 3.3 and 5.2). An asymptotic expression for the distribution of the global maximum is a simple corollary to these results. Before tackling the DT characteristic, we must establish some lemmata on the expected determinants of linear combinations of

Wishart matrices and a normally distributed matrix with the same invariance properties as a Wishart matrix (Lemmata 7.1 to 7.5 in the Appendix). After this we are ready for our final results on the DT characteristic of the $F$ field (Theorem 4.6) which is immediately applied to the $\chi^2$ and $t$ fields (Theorems 3.5 and 5.4).

# 3 The $\chi^2$ field

## 3.1 Definition

Let $X_1(\mathbf{t}), \ldots, X_n(\mathbf{t}),\ \mathbf{t} \in \mathrm{I\!R}^N$, be independent, identically distributed, homogeneous, real-valued Gaussian random fields with zero mean and unit variance. Then Adler (1981), page 169, defines the $\chi^2$ field $U(\mathbf{t})$ as

$$U(\mathbf{t}) = \sum_{i=1}^{n} X_i(\mathbf{t})^2,\ \mathbf{t} \in \mathrm{I\!R}^N.$$

Clearly the marginal distribution of $U(\mathbf{t})$ at each $\mathbf{t}$ is $\chi^2$ with $n$ degrees of freedom. Adler (1981), page 169, notes that a chi-squared field is twice as 'rough' as its component Gaussian fields, in the sense that the variance of its partial derivatives relative to its variance is twice as great as that of its component fields.

## 3.2 Representations of derivatives

We shall need the following two lemmata, the first proved by Adler (1981), page 114, which give a representation of the first and second derivatives of a Gaussian and a $\chi^2$ field in terms of independent random variables. We shall use the notation $\mathrm{Normal}_d(\mu, \mathbf{\Sigma})$ to represent the multivariate normal distribution on $\mathrm{I\!R}^d$ with mean $\mu$ and variance $\mathbf{\Sigma}$, $\chi^2_\nu$ to represent the $\chi^2$ distribution with $\nu$ degrees of freedom, and $\mathrm{Wishart}_d(\mathbf{\Sigma}, \nu)$ to represent the Wishart distribution of a $d \times d$ matrix with expectation $\nu\mathbf{\Sigma}$ and degrees of freedom $\nu$.

Let $\mathbf{\Lambda} = \mathrm{Var}(\partial X_i(\mathbf{t})/\partial \mathbf{t})$ be the $N \times N$ variance-covariance matrix of the partial derivatives of $X_i(\mathbf{t})$ with $(k,l)$ element $\lambda_{kl} = \mathrm{Cov}(X_i^{(k)}, X_i^{(l)})$, $k, l = 1, \ldots, N$, $i = 1, \ldots, n$.

**Lemma 3.1** *Let $X(\mathbf{t}) = X_i(\mathbf{t})$ for any $i = 1, \ldots, n$, then*

*(a)* $\dfrac{\partial X}{\partial \mathbf{t}} \sim \mathrm{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$ *independent of $X$ and* $\dfrac{\partial^2 X}{\partial \mathbf{t} \partial \mathbf{t}'}$,

*(b) Conditional on $X$,*

$$\frac{\partial^2 X}{\partial \mathbf{t} \partial \mathbf{t}'}\Big| X \sim \mathrm{Normal}_{N \times N}(-X\mathbf{\Lambda}, \mathbf{M}(\mathbf{\Lambda}))$$

*where the elements of $\mathbf{M}(\mathbf{\Lambda})$ are such that*

$$\mathrm{Cov}\left(\frac{\partial^2 X}{\partial t_i \partial t_j}\,,\ \frac{\partial^2 X}{\partial t_k \partial t_l}\Big| X\right) = \varepsilon(i,j,k,l) - \lambda_{ij}\lambda_{kl}$$

*where $\varepsilon(i,j,k,l)$ is symmetric in its arguments.*

**Lemma 3.2** *We can write the first two derivatives of $U = U(\mathbf{t})$ in terms of independent random variables as follows, where the equalities are equalities in law:*

$$\text{(a)}\ \frac{\partial U}{\partial \mathbf{t}} = 2U^{\frac{1}{2}}\mathbf{z}, \quad \text{(b)}\ \frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} = 2(\mathbf{P} + \mathbf{z}\mathbf{z}' - U\mathbf{\Lambda} + U^{\frac{1}{2}}\mathbf{H}),$$

*where* $U \sim \chi^2_n$, $\mathbf{z} \sim \text{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$, $\mathbf{P} \sim \text{Wishart}_N(\mathbf{\Lambda}, n-1)$ *and* $\mathbf{H} \sim \text{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$, *all independently.*

*Proof.* Let $X_i = X_i(\mathbf{t})$ and $\mathbf{X}' = (X_1, \ldots, X_n)$ so that $U = \mathbf{X}'\mathbf{X}$. Let $\partial\mathbf{X}'/\partial\mathbf{t}$ denote the $N \times n$ matrix with $(k,i)$ element $X_i^{(k)}$, and $\partial\mathbf{X}/\partial\mathbf{t}'$ its transpose. Then conditioning on $\mathbf{X}$ and applying Lemma 3.1(a) to each $X_i$ we have

$$\frac{\partial U}{\partial \mathbf{t}} = 2\sum_{i=1}^n X_i \frac{\partial X_i}{\partial \mathbf{t}} = 2\frac{\partial \mathbf{X}'}{\partial \mathbf{t}}\mathbf{X} \sim \text{Normal}_N(\mathbf{0}, 4U\mathbf{\Lambda}).$$

Since this depends on $\mathbf{X}$ only through $U$, then it is the distribution of $\partial U/\partial\mathbf{t}$ conditional on $U$ alone. Letting $\mathbf{z} = U^{-\frac{1}{2}}(\partial\mathbf{X}'/\partial\mathbf{t})\mathbf{X}$ gives the first result (a).

For the second result (b) we have

$$\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} = 2\sum_{i=1}^n \frac{\partial X_i}{\partial \mathbf{t}}\frac{\partial X_i}{\partial \mathbf{t}'} + 2\sum_{i=1}^n X_i \frac{\partial^2 X_i}{\partial \mathbf{t}\partial \mathbf{t}'}.$$

Conditional on $\mathbf{X}$ and $\partial\mathbf{X}'/\partial\mathbf{t}$ we have, from Lemma 3.1(b) applied to each $X_i$,

$$\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'}\bigg|\mathbf{X}, \frac{\partial \mathbf{X}'}{\partial \mathbf{t}} \sim \text{Normal}_{N\times N}\left(2\frac{\partial \mathbf{X}'}{\partial \mathbf{t}}\frac{\partial \mathbf{X}}{\partial \mathbf{t}'} - 2U\mathbf{\Lambda}, 4U\mathbf{M}(\mathbf{\Lambda})\right),$$

so that we can write

$$\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} = 2\left(\frac{\partial \mathbf{X}'}{\partial \mathbf{t}}\frac{\partial \mathbf{X}}{\partial \mathbf{t}'} - U\mathbf{\Lambda} + U^{\frac{1}{2}}\mathbf{H}\right).$$

Let $\mathbf{I}_n$ be the $n \times n$ identity matrix, $\mathbf{A} = \mathbf{I}_n - \mathbf{X}\mathbf{X}'/U$ and

$$\mathbf{P} = \frac{\partial \mathbf{X}'}{\partial \mathbf{t}}\mathbf{A}\frac{\partial \mathbf{X}}{\partial \mathbf{t}'}.$$

Since $\mathbf{z} = U^{-\frac{1}{2}}(\partial\mathbf{X}'/\partial\mathbf{t})\mathbf{X}$ then we can write

$$\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} = 2(\mathbf{P} + \mathbf{z}\mathbf{z}' - U\mathbf{\Lambda} + U^{\frac{1}{2}}\mathbf{H}).$$

Assume for the moment that $\mathbf{X}$ is fixed. Then $\mathbf{A}$ is fixed and $\mathbf{P} \sim \text{Wishart}_N(\mathbf{\Lambda}, n-1)$ since $\mathbf{A}$ is idempotent of rank $n-1$. Since $\mathbf{A}\mathbf{X} = \mathbf{0}$ and $\mathbf{z}$ is a linear combination of $\partial\mathbf{X}'/\partial\mathbf{t}$ then $\mathbf{z}$ is independent of $\mathbf{P}$ conditional on $\mathbf{X}$. Since the distributions of $\mathbf{P}$ and $\mathbf{z}$ do not depend on $\mathbf{X}$, then $\mathbf{P}$ and $\mathbf{z}$ are independent unconditionally. □

## 3.3 Expectations

We shall assume that each of the component fields $X_i(\mathbf{t})$ satisfies the same regularity conditions inside a compact set $C \subset \mathbb{R}^N$ as are required for Theorem 2.1 to hold. Then Adler (1981), Lemma 7.1.1., page 171, shows that $U(\mathbf{t})$ is also suitably regular for Theorem 2.1 to hold, and Theorem 2.1 is used to derive the expected DT characteristic in two dimensions. In this section we shall extend this to higher dimensions.

First we shall give an asymptotic expression for the expectation of $M_u^+(U,C)$, the mean number of local maxima greater than $u$, and $M_u^-(U,C)$, the mean number of local minima less than $u$. We shall postpone the proofs of the results in this subsection until the end of section four where we shall derive them as special cases of the $F$ field (Corollaries 4.5 and 4.8).

**Theorem 3.3** *Under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields* $X_i(\mathbf{t})$,

$$\mathrm{E}(M_u^+(U,C)) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-N)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{1}{2}N} 2^{\frac{1}{2}(n-2)} \Gamma\left(\frac{n}{2}\right)} u^{N-1}\{1 + O(u^{-\frac{1}{2}})\},$$

*and provided* $n > N$ *then*

$$\mathrm{E}(M_u^-(U,C)) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-N)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{1}{2}N} 2^{\frac{1}{2}(n-2)} \Gamma\left(\frac{n}{2}\right)} \frac{(n-1)!}{(n-N)!}\{1 + O(u^{\frac{1}{2}})\}.$$

We can immediately find an asymptotic expression for the global maximum and global minimum of $U(\mathbf{t})$ using the same argument as Hasofer (1978):

**Corollary 3.4** *Let* $U_{\max} = \sup\{U(\mathbf{t}) : \mathbf{t} \in C\}$ *and* $U_{\min} = \inf\{U(\mathbf{t}) : \mathbf{t} \in C\}$. *Then*

$$\mathrm{P}(U_{\max} \geq u) \to \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-N)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{1}{2}N} 2^{\frac{1}{2}(n-2)} \Gamma\left(\frac{n}{2}\right)} u^{N-1} \quad \text{as } u \to \infty,$$

*and provided* $n > N$ *then*

$$\mathrm{P}(U_{\min} \leq u) \to \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-N)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{1}{2}N} 2^{\frac{1}{2}(n-2)} \Gamma\left(\frac{n}{2}\right)} \frac{(n-1)!}{(n-N)!} \quad \text{as } u \to 0.$$

**Theorem 3.5** *For* $N \geq 2$, *and under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields* $X_i(\mathbf{t})$,

$$\mathrm{E}(\chi(A_u(U,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-N)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{1}{2}N} 2^{\frac{1}{2}(n-2)} \Gamma\left(\frac{n}{2}\right)} P_{N,n}(u),$$

*where* $P_{N,n}(u)$ *is a polynomial of degree* $N-1$ *in* $u$ *with integer coefficients, given by*

$$P_{N,n}(u) = \sum_{j=0}^{[(N-1)/2]} \sum_{k=0}^{N-1-2j} \binom{n-1}{N-1-2j-k} \frac{(-1)^{N-1+j+k}(N-1)!}{2^j j! k!} u^{j+k},$$

*where division by the factorial of a negative integer is treated as multiplication by zero.*

**Corollary 3.6** *If* $N = 2$

$$\mathrm{E}(\chi(A_u(U,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-2)} e^{-\frac{1}{2}u}}{2\pi 2^{\frac{1}{2}(n-2)}\Gamma\left(\frac{n}{2}\right)}\{u-(n-1)\},$$

*and if* $N = 3$

$$\mathrm{E}(\chi(A_u(U,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} u^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}u}}{(2\pi)^{\frac{3}{2}} 2^{\frac{1}{2}(n-2)}\Gamma\left(\frac{n}{2}\right)}\{u^2-(2n-1)u+(n-1)(n-2)\}.$$

## 3.4 Comments

The two dimensional result ($N = 2$) for $C$ a unit square was obtained by Adler(1981), Theorem 7.1.2, page 172. Hamilton (1988) has derived the expected DT characteristic of excursion sets of a random field generated by the Rayleigh-Lévy random-walk fractal of Mandelbrot (1982). This turns out to be identical to that of a $\chi^2$ field in three dimensions with two degrees of freedom ($N = 3$, $n = 2$).

As Adler(1981), page 176, points out, Theorem 3.3 and 3.5 tell us a lot about the zeros of $\chi^2$ fields and Gaussian fields. For small, but non-zero, $u$ it is clear that the excursion set $A_u(U,C)$ consists essentially of the whole of $C$ except for a few 'holes' where the field drops below the level $u$, of which there are approximately $(-1)^{N-1}\chi(A_u(U,C))$ in number. Note that $\mathrm{E}(\chi(A_u(U,C)))$ is proportional to $e^{-\frac{1}{2}u}u^{\frac{1}{2}(n-N)}P_{N,n}(u)$. Taking the limit of $\mathrm{E}(\chi(A_u(U,C)))$ as $u \to 0$ tells us that if $n > N$ then there are with probability one, no zeros, confirming the result of Theorem 3.3. If $n = N$ there are an almost surely finite number of zeros, and on average

$$\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}\Gamma\{(N+1)/2\}\pi^{-\frac{1}{2}(N+1)},$$

since $P_{N,N}(0) = (N-1)! = 2^{N-1}\Gamma\{(N+1)/2\}\Gamma\{N/2\}/\pi^{\frac{1}{2}}$. Thus Theorem 3.3 can be extended to the case $n = N$. If $n = N-1$ then there are an infinite number of zeros which form 'strings' with an expected DT characteristic of zero, since $P_{N,N-1}(0) = 0$ and the coefficient of $u$ in $P_{N,N-1}(u)$ is non-zero. In particular, the zeros of a $\chi^2$ field in three dimensions with two degrees of freedom, or the Rayleigh-Lévy random-walk fractal, form closed 'loops' with an expected DT characteristic of zero. These observations will be important for defining $F$ and $t$ fields in the following sections.

# 4 The $F$ field

## 4.1 Definition

Let $X_1(\mathbf{t}),\ldots,X_n(\mathbf{t}),Y_1(\mathbf{t}),\ldots,Y_m(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^N$, be independent, identically distributed, homogeneous, real-valued Gaussian random fields with zero mean, unit variance, and let $\mathbf{\Lambda}$

$= \mathrm{Var}(\partial X_i(\mathbf{t})/\partial\mathbf{t}) = \mathrm{Var}(\partial Y_j(\mathbf{t})/\partial\mathbf{t})$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Then define the $F$ field as

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

The marginal distribution of $F(\mathbf{t})$ for fixed $\mathbf{t}$ is an $F$-distribution with $n$ and $m$ degrees of freedom. However there is a serious difficulty with the above definition which does not arise in the univariate case. If the degrees of freedom are small there may be many values of $\mathbf{t}$ inside a compact set $C$ where the numerator and denominator in the definition of $F(\mathbf{t})$ both take the value zero, with non-zero probability, and so $F(\mathbf{t})$ is not defined. This will happen when each of the component Gaussian fields takes the value zero, that is when $W(\mathbf{t}) = X_1(\mathbf{t})^2 + \ldots + X_n(\mathbf{t})^2 + Y_1(\mathbf{t})^2 + \ldots + Y_m(\mathbf{t})^2 = 0$. However $W(\mathbf{t})$ is a $\chi^2$ field with $m+n$ degrees of freedom and we have seen from the previous section that $W(\mathbf{t})$ has almost surely no zeros provided that $m+n > N$. Thus the definition of the $F$-field will be restricted to $m+n > N$.

## 4.2 Representations of derivatives

To make the algebra simpler we shall work with the field $G(\mathbf{t}) = (n/m)F(\mathbf{t}) = U(\mathbf{t})/V(\mathbf{t})$, where $U(\mathbf{t}) = \sum_{i=1}^{n} X_i(\mathbf{t})^2$ and $V(\mathbf{t}) = \sum_{i=1}^{m} Y_i(\mathbf{t})^2$ are independent $\chi^2$ fields with $n$ and $m$ degrees of freedom, respectively. We shall need the following Lemmata, similar to Lemma 3.2.

**Lemma 4.1** *We can express the first and second derivatives of $G = G(\mathbf{t})$ in terms of independent random variables as follows, where the equalities are equalities in law:*

*(a)* $\dfrac{\partial G}{\partial \mathbf{t}} = 2G^{\frac{1}{2}}(1+G)W^{-\frac{1}{2}}\mathbf{z}_1$

*(b)* $\dfrac{\partial^2 G}{\partial \mathbf{t}\partial \mathbf{t}'} = 2(1+G)[W^{-1}\{\mathbf{P} - G\mathbf{Q} + (1+3G)\mathbf{z}_1\mathbf{z}'_1 - G^{\frac{1}{2}}(\mathbf{z}_1\mathbf{z}'_2 + \mathbf{z}_2\mathbf{z}'_1)\} + G^{\frac{1}{2}}W^{-\frac{1}{2}}\mathbf{H}]$

*where* $(m/n)G \sim F_{n,m}$, $W \sim \chi^2_{m+n}$, $\mathbf{z}_1, \mathbf{z}_2 \sim \mathrm{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$, $\mathbf{P} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n-1)$, $\mathbf{Q} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m-1)$ *and* $\mathbf{H} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$*, all independently.*

*Proof.* Let $U = U(\mathbf{t})$, $V = V(\mathbf{t})$ and $W = W(\mathbf{t}) = U + V$. It is easily verified that $W \sim \chi^2_{m+n}$ and that $G$ and $W$ are independent. For the first derivative we have

$$\frac{\partial G}{\partial \mathbf{t}} = \frac{1}{V}\frac{\partial U}{\partial \mathbf{t}} - \frac{U}{V^2}\frac{\partial V}{\partial \mathbf{t}}.$$

From Lemma 3.2(a) we can make the substitutions

$$\frac{\partial U}{\partial \mathbf{t}} = 2U^{\frac{1}{2}}\mathbf{z}_U \text{ and } \frac{\partial V}{\partial \mathbf{t}} = 2V^{\frac{1}{2}}\mathbf{z}_V,$$

where $\mathbf{z}_U, \mathbf{z}_V \sim \mathrm{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$, to give

$$\frac{\partial G}{\partial \mathbf{t}} = 2U^{\frac{1}{2}}V^{-\frac{3}{2}}\left(V^{\frac{1}{2}}\mathbf{z}_U - U^{\frac{1}{2}}\mathbf{z}_V\right).$$

Letting

$$\mathbf{z}_1 = \frac{1}{2}U^{-\frac{1}{2}}V^{\frac{3}{2}}W^{-\frac{1}{2}}\frac{\partial G}{\partial \mathbf{t}} = W^{-\frac{1}{2}}\left(V^{\frac{1}{2}}\mathbf{z}_U - U^{\frac{1}{2}}\mathbf{z}_V\right),\ U = \frac{GW}{1+G},\ \text{and } V = \frac{W}{1+G}$$

gives the first result (a). It is easily verified that $\mathbf{z}_1 \sim \text{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$.

For the second derivative (b) we have

$$\frac{\partial^2 G}{\partial \mathbf{t}\partial \mathbf{t}'} = \frac{1}{V}\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} - \frac{U}{V^2}\frac{\partial^2 V}{\partial \mathbf{t}\partial \mathbf{t}'} - \frac{1}{V^2}\left(\frac{\partial U}{\partial \mathbf{t}}\frac{\partial V}{\partial \mathbf{t}'} + \frac{\partial V}{\partial \mathbf{t}}\frac{\partial U}{\partial \mathbf{t}'}\right) + \frac{2U}{V^3}\frac{\partial V}{\partial \mathbf{t}}\frac{\partial V}{\partial \mathbf{t}'}.$$

From Lemma 3.2(b) we can make the substitutions

$$\frac{\partial^2 U}{\partial \mathbf{t}\partial \mathbf{t}'} = 2(\mathbf{P} + \mathbf{z}_U\mathbf{z}_U' - U\mathbf{\Lambda} + U^{\frac{1}{2}}\mathbf{H}_U) \text{ and } \frac{\partial^2 V}{\partial \mathbf{t}\partial \mathbf{t}'} = 2(\mathbf{Q} + \mathbf{z}_V\mathbf{z}_V' - V\mathbf{\Lambda} + V^{\frac{1}{2}}\mathbf{H}_V),$$

where $\mathbf{P} \sim \text{Wishart}_N(\mathbf{\Lambda}, n-1)$, $\mathbf{Q} \sim \text{Wishart}_N(\mathbf{\Lambda}, m-1)$ and $\mathbf{H}_U, \mathbf{H}_V \sim \text{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$. This gives

$$\frac{\partial^2 G}{\partial \mathbf{t}\partial \mathbf{t}'} = 2V^{-1}\left(\mathbf{P} - G\mathbf{Q} + \mathbf{z}_U\mathbf{z}_U' - 2G^{\frac{1}{2}}(\mathbf{z}_U\mathbf{z}_V' + \mathbf{z}_V\mathbf{z}_U') + 3G\mathbf{z}_V\mathbf{z}_V' + G^{\frac{1}{2}}(V^{\frac{1}{2}}\mathbf{H}_U - U^{\frac{1}{2}}\mathbf{H}_V)\right).$$

We now find a linear combination of $\mathbf{z}_U$ and $\mathbf{z}_V$ that is independent of $\mathbf{z}_1$:

$$\mathbf{z}_2 = \frac{1}{2}W^{-\frac{1}{2}}\frac{\partial W}{\partial \mathbf{t}} = W^{-\frac{1}{2}}\left(U^{\frac{1}{2}}\mathbf{z}_U + V^{\frac{1}{2}}\mathbf{z}_V\right).$$

Then we can write

$$\mathbf{z}_U = W^{-\frac{1}{2}}\left(V^{\frac{1}{2}}\mathbf{z}_1 + U^{\frac{1}{2}}\mathbf{z}_2\right) \text{ and } \mathbf{z}_V = W^{-\frac{1}{2}}\left(-U^{\frac{1}{2}}\mathbf{z}_1 + V^{\frac{1}{2}}\mathbf{z}_2\right).$$

If we let

$$\mathbf{H} = W^{-\frac{1}{2}}(V^{\frac{1}{2}}\mathbf{H}_U - U^{\frac{1}{2}}\mathbf{H}_V),$$

and substitute these into the above then we obtain the result (b). It is easily verified that $\mathbf{H} \sim \text{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$, $\mathbf{z}_2 \sim \text{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$ and that $\mathbf{z}_1$ and $\mathbf{z}_2$ are independent. □

## 4.3 Expectations

**Lemma 4.2** *Provided that $m + n > N$ and the component fields $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ satisfy the conditions of Theorem 2.1 then $F$ is suitably regular for Theorem 2.1 to hold.*

*Proof.* For the regularity conditions (i) of Theorem 2.1 we follow the arguments of Lemma 7.1.1 of Adler(1981), page 171. Let $M$ be the supremum over all $\mathbf{t} \in C$ of the absolute value of $X_i(\mathbf{t})$ and all its derivatives up to second order, $i = 1, \ldots, n$, $Y_i(\mathbf{t})$ and all its derivatives up to second order, $i = 1, \ldots, m$, and $V(\mathbf{t})^{-1}$. The moduli of continuity of $G(\mathbf{t})$ will converge to zero at the right probabilistic rate if we condition on $M < K$ throughout the proof of Lemma 5.2.2 of Adler(1981), page 100, eventually letting $K \to \infty$.

For conditions (iii) of Theorem 2.1 we shall use the representation of Lemma 4.1 for the first and second derivatives of $G$ in terms of independent random variables. The conditions (iii) are not directly satisfied by $G$, but we can nevertheless prove that Theorem 2.1 holds by conditioning on $W$ while working through the proof of Lemma 5.2.1 of Adler(1981), page 98. Now conditional on $G, W, \partial G/\partial t_1, \ldots, \partial G/\partial t_{N-1}$, we see from Lemma 4.1 that $\partial G/\partial t_N$ is $O(W^{-\frac{1}{2}})$, the determinant of any $(N-1)\times(N-1)$ minor of $\partial^2 G/\partial\mathbf{t}\partial\mathbf{t}'$ is $W^{-\frac{1}{2}(N-1)}$ times a polynomial in $W^{-\frac{1}{2}}$ of degree $N-1$, and the joint density of $G, \partial G/\partial t_1, \ldots, \partial G/\partial t_{N-1}$ conditional on $W$ is $O(W^{\frac{1}{2}(N-1)})$. The bound required by the proof is the product of these, which is thus a polynomial of degree $N$ in $W^{-\frac{1}{2}}$ with coefficients that have finite variance. Taking expectations over $W \sim \chi^2_{m+n}$, we see that this bound has finite variance provided that $m+n-N \geq 1$, or $m+n > N$. Similar arguments apply to the conditions (ii). □

**Theorem 4.3** *For $m > N$, and under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields $X_i(\mathbf{t})$ and $Y_i(\mathbf{t})$, then*

$$\mathrm{E}(M_f^+(F,C)) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}N}2^{\frac{1}{2}(N-2)}}\frac{\Gamma\left(\frac{m+n-N}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\frac{(m-1)!}{(m-N)!}\left(\frac{nf}{m}\right)^{-\frac{1}{2}(m-N)}\{1+O(f^{-\frac{1}{2}})\}.$$

*Proof.* Since $m > N$ implies $m+n > N$ then by Lemma 4.2 we can apply Theorem 2.1 to $F(\mathbf{t})$. Let $G = G(\mathbf{t}) = (n/m)F(\mathbf{t}) = U/V$ and $\mathbf{G} = \partial G/\partial\mathbf{t}$. We shall evaluate the expectations in Theorem 2.1 by conditioning on both $G$ and $W$ and then taking expectations over $W$. Now since $W \sim \chi^2_{m+n}$, independent of $G$, using Theorem 2.1 we can write

$$\mathrm{E}(M_g^+(G,C)) = \lambda(C)\int_g^\infty \mathrm{E}_W\{\mathrm{E}\{-\det(\mathbf{D}_N^-)|G=h, W, \mathbf{G}=\mathbf{0}\}\ \psi_N(\mathbf{0};h,W)\}\psi_0(h)\ dh$$

where $\mathbf{D}_N$ is the $N\times N$ matrix of all second order partial derivatives of $G$, $\psi_N(\mathbf{g};h,W)$ is the density of $\mathbf{G}$ conditional on $G=h$ and $W$, and $\psi_0(h)$ is the density of $G$. By Lemma 4.1 we can see that if $\mathbf{G}=\mathbf{0}$ then we can write

$$\mathbf{D}_N = 2(1+G)[W^{-1}(\mathbf{P}-G\mathbf{Q}) + G^{\frac{1}{2}}W^{-\frac{1}{2}}\mathbf{H}],$$

where $\mathbf{P} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, n-1)$, $\mathbf{Q} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m-1)$ and $\mathbf{H} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$, independently of $G$ and $W$. Thus $\det(\mathbf{D}_N^-)$ is a polynomial in $G^{\frac{1}{2}}$ of degree $2N$, multiplied by $(1+G)^N$. Following the same arguments as in the proof of Theorem 6.3.1 of Adler(1981), page 134, it can be shown that as $G\to\infty$ then $\mathbf{D}_N$ approaches the negative definite matrix $-2(1+G)GW^{-1}\mathbf{Q}$ to give

$$\mathrm{E}\{-\det(\mathbf{D}_N^-)|G=h, W, \mathbf{G}=\mathbf{0}\} = \mathrm{E}(\det(\mathbf{Q}))(2(1+h)hW^{-1})^N\{1+O(h^{-\frac{1}{2}})\}.$$

From standard multivariate statistics we have $\mathrm{E}\{\det(\mathbf{Q})\} = \det(\mathbf{\Lambda})(m-1)!/(m-1-N)!$ (see for example Anderson(1984), page 265), which is non-zero since $m > N$. From Lemma 4.1(a) the density of $\mathbf{G}$ at zero conditional on $G=h$ and $W$ is

$$\psi_N(\mathbf{0};h,W) = (2\pi)^{-\frac{1}{2}N}\det(\mathbf{\Lambda})^{-\frac{1}{2}}2^{-N}(1+h)^{-N}h^{-\frac{1}{2}N}W^{\frac{1}{2}N}.$$

Multiplying these together gives

$$\begin{aligned} & \mathrm{E}\{-\det(\mathbf{D}_N^-)|G=h,W,\mathbf{G}=\mathbf{0}\}\ \psi_N(\mathbf{0};h,W) \\ &= \frac{\det(\mathbf{\Lambda})^{\frac{1}{2}}(m-1)!}{(2\pi)^{\frac{1}{2}N}(m-1-N)!}h^{\frac{1}{2}N}\{1+O(h^{-\frac{1}{2}})\}W^{-\frac{1}{2}N}. \end{aligned}$$

Using the fact that $\mathrm{E}(W^d)=2^d\Gamma\{(m+n)/2+d\}/\Gamma\{(m+n)/2\}$ and multiplying by the density of $G$:

$$\psi_0(h)=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}h^{\frac{1}{2}n-1}(1+h)^{-\frac{1}{2}(m+n)}$$

gives:

$$\begin{aligned} & \mathrm{E}_W\{\mathrm{E}\{-\det(\mathbf{D}_N^-)|G=h,W,\mathbf{G}=\mathbf{0}\}\ \psi_N(\mathbf{0};h,W)\}\psi_0(h) \\ &= \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}N}2^{\frac{1}{2}N}}\frac{\Gamma\left(\frac{m+n-N}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\frac{(m-1)!}{(m-1-N)!}h^{-\frac{1}{2}(m-N)-1}\{1+O(h^{-\frac{1}{2}})\}. \end{aligned}$$

Integrating over $h$ and converting back from $G$ to $F$ gives the result. □

**Corollary 4.4** *Let $F_{\max}=\sup\{F(\mathbf{t}):\mathbf{t}\in C\}$. For $m>N$ and as $f\to\infty$*

$$\mathrm{P}(F_{\max}\geq f)\to\frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}N}2^{\frac{1}{2}(N-2)}}\frac{\Gamma\left(\frac{m+n-N}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\frac{(m-1)!}{(m-N)!}\left(\frac{nf}{m}\right)^{-\frac{1}{2}(m-N)}.$$

*Proof.* Following Hasofer (1978) we have

$$\mathrm{P}(F_{\max}\geq f)=\mathrm{P}(M_f^+(F,C)\geq 1)\leq \mathrm{E}(M_f^+(F,C))$$

with convergence as $f\to\infty$. □

**Corollary 4.5** *Theorem 3.3 holds.*

*Proof.* Since $U(\mathbf{t})=nF(\mathbf{t})$ in the limit as $m\to\infty$ then the first result of Theorem 3.3 for $\mathrm{E}(M_u^+(U,C))$ can be obtained by letting $m\to\infty$ in the result of Theorem 4.3. The second result of Theorem 3.3 can be obtained by noting that $\mathrm{E}(M_v^-(V,C))=\mathrm{E}(M_{1/v}^+(1/V,C))$ and $1/V(\mathbf{t})=F(\mathbf{t})/m$ as $n\to\infty$. Provided $m>N$, Theorem 4.3 gives the desired result, after a change of parameters from $m$ to $n$. □

**Theorem 4.6** *For $N\geq 2$, $m+n>N$, and under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields $X_i(\mathbf{t})$ and $Y_i(\mathbf{t})$, then*

$$\mathrm{E}(\chi(A_f(F,C)))=\frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}N}2^{\frac{1}{2}(N-2)}}\frac{\Gamma\left(\frac{m+n-N}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{nf}{m}\right)^{\frac{1}{2}(n-N)}\left(1+\frac{nf}{m}\right)^{-\frac{1}{2}(m+n-2)}K_{N,m,n}(f),$$

*where $K_{N,m,n}(f)$ is a polynomial of degree $N-1$ in $nf/m$ with integer coefficients given by*

$$K_{N,m,n}(f) = (-1)^{N-1}(N-1)! \sum_{j=0}^{[(N-1)/2]} \frac{\Gamma\left(\frac{m+n-N}{2}+j\right)}{\Gamma\left(\frac{m+n-N}{2}\right) j!}$$

$$\sum_{k=0}^{N-1-2j} \binom{m-1}{k} \binom{n-1}{N-1-2j-k} (-1)^{j+k} \left(\frac{nf}{m}\right)^{j+k}.$$

*Proof.* Let $G = G(\mathbf{t}) = (n/m)F(\mathbf{t}) = U/V$ and $G^{(k)} = \partial G/\partial t_k$. We shall evaluate the expectations in Theorem 2.1 by conditioning on both $G$ and $W$ and then taking expectations over $W$. Now since $W \sim \chi^2_{m+n}$, independent of $G$, we can write

$$\mathrm{E}(\chi(A_g(G,C))) = (-1)^{N-1}\lambda(C)\mathrm{E}_W\{\mathrm{E}(G^{(N)+}\det(\mathbf{D}_{N-1})|G=g, W, G^{(1)}=0,\ldots,G^{(N-1)}=0)$$

$$\psi_{N-1}(0,\ldots,0;g,W)\}\psi_0(g),$$

where $\mathbf{D}_{N-1}$ is the $(N-1)\times(N-1)$ matrix of second order partial derivatives of $G$ with respect to $t_1,\ldots,t_{N-1}$, $\psi_{N-1}(g_1,\ldots,g_{N-1};G,W)$ is the density of $G^{(1)},\ldots,G^{(N-1)}$ conditional on $G$ and $W$, and $\psi_0(g)$ is the density of $G$. By Lemma 4.1 we can see that if $G^{(1)}=0,\ldots,G^{(N-1)}=0$ then we can write

$$\mathbf{D}_{N-1} = c(\mathbf{P}^* + a\mathbf{Q}^* + b\mathbf{H}^*),$$

where $c = 2(1+G)W^{-1}$, $a=-G$, $b = G^{\frac{1}{2}}W^{\frac{1}{2}}$, $\mathbf{\Lambda}^*$ is the $(N-1)\times(N-1)$ matrix of the first $N-1$ rows and columns of $\mathbf{\Lambda}$, and $\mathbf{P}^* \sim \mathrm{Wishart}_{N-1}(\mathbf{\Lambda}^*, n-1)$, $\mathbf{Q}^* \sim \mathrm{Wishart}_{N-1}(\mathbf{\Lambda}^*, m-1)$ and $\mathbf{H}^* \sim \mathrm{Normal}_{(N-1)\times(N-1)}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}^*))$, independently. Thus $G^{(N)}$ is independent of $\mathbf{D}_{N-1}$ and

$$\mathrm{E}(G^{(N)+}\det(\mathbf{D}_{N-1})|G,W,G^{(1)}=0,\ldots,G^{(N-1)}=0)$$
$$= \mathrm{E}(G^{(N)+}|G,W,G^{(1)}=0,\ldots,G^{(N-1)}=0)\mathrm{E}(\det(\mathbf{D}_{N-1})|G,W,G^{(1)}=0,\ldots,G^{(N-1)}=0).$$

We shall start with the first term. Let $\mathbf{z}_1 = (z_1,\ldots,z_N)' \sim \mathrm{Normal}_N(\mathbf{0},\mathbf{\Lambda})$ independent of $\mathbf{P}^*$, $\mathbf{Q}^*$ and $\mathbf{H}^*$, and let $\lambda_N = \mathrm{Var}(z_N|z_1,\ldots,z_{N-1})$. Then from Lemma 4.1(a) and Adler (1981), Lemma 5.3.3, page 111, we have

$$\mathrm{E}(G^{(N)+}|G,W,G^{(1)}=0,\ldots,G^{(N-1)}=0) = bc\mathrm{E}(z_N^+|z_1=0,\ldots,z_{N-1}=0) = bc(2\pi)^{-\frac{1}{2}}\lambda_N^{\frac{1}{2}}.$$

For the second term, let $\mathbf{B}$ be an orthogonal $(N-1)\times(N-1)$ matrix such that $\mathbf{B}'\mathbf{\Lambda}^*\mathbf{B} = \mathbf{I}_{N-1}$. Then

$$\mathrm{E}(\det(\mathbf{D}_{N-1})|G,W,G^{(1)}=0,\ldots,G^{(N-1)}=0) = c^{N-1}\mathrm{E}(\det(\mathbf{P}^*+a\mathbf{P}^*+b\mathbf{H}^*))$$

$$= \det(\mathbf{\Lambda}^*)c^{N-1}\mathrm{E}(\det(\mathbf{B}'(\mathbf{P}^*+a\mathbf{Q}^*+b\mathbf{H}^*)\mathbf{B})$$

$$= \det(\mathbf{\Lambda}^*)c^{N-1}\mathrm{E}(\det(\mathbf{P}+a\mathbf{Q}+b\mathbf{H}))$$

where $\mathbf{P} = \mathbf{B}'\mathbf{P}^*\mathbf{B} \sim \text{Wishart}_N(\mathbf{I}_{N-1}, n-1)$, $\mathbf{Q} = \mathbf{B}'\mathbf{Q}^*\mathbf{B} \sim \text{Wishart}_N(\mathbf{I}_{N-1}, m-1)$ and $\mathbf{H} = \mathbf{B}'\mathbf{H}^*\mathbf{B} \sim \text{Normal}_{(N-1)\times(N-1)}(\mathbf{0}, \mathbf{M}(\mathbf{I}_{N-1}))$ independently, by Lemma 7.1(a). We can now apply Lemma 7.5 to obtain

$$\mathrm{E}(\det(\mathbf{D}_{N-1})|G, W, G^{(1)} = 0, \ldots, G^{(N-1)} = 0) = \det(\mathbf{\Lambda}^*)2^{N-1}(1+G)^{N-1}$$
$$\sum_{j=0}^{[(N-1)/2]} \sum_{k=0}^{N-1-2j} \binom{m-1}{k} \binom{n-1}{N-1-2j-k} \frac{(-1)^{j+k}(N-1)!}{2^j j!} G^{j+k} W^{-(N-1-j)}.$$

From Lemma 4.1(a) the density of $G^(1), \ldots, G^{(N-1)}$ at zero conditional on $G = g$ and $W$ is

$$\psi_{N-1}(0, \ldots, 0; g, W) = (2\pi 4g(1+g)^2 W^{-1})^{-\frac{1}{2}(N-1)} \det(\mathbf{\Lambda}^*)^{-\frac{1}{2}}.$$

Multiplying these together gives

$$\mathrm{E}(G^{(N)+}\det(\mathbf{D}_{N-1})|G = g, W, G^{(1)} = 0, \ldots, G^{(N-1)} = 0)\psi_{N-1}(0, \ldots, 0; g, W)$$
$$= (2\pi)^{-\frac{1}{2}N}\det(\mathbf{\Lambda})^{\frac{1}{2}} 2(1+g) g^{-\frac{1}{2}N+1}$$
$$\sum_{j=0}^{[(N-1)/2]} \sum_{k=0}^{N-1-2j} \binom{m-1}{k} \binom{n-1}{N-1-2j-k} \frac{(-1)^{j+k}(N-1)!}{2^j j! k!} g^{j+k} W^{-\frac{1}{2}(N-2j)},$$

since $\lambda_N \det(\mathbf{\Lambda}^*) = \det(\mathbf{\Lambda})$. Using the fact that $\mathrm{E}(W^d) = 2^d\Gamma\{(m+n)/2+d\}/\Gamma\{(m+n)/2\}$ and multiplying by the density of $G$:

$$\psi_0(g) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} g^{\frac{1}{2}n-1}(1+g)^{-\frac{1}{2}(m+n)}$$

gives:

$$\mathrm{E}(\chi(A_g(G, C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}} g^{\frac{1}{2}(n-N)}(1+g)^{-\frac{1}{2}(m+n-2)}}{(2\pi)^{\frac{1}{2}N}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}$$
$$\sum_{j=0}^{[(N-1)/2]} \sum_{k=0}^{N-1-2j} \binom{m-1}{k} \binom{n-1}{N-1-2j-k} \frac{(-1)^{N-1+j+k}(N-1)!\Gamma\left(\frac{m+n-N+2j}{2}\right)}{2^{\frac{1}{2}(N-2)} j!} g^{j+k}.$$

Converting back from $G$ to $F$ gives the result. □

**Corollary 4.7** *If $N = 2$ and $m + n \geq 3$ then*

$$\mathrm{E}(\chi(A_f(F, C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{2\pi} \frac{\Gamma\left(\frac{m+n-2}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{nf}{m}\right)^{\frac{1}{2}(n-2)} \left(1+\frac{nf}{m}\right)^{-\frac{1}{2}(m+n-2)}$$
$$\left\{(m-1)\frac{nf}{m} - (n-1)\right\},$$

*and if $N = 3$ and $m + n \geq 4$ then*

$$\mathrm{E}(\chi(A_f(F, C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}} 2^{\frac{1}{2}}} \frac{\Gamma\left(\frac{m+n-3}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{nf}{m}\right)^{\frac{1}{2}(n-3)} \left(1+\frac{nf}{m}\right)^{-\frac{1}{2}(m+n-2)}$$
$$\left\{(m-1)(m-2)\left(\frac{nf}{m}\right)^2 - (2mn-m-n-1)\frac{nf}{m} + (n-1)(n-2)\right\}.$$

**Corollary 4.8** *Theorem 3.5 holds.*

*Proof.* Since $U(\mathbf{t}) = nF(\mathbf{t})$ in the limit as $m \to \infty$ then the result of Theorem 3.5 can be obtained by letting $m \to \infty$ in the result of Theorem 4.6. $\square$

## 4.4 Comments

It is worth noting three things about these results. First, they only depend on the distribution of the component fields through the variance of their first order derivatives, even though the definition of $M^+$ and $\chi(A)$ depend on second order derivatives. Second, $\mathrm{E}(\chi(A))$ is invariant under rotations of the coordinate system, even though $\chi(A)$ is not invariant under rotations if $A$ touches the boundary of the region $C$. Third, $\mathrm{E}(M_f^+(F,C))$ and $\mathrm{E}(\chi(A_f(F,C)))$ converge to the same limit as $f \to \infty$, but the former at the rate $O(f^{-\frac{1}{2}})$ and the latter at the faster rate $O(f^{-1})$. These comments also apply to Gaussian, $\chi^2$, and, as we shall see in the next section, $t$ fields.

We end this section with some comments on infinite values of $F(\mathbf{t})$. For large $f$ it is clear that the excursion set $A_f(F,C)$ consists essentially of the whole of $C$ except for a few 'peaks' where the field $F(\mathbf{t})$ exceeds the level $f$, of which there are approximately $\chi(A_f(F,C))$ in number. Note that $\mathrm{E}(\chi(A_f(F,C)))$ is proportional to $f^{-\frac{1}{2}(m-N)}$. Taking the limit of $\mathrm{E}(\chi(A_f(F,C)))$ as $f \to \infty$ tells us that if $m > N$ then there are with probability one, no points where the field is infinite. If $m = N$ there are an almost surely finite number of infinities, and on average

$$\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}\Gamma\{(N+1)/2\}\pi^{-\frac{1}{2}(N+1)}.$$

Thus Theorem 4.3 can be extended to the case $m = N$. It is not surprising that this result is the same as the expected number of zeros of a $\chi^2$ field with $N$ degrees of freedom, from Theorem 3.3; in fact we can see that the behaviour of an $F$ field near infinity depends largely on the behaviour of its denominator $\chi^2$ field near zero. Similar results apply to the zeros of $F$ fields, as can be seen by taking the reciprocal of the $F$ field.

# 5 The $t$ field

## 5.1 Definition

Let $X(\mathbf{t}), Y_1(\mathbf{t}), \ldots, Y_m(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^N$, be independent, identically distributed, homogeneous, real-valued Gaussian random fields with zero mean, unit variance, and $\mathbf{\Lambda} = \mathrm{Var}(\partial X(\mathbf{t})/\partial \mathbf{t})$ $= \mathrm{Var}(\partial Y_i(\mathbf{t})/\partial \mathbf{t})$, $i = 1, \ldots, m$. Define the $t$ field as

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

The marginal distribution of $T(\mathbf{t})$ for fixed $\mathbf{t}$ is a $t$ distribution with $m$ degrees of freedom, written $\mathrm{t}_m$. Note that $T(\mathbf{t})^2$ is an $F$ field with 1 and $m$ degrees of freedom. Once again we must avoid the possibility that the numerator and denominator both take the value zero inside a compact set $C$. This will happen when each of the component Gaussian fields takes the value zero, or when $S(\mathbf{t}) = X(\mathbf{t})^2 + Y_1(\mathbf{t})^2 + \ldots + Y_m(\mathbf{t})^2 = 0$. However $S(\mathbf{t})$ is a $\chi^2$ field with $m+1$ degrees of freedom and we have seen from the section three that $S(\mathbf{t})$ has almost surely no zeros provided that $m+1 > N$. Thus the definition of the $t$-field will be restricted to $m \geq N$.

## 5.2 Representations of derivatives

We can find a representation for the derivatives of a $t$ field, similar to that of Lemma 3.2. The proof is similar and is omitted.

**Lemma 5.1** *We can express the first and second derivatives of $T = T(\mathbf{t})$ in terms of independent random variables as follows, where equalities are equlities in law:*

*(a)* $\dfrac{\partial T}{\partial \mathbf{t}} = m^{\frac{1}{2}}(1 + T^2/m) S^{-\frac{1}{2}} \mathbf{z}_1$

*(b)* $\dfrac{\partial^2 T}{\partial \mathbf{t} \partial \mathbf{t}'} = m(1 + T^2/m) S^{-1}\{-m^{-\frac{1}{2}} T(\mathbf{Q} - 2\mathbf{z}_1 \mathbf{z}'_1) - \mathbf{z}_1 \mathbf{z}'_2 - \mathbf{z}_2 \mathbf{z}'_1 + S^{\frac{1}{2}} \mathbf{H}\}$

*where* $T \sim \mathrm{t}_m$, $S \sim \chi^2_{m+1}$, $\mathbf{z}_1, \mathbf{z}_2 \sim \mathrm{Normal}_N(\mathbf{0}, \mathbf{\Lambda})$, $\mathbf{Q} \sim \mathrm{Wishart}_N(\mathbf{\Lambda}, m-1)$ *and* $\mathbf{H} \sim \mathrm{Normal}_{N \times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Lambda}))$, *all independently.*

## 5.3 Expectations

**Theorem 5.2** *For $N \geq 2$ and $m \geq N$, and under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields $X(\mathbf{t})$ and $Y_i(\mathbf{t})$,*

$$\mathrm{E}(M_t^+(T, C)) = \frac{\lambda(C) \det(\mathbf{\Lambda})^{\frac{1}{2}} \Gamma\left(\frac{m+1}{2}\right) m^{\frac{1}{2}(m-N)}}{2(\pi)^{\frac{1}{2}(N+1)} \Gamma\left(\frac{m+2-N}{2}\right)} t^{-(m-N)} \{1 + O(t^{-1})\}.$$

*Proof.* This result follows from Theorem 4.3 and the fact that $T(\mathbf{t})^2$ is an $F$ field with 1 and $m$ degrees of freedom. Since $T(\mathbf{t})$ has the same distribution as $-T(\mathbf{t})$ then for $t > 0$ $\mathrm{E}(M_t^+(T, C)) = \mathrm{E}(M_{-t}^+(-T, C)) = \mathrm{E}(M_{t^2}^+(T^2, C))/2$. For $m > N$ the result is obtained using Theorem 4.3 with $n = 1$ and $f = t^2$, and the relation $(m-N)! = 2^{m-N} \Gamma\{(m+2-N)/2\} \Gamma\{(m+1-N)/2\} / \Gamma(1/2)$. We can extend the result to the case $m = N$ following the comments at the end of Theorem 4.6. $\square$

**Corollary 5.3** *Let $T_{\max} = \sup\{T(\mathbf{t}) : \mathbf{t} \in C\}$. For $m > N$ and as $t \to \infty$*

$$\mathrm{P}(T_{\max} \geq t) \to \frac{\lambda(C) \det(\mathbf{\Lambda})^{\frac{1}{2}} \Gamma\left(\frac{m+1}{2}\right) m^{\frac{1}{2}(m-N)}}{2(\pi)^{\frac{1}{2}(N+1)} \Gamma\left(\frac{m+2-N}{2}\right)} t^{-(m-N)}.$$

**Theorem 5.4** *For $N \geq 2$ and $m \geq N$ , and under the same regularity conditions as are required for Theorem 2.1 to hold for each of the component fields $X(\mathbf{t})$ and $Y_i(\mathbf{t})$,*

$$\mathrm{E}(\chi(A_t(T,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}(N+1)}}\left(1+\frac{t^2}{m}\right)^{-\frac{1}{2}(m-1)} Q_{N,m}(t),$$

*where $Q_{N,m}(t)$ is a polynomial of degree $N-1$ in $t$ given by*

$$Q_{N,m}(t) = \sum_{j=0}^{[(N-1)/2]} \frac{(-1)^j (N-1)!}{2^j j!(N-1-2j)!} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m+2-N+2j}{2}\right)\left(\frac{m}{2}\right)^{\frac{1}{2}(N-1-2j)}} t^{N-1-2j}.$$

*Proof.* Suppose first that $t > 0$, so that the two excursion sets $A_t(T,C)$ and $A_{-t}(-T,C)$ are disjoint and their union is $A_{t^2}(T^2,C)$. Since the DT characteristic of the union of two disjoint sets is the sum of the DT characteristics of the two sets, and $T(\mathbf{t})$ has the same distribution as $-T(\mathbf{t})$, then $\mathrm{E}(\chi(A_t(T,C))) = \mathrm{E}(\chi(A_{t^2}(T^2,C)))/2$. Since $T^2$ is an $F$ field with $n = 1$ then the result is obtained using Theorem 4.6, and the relation $(m-N+2j)! = 2^{m-N+2j}\Gamma\{(m+2-N)/2+j\}\Gamma\{(m+1-N)/2+j\}/\Gamma(1/2)$. Suppose now that $t < 0$. It is straightforward to show by a proof similar to that of Lemma 4.2 that the conditions of Theorem 2.1 are satisfied by $T(\mathbf{t})$ so by Corollary 2.2 we have

$$\mathrm{E}(\chi(A_t(T,C))) = (-1)^{N-1}\mathrm{E}(\chi(A_{-t}(-T,C))) = (-1)^{N-1}\mathrm{E}(\chi(A_{-t}(T,C)))$$

since $T(\mathbf{t})$ has the same distribution as $-T(\mathbf{t})$. Since $Q_{N,m}(-t) = (-1)^{N-1}Q_{N,m}(t)$ then the result of the Theorem can be extended to $t \neq 0$. For $t = 0$, $T = 0$ implies $X = 0$ almost surely. It can be checked that the expected DT characteristic of a Gaussian field at zero, from Adler (1981), Theorem 5.3.1, page 111, agrees with the result of the Theorem for $t = 0$. □

**Corollary 5.5** *If $N = 2$ and $m \geq 2$ then*

$$\mathrm{E}(\chi(A_t(T,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\left(1+\frac{t^2}{m}\right)^{-\frac{1}{2}(m-1)} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{2}\right)^{\frac{1}{2}}} t,$$

*and if $N = 3$ and $m \geq 3$ then*

$$\mathrm{E}(\chi(A_t(T,C))) = \frac{\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}}{(2\pi)^2}\left(1+\frac{t^2}{m}\right)^{-\frac{1}{2}(m-1)} \left\{\frac{m-1}{m}t^2 - 1\right\}.$$

## 5.4 Comments

Note that the coefficients of $Q_{N,m}(t)$ are equal to those of the Hermite polynomial of degree $N-1$, $\mathrm{He}_{N-1}(t)$, mutiplied by a term that is less than one but which converges to one as $m \to \infty$. Thus as the degrees of freedom approaches infinity, the expected DT characteristic of the $t$ field converges to that of the Gaussian field, as given by Adler (1981), Theorem

5.3.1, page 111. At extreme thresholds, the behaviour of the $t$ field is largely dominated by the behaviour near zero of the denominator $\chi^2$ field $V(\mathbf{t})$. If $m = N$ then we saw in section three that $V(\mathbf{t})$ takes the value zero at a finite number of points, so that the $t$ field can be infinite at a finite number of points. Since the numerator $X(\mathbf{t})$ takes positive values with a probability of one half, then we expect that the number of positive infinite values of $T(\mathbf{t})$ to be one half the expected number of zeros of $V(\mathbf{t})$. It is reassuring to note that letting $t \to \infty$ in the result of Theorem 5.4 confirms this.

# 6 Applications to positron emission tomography images

The PET technique uses positron-emitting isotope labelled carriers, created in an on-site cyclotron, to produce an image of brain activity such as glucose utilization, oxygen consumption or blood flow. In the last three years, researchers have succeeded in carrying out a new type of experiment in which a group of subjects in a randomised block design are given a set of stimuli such as word-recognition or a painfull heat stimulus. The number of subjects and stimuli is small, typically 10 and 6 respectively. The unusual statistical feature is that each observation is a three-dimensional image of blood flow in the brain. By careful alignment it is possible to subtract the blood flow image under one stimulus condition from that under another to look for changes in blood flow, or activation between the two stimuli. Such an experiment is fully described in Worsley et al. (1992) and the relevant details of the analysis will now be given.

Suppose that there are $p$ subjects and that for the $i$th subject, $\Delta_i(\mathbf{t})$ is the value of this differenced image at a point $\mathbf{t} \in C$, where $C \subset \mathrm{I\!R}^3$ is the brain, $i = 1, \ldots, p$. Let $\mu(\mathbf{t}) = \mathrm{E}(\Delta_i(\mathbf{t}))$ be the mean change in blood flow, or mean activation, between the two stimuli, and let $\sigma(\mathbf{t})^2 = \mathrm{Var}(\Delta_i(\mathbf{t}))$, $i = 1, \ldots, p$. Then it was assumed that for each subject, $\{\Delta_i(\mathbf{t}) - \mu(\mathbf{t})\}/\sigma(\mathbf{t})$ was an independent homogeneous Gaussian field with zero mean and unit variance, $i = 1, \ldots, p$. We are interested in testing the null hypothesis of no activation, $\mu(\mathbf{t}) = 0$. The sample mean $\bar{\Delta}(\mathbf{t})$ and sample variance $S(\mathbf{t})$ of the differenced images over all subjects were then calculated:

$$\bar{\Delta}(\mathbf{t}) = \sum_{i=1}^{p} \Delta_i(\mathbf{t})/p, \text{ and } S^2(\mathbf{t}) = \sum_{i=1}^{p} \left(\Delta_i(\mathbf{t}) - \bar{\Delta}(\mathbf{t})\right)^2/(p-1).$$

Worsley et al. (1992) made the assumption that the standard deviation was stationary, that is $\sigma(\mathbf{t}) = \sigma$, say. The variance $\sigma^2$ was then estimated by pooling the sample variance $S^2(\mathbf{t})$ over all $\mathbf{t} \in C$ to give

$$\hat{\sigma}^2 = \int_{\mathbf{t} \in C} S^2(\mathbf{t}) d\mathbf{t}/\lambda(C).$$

The sample mean image was then standardised to give

$$X(\mathbf{t}) = p^{\frac{1}{2}}\bar{\Delta}(\mathbf{t})/\hat{\sigma}.$$

Since the brain volume was large relative to the resolution of the image, then $\hat{\sigma}$ is approximately constant, so that under the null hypothesis $\mu(\mathbf{t}) = 0$, $X(\mathbf{t})$ can be well approximated by a homogeneous Gaussian field with zero mean and unit variance. Since activation was expected to produce a few isolated regions of high mean, then the null hypothesis was tested by calculating the probability that the global maximum $X_{\max}$ exceeded its observed value, using Adler (1981), Theorem 6.9.1, page 160. The number of isolated regions of activation was investigated by comparing the observed DT characteristic of excursion sets of $X(\mathbf{t})$ with its expected value as given by Adler (1981), Theorem 5.3.1, page 111.

It has been claimed by some workers that the standard deviation $\sigma(\mathbf{t})$ of the individual images is not stationary. This claim can be investigated using the standardised sums of squares image

$$U(\mathbf{t}) = (p-1)S^2(\mathbf{t})/\hat{\sigma}^2.$$

If $\sigma(\mathbf{t})$ is stationary then $U(\mathbf{t})$ is a $\chi^2$ field with $n = p - 1$ degrees of freedom, independent of $X(\mathbf{t})$. The global maximum $U_{\max}$ and the global minimum $U_{\min}$ can be used as a test statistic for a few regions of high or low standard deviation, respectively; their approximate null distributions are given by Corollary 3.4. The number of isolated regions of high or low standard deviation can be investigated by comparing the observed DT characteristic of excursion sets of $U(\mathbf{t})$ with its expected value as given by Theorem 3.5.

If the standard deviation $\sigma(\mathbf{t})$ is not stationary then we can use the image

$$T(\mathbf{t}) = p^{\frac{1}{2}}\bar{\Delta}(\mathbf{t})/S(\mathbf{t})$$

which is a $t$ field with $m = p-1$ degrees of freedom under the null hypothesis $\mu(\mathbf{t}) = 0$. The global maximum $T_{\max}$ can be used as a test statistic for regions of high mean; its approximate null distribution is given by Corollary 5.3. The number of isolated regions of high mean can be investigated by comparing the observed DT characteristic of excursion sets of $T(\mathbf{t})$ with its expected value as given by Theorem 5.4.

Talbot et al. (1990) investigated the regions of the brain showing an increased blood flow in response to a painfull heat stimulus. There were $p = 8$ subjects and the baseline condition of no heat stimulus was repeated twice on each subject. Worsley et al. (1992) used these images as a convenient control experiment to validate the models, since the null hypothesis of zero mean difference between the same two stimuli, $\mu(\mathbf{t}) = 0$, should be satisfied in this case, although the standard deviation $\sigma(\mathbf{t})$ may still be non-stationary. Horizontal slices of the images $X(\mathbf{t})$, $U(\mathbf{t})$ and $T(\mathbf{t}) = X(\mathbf{t})/\{U(\mathbf{t})/n\}^{\frac{1}{2}}$ are shown in Figure 2. Note that $U(\mathbf{t})$ is twice as 'rough' as $X(\mathbf{t})$, and that $T(\mathbf{t})$ has much sharper peaks than $X(\mathbf{t})$.

The $3 \times 3$ variance-covariance matrix $\mathbf{\Lambda}$ of the partial derivatives of $X(\mathbf{t})$ was estimated numerically and found to be in good agreement with a theoretical estimate based on the known resolution of the PET camera. The unitless parameter $\lambda(C)\det(\mathbf{\Lambda})^{\frac{1}{2}}$ was estimated to be 1564. The observed global maxima and minima of these fields, together with their exceedence probabilities, are given in Table 1. There is no evidence against the null hypotheses, apart from some evidence of regions of high standard deviation ($U_{\max} = 52.4$). The observed DT characteristic of the excursion sets for several values of the threshold level, together with their expected values, are given in Figure 3. In all cases they seem to be in reasonable agreement.

The analyses were repeated on data from the same study, but for the difference between a painful heat stimulus minus a control stimulus. Horizontal slices of the images $X(\mathbf{t})$, $U(\mathbf{t})$ and $T(\mathbf{t}) = X(\mathbf{t})/\{U(\mathbf{t})/n\}^{\frac{1}{2}}$ are shown in Figure 4; one subject had missing values and so $p = 7$. The test statistics are given in Table 1, and they indicate that there has been an increase in mean activation ($X_{\max} = 4.99$) and some regions of high standard deviation ($U_{\max} = 38.9$). The test based on $T_{\max}$ failed to detect an increase in mean activation. This can be explained by some simulation results of Worsley et al. (1992) which showed that a test based on $T_{\max}$ was not as powerful as that based on $X_{\max}$ if indeed the standard deviation was stationary. Plots of the DT characteristics, shown in Figure 5, show considerable discrepancies between the observed values of $X(\mathbf{t})$ and $T(\mathbf{t})$ and their expectations under the null hypothesis $\mu(\mathbf{t}) = 0$. Worsley et al. (1992) concluded on the basis of Figure 5(a) that there were three isolated regions of activation caused by the heat stimulus. The horizontal slices in Figures 2 and 4 were in fact chosen to pass through one of these regions, the left cingulate, which is approximately 3 centimetres left frontal of the centre of the slice.

TABLE 1. *Global maxima and minima, $z$, of the fields $Z(\mathbf{t})$ from the pain study, together with the expected DT characteristic of excursion sets above $z$.*

| | | Global minimum | | Global maximum | |
|---|---|---|---|---|---|
| $Z(\mathbf{t})$ | stimuli | $z$ | $\mathrm{E}(\chi(A_z(Z,C)))$ | $z$ | $\mathrm{E}(\chi(A_z(Z,C)))$ |
| $X(\mathbf{t})$ | no activation | -3.47 | 1.14 | 4.16 | 0.121 |
| $X(\mathbf{t})$ | heat stimulus | -4.32 | 0.066 | 4.99 | 0.0039 |
| $U(\mathbf{t})$ | no activation | 0.0814 | 0.973 | 52.4 | 0.00013 |
| $U(\mathbf{t})$ | heat stimulus | 0.0375 | 1.703 | 38.9 | 0.012 |
| $T(\mathbf{t})$ | no activation | -17.9 | 0.105 | 7.4 | 2.65 |
| $T(\mathbf{t})$ | heat stimulus | -42.6 | 0.037 | 15.6 | 0.72 |

# 7 Appendix

For the expected DT characteristic, we shall need the following lemmata in order to find the expected determinant of the matrix of second order derivatives, $\mathbf{D}_{N-1}$. Let $\mathbf{A}$ be an $N \times N$

matrix, and define $\mathrm{detr}_j(\mathbf{A})$ to be the sum of the determinant of all the $j \times j$ principal minors of $\mathbf{A}$, so that $\mathrm{detr}_N(\mathbf{A}) = \det(\mathbf{A})$, $\mathrm{detr}_1(\mathbf{A}) = \mathrm{tr}(\mathbf{A})$ and $\mathrm{detr}_0(\mathbf{A})$ is defined to be one. Note that $(-1)^j \mathrm{detr}_{N-j}(\mathbf{A})$ is the coefficient of $x^j$ in the characteristic polynomial $\det(\mathbf{A} - x\mathbf{I}_N)$, and $\mathrm{detr}_j(\mathbf{A})$ equals the product of $j$ eigen values of $\mathbf{A}$, summed over all possible subsets of $j$ eigen values of $\mathbf{A}$.

**Lemma 7.1** *(a) If* $\mathbf{H} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{\Sigma}))$ *and* $\mathbf{B}$ *is a fixed* $N \times k$ *matrix then*

$$\mathbf{B}'\mathbf{H}\mathbf{B} \sim \mathrm{Normal}_{k\times k}(\mathbf{0}, \mathbf{M}(\mathbf{B}'\mathbf{\Sigma}\mathbf{B})).$$

*(b) If* $\mathbf{H} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{I}_N))$ *then*

$$\mathrm{E}(\det(\mathbf{H})) = \frac{(-1)^j (2j)!}{2^j j!}$$

*if* $N = 2j$ *is even, and zero if* $N$ *is odd.*

*Proof.* The result (a) is proved by expanding and using Lemma 3.1(b), and result (b) is given by Adler (1981), Lemma 5.3.2, page 110. □

The following result is a generalisation of the Corollary to Lemma 5.3.2 of Adler (1981), page 110:

**Lemma 7.2** *Let* $\mathbf{H} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{I}_N))$ *and let* $\mathbf{A}$ *be a fixed symmetric* $N \times N$ *matrix. Then*

$$\mathrm{E}(\det(\mathbf{A} + \mathbf{H})) = \sum_{j=0}^{[N/2]} \frac{(-1)^j (2j)!}{2^j j!} \mathrm{detr}_{N-2j}(\mathbf{A}).$$

*Proof.* Let $\mathbf{B}$ be an orthonormal matrix such that $\mathbf{B}'\mathbf{A}\mathbf{B} = \mathbf{L} = \mathrm{diag}(l_1, \ldots, l_N)$ is a diagonal matrix of eigen values of $\mathbf{A}$. Then

$$\det(\mathbf{A} + \mathbf{H}) = \det(\mathbf{B}'(\mathbf{A} + \mathbf{H})\mathbf{B}) = \det(\mathbf{D} + \mathbf{H}^*)$$

where $\mathbf{H}^* = \mathbf{B}'\mathbf{H}\mathbf{B} \sim \mathrm{Normal}_{N\times N}(\mathbf{0}, \mathbf{M}(\mathbf{I}_N))$ by Lemma 7.1(a). Now since $\mathbf{L}$ is diagonal then $\det(\mathbf{L} + \mathbf{H}^*)$ can be expanded in terms of products of the determinant of each $k \times k$ principal minor of $\mathbf{H}^*$ with the $N-k$ members of $\{l_1, \ldots, l_N\}$ corresponding to the remaining rows and columns not included in the principal minor. By Lemma 7.1(a) the distribution of any $k \times k$ principal minor of $\mathbf{H}^*$ is $\mathrm{Normal}_{k\times k}(\mathbf{0}, \mathbf{M}(\mathbf{I}_k))$. Since the expected value of its determinant depends only on $k$, then using Lemma 7.1(b) with $k = 2j$ we obtain the result. □

**Lemma 7.3** *Let* $\mathbf{P} \sim \mathrm{Wishart}_N(\mathbf{I}_N, \nu)$ *and let* $a$ *be a fixed scalar. Then*

$$\mathrm{E}(\mathrm{detr}_j(\mathbf{P})) = \binom{N}{j} \frac{\nu!}{(\nu - j)!},$$

*where division by the factorial of a negative integer is treated as multiplication by zero.*

*Proof.* Each $j \times j$ principal minor of $\mathbf{P}$ has a $\text{Wishart}_j(\mathbf{I}_j, \nu)$ distribution. From standard multivariate statistics, the expected determinant of such a matrix is $\nu!/(\nu - j)!$ (see for example Anderson(1984), page 265). $\square$

**Lemma 7.4** *Let* $\mathbf{P} \sim \text{Wishart}_N(\mathbf{I}_N, \nu)$ *and* $\mathbf{Q} \sim \text{Wishart}_N(\mathbf{I}_N, \eta)$ *independently, and let* $a$ *be a fixed scalar. Then*

$$\mathrm{E}(\mathrm{detr}_j(\mathbf{P} + a\mathbf{Q})) = \frac{N!}{(N-j)!} \sum_{k=0}^{j} \binom{\eta}{k} \binom{\nu}{j-k} a^k,$$

*where division by the factorial of a negative integer is treated as multiplication by zero.*

*Proof.* Let $\mathbf{B}$ be an orthonormal matrix such that $\mathbf{B}'\mathbf{Q}\mathbf{B} = \mathbf{L} = \mathrm{diag}(l_1, \ldots, l_N)$ is a diagonal matrix of eigen values of $\mathbf{Q}$. Then

$$\mathrm{detr}_j(\mathbf{P} + a\mathbf{Q}) = \mathrm{detr}_j(\mathbf{B}'(\mathbf{P} + a\mathbf{Q})\mathbf{B}) = \mathrm{detr}_j(\mathbf{P}^* + a\mathbf{L}),$$

where $\mathbf{P}^* = \mathbf{B}'\mathbf{P}\mathbf{B}$. If $\mathbf{Q}$ is fixed then $\mathbf{B}$ is fixed and so $\mathbf{P}^* \sim \text{Wishart}_N(\mathbf{I}_N, \nu)$. From Lemma 7.3 the expected determinant of each $(j-k) \times (j-k)$ principal minor of $\mathbf{P}^*$ is $\nu!/(\nu - j + k)!$. Thus we can expand $\mathrm{E}(\mathrm{detr}_j(\mathbf{P} + a\mathbf{Q}))$ in powers of $a$ as follows:

$$\mathrm{E}(\mathrm{detr}_j(\mathbf{P}^* + a\mathbf{L})|\mathbf{Q}) = \sum_{k=0}^{j} \binom{N-k}{j-k} \frac{\nu!}{(\nu - j + k)!} \mathrm{detr}_k(\mathbf{L}) a^k.$$

Using the fact that $\mathrm{detr}_k(\mathbf{L}) = \mathrm{detr}_k(\mathbf{Q})$ and taking expectations over $\mathbf{Q}$ using Lemma 7.3, gives the result. $\square$

**Lemma 7.5** *Let* $\mathbf{P} \sim \text{Wishart}_N(\mathbf{I}_N, \nu)$, $\mathbf{Q} \sim \text{Wishart}_N(\mathbf{I}_N, \eta)$ *and* $\mathbf{H} \sim \text{Normal}_{N \times N}(\mathbf{0}, \mathbf{M}(\mathbf{I}_N))$ *independently, and let* $a$ *and* $b$ *be fixed scalars. Then*

$$\mathrm{E}(\det(\mathbf{P} + a\mathbf{Q} + b\mathbf{H})) = \sum_{j=0}^{[N/2]} \frac{(-1)^j N!}{2^j j!} b^{2j} \sum_{k=0}^{N-2j} \binom{\eta}{k} \binom{\nu}{N-2j-k} a^k.$$

*Proof.* Holding $\mathbf{P}$ and $\mathbf{Q}$ fixed and applying Lemma 7.2 with $\mathbf{A} = \mathbf{P} + a\mathbf{Q}$ we get

$$\mathrm{E}(\det(\mathbf{P} + a\mathbf{Q} + b\mathbf{H})) = \sum_{j=0}^{[N/2]} \frac{(-1)^j (2j)!}{2^j j!} b^{2j} \mathrm{E}(\mathrm{detr}_{N-2j}(\mathbf{P} + a\mathbf{Q})).$$

Applying Lemma 7.4 gives the result. $\square$

# References


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

Adler, R.J. and Hasofer, A.M. (1976). Level crossings for random fields. *Annals of Probability,* **4**, 1-12.

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd edition). Wiley, New York.

Beaky, M.M., Scherrer, R.J. and Villumsen, J.V. (1992). Topology of large scale structure in seeded hot dark matter models. *Astrophysical Journal*, **387**, 443-448.

Gott, J.R., Park, C., Juskiewicz, R., Bies, W.E., Bennett, D.P., Bouchet, F.R. and Stebbins, A. (1986). Topology of microwave background fluctuations: theory. *Astrophysical Journal*, **352**, 1-14.

Hamilton, A.J.S. (1988). The topology of fractal universes. *Publications of the Astronomical Society of the Pacific*, **100**, 1343-1350.

Hamilton, A.J.S., Gott, J.R. and Weinberg, D. (1986). The topology of large-scale structure in the universe. *Astrophysical Journal*, **309**, 1-12.

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

Mandelbrot, B.B. (1982). *The Fractal Geometry of Nature.* Freeman, New York.

Smoot, G.F., Bennett, C.L., Kogut, A., Wright, E.L., Aymon, J., Boggess, N.W., Cheng, E.S., De Amici, G., Gulkis, S., Hauser, M.G., Hinshaw, G., Jackson, P.D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Lowenstein, K., Lubin, P., Mather, J., Meyer, S.S., Moseley, S.H., Murdock, T., Rokke, L., Silverberg, R.F., Tenorio, L., Weiss, R. and Wilkinson, D.T. (1992). Structure in the *COBE* differential microwave radiometer first-year maps. *Astrophysical Journal*, **396**, L1-L5.

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., Evans, A.C., Marrett, S. and Neelin, P. (1993). Detecting changes in random fields and applications to medical images. *Journal of the American Statistical Association*, (submitted for publication).

**Figure 1**. The DT characteristic of an artificial image in $\mathbb{R}^2$. (a) The image, with colour bar in (e); darker shading denotes higher values. Local maxima are indicated by L and the global maximum is indicated by G. (b) Excursion set A (shaded areas) above a threshold $z = 3.2$. Points which are counted in $\chi_0(A)$, and contribute -1 to $\chi(A)$, are indicated by O. Points which are counted in $\chi_1(A)$, and contribute +1 to $\chi(A)$, are indicated by X. Since the excursion set does not touch the boundary the DT characteristic equals the Euler characteristic, which counts the number of isolated regions minus the number of 'holes', giving $\chi(A) = 2$. The number of local maxima greater then 3.2 is $M^+ = 4$. (b) As the threshold is increased to $z = 4.0$ the holes disappear and the DT characteristic counts the number of local maxima, giving $\chi(A) = M^+ = 4$. (c) At even higher levels $z = 5.6$, the DT characteristic takes the value one if the global maximum exceeds $z$ and zero otherwise, giving $\chi(A) = M^+ = 1$. (e) A plot of the DT characteristic $\chi(A)$ against $z$. For $z > 2.7$ the excursion set does not touch the boundary and the DT characteristic equals the Euler characteristic.

**Figure 2** Horizontal slices of the three dimensional Gaussian field $X(\mathbf{t})$ (a), the $\chi^2$ field $U(\mathbf{t})$ (b), and the $t$ field $T(\mathbf{t}) = X(\mathbf{t})/\{U(\mathbf{t})/n\}^{\frac{1}{2}}$ (c), as defined in section 6. The slice is taken roughly mid way through the brain, 3.2 cms above the anterior commissure-posterior commissural line. Darker shading denotes higher values. Note that the scale on the $\chi^2$ field with $n$ degrees of freedom (d.f.) has been transformed to $\{U(\mathbf{t})/n\}^{\frac{1}{2}}$ to give a better rendering of the image. The data shown is from a study with $p = 8$ subjects in which the two stimuli were identical, so that we expect to see no increase or decrease in the means of $X(\mathbf{t})$ and $T(\mathbf{t})$.

**Figure 3** The DT characteristics of excursion sets of the three dimensional fields in Figure 2, as a function of the level of the excursion set, together with the expected DT characteristic under the null hypothesis of no mean activation. The maxima and minima are marked with arrows. Note that the scale on the $\chi^2$ field with $n$ degrees of freedom (d.f.) has been transformed to $\{U(\mathbf{t})/n\}^{\frac{1}{2}}$. The observed and expected values appear to be in reasonable agreement.

**Figure 4** Horizontal slices of the three dimensional fields, as in Figure 2, but for the difference between a painfull heat stimulus and a control stimulus. Worsley et al. (1992) found evidence of significant activation in the Gaussian field $X(\mathbf{t})$ (a) in the left cingulate region, 3 centimetres left frontal of the centre of the slice.

**Figure 5** The DT characteristics of excursion sets of the three dimensional fields in Figure 4, for the difference between a painfull heat stimulus and a control stimulus. There are some discrepancies between between the observed and expected DT characteristics of $X(\mathbf{t})$ (a) and $T(\mathbf{t})$ (c). On the basis of (a), Worsley et al. (1992) estimated that there were three regions of activation caused by the heat stimulus.