Stoch Environ Res Risk Assess (2008) 22:809–822
DOI 10.1007/s00477-008-0242-6

ORIGINAL PAPER

Some new random field tools for spatial analysis

Robert J. Adler

Published online: 15 May 2008

Abstract This is a brief review, in relatively non-technical terms, of recent rather technical advances in the theory of random field geometry. These advances have provided a collection of explicit new formulae describing mean values of a variety of geometric characteristics of excursion sets of random fields, such as their volume, surface area and Euler characteristics. What is particularly important in these formulae is that whereas the previous theory covered only stationary, Gaussian random fields, the new theory requires neither stationarity nor, a fortiori, isotropy. Furthermore, it covers a wide class of non-Gaussian random fields. The formulae provided by these advances have a wide range of potential applications, including new techniques of parameter estimation, model testing, and thresholding for spatial and space-time functions. The paper reviews this theory, and provides brief descriptions of some of the applications.

Keywords Random field · Excursion set · Nodal domain · Euler characteristic · Stationary processes · Spatial analysis

1 Introduction

Random fields, a generic term that I shall use to describe random processes defined over parameter spaces with dimension great than one, appear throughout the modelling of spatial and space-time phenomena, whether space is two or three dimensional. To note just a handful of examples, we have

- Physical oceanography and hydrology, where, as in Adler et al. (1996), Rubin (2002), the spatial parameter might be the two dimensional water surface or a three dimensional body of water and the variable being measured might be water temperature and/or pressure. If both are being measured, then we have an example of a vector-valued random field.
- Atmospheric studies, where the random field might provide a model for wind spreads or airborne contaminants, with a theory as described in Daley (1991).
- Geostatistics and other earth sciences, where, as for oceanography, the spatial aspects of parameter spaces may be two or three dimensional (cf. Christakos 1991; Christakos 2000; Olea 1999).
- Astrophysics, where random field techniques have been heavily used in analysing the COBE (Cosmic Background Explorer) data, which measures the 'signature radiation' from the universe of 15 billion years ago (Smoot and Davidson 1993; Torres 1994; Vogeley et al. 1994). This is directional data, and so is realised as a random field on the two-dimensional sphere. Three dimensional astrophysical data has been generated by the Sloan digital sky survey (Gott et al. 2008) where, as we shall describe later, heavy use has been made of the geometrical theory that is at the centre of this review.
- Analysis of functional magnetic resonance imaging (fMRI) data. This is the archtypical example that will be used throughout this paper, and so I shall leave the details until later.

The theory of random fields has undergone a minor revolution over the past few years in terms of its ability to provide precise geometric information about the global

R. J. Adler (✉)
Faculties of Industrial Engineering & Management and Electrical Engineering, Technion-Israel Institute of Technology, 32000 Haifa, Israel
e-mail: robert@ieadler.technion.ac.il

behaviour of sample paths. While this work was motivated by applications, it has appeared primarily in the pure mathematical literature and, even by its esoteric standards, is not always very easy to read.

In this brief review we want to describe some of these results in simpler language, as well as discussing some of their applications.

The new results centre on the geometry of the so-called *excursion sets* or *nodal domains* of random fields. To describe these, we first define a *random field* to be a random function f defined over a parameter space M, which we shall always take to be a bounded region in N-dimensional Euclidean space, $\mathbb{R}^N$. The random field itself might be real or vector valued. In general, we let it take values in $\mathbb{R}^k, k \geq 1$. If D is a subset of $\mathbb{R}^k$, then the excursion set of f in D and over M is defined by

$$A_D \equiv A_D(f, M) \triangleq \{x \in M : f(x) \in D\}. \tag{1}$$

In the particular case that f is real-valued and D is the set $[u,\infty)$, we are looking at those points in the parameter space M at which f takes values larger than u, and write

$$A_u \equiv A_u(f, M) \triangleq \{x \in M : f(x) \geq u\}. \tag{2}$$

In general, we shall call D a *hitting set*.

Perhaps the first studied example of a real valued random field in two dimensions is due to Matérn (1960), who modelled forestry yields $f(x)$ as a function of a two-dimensional positional variable x. The excursion sets were then high yield geographical regions. Real valued random fields defined over $\mathbb{R}^3$ abound in environmental data, given, for example, by pollutant levels in space. If more than one pollutant is measured, then the random field becomes vector valued and the sets D defining excursion sets can become quite involved. For example, if interest is in regions where at least one of the pollutants takes some minimal value—say u_i for the ith pollutant, $i = 1,...,k$—then D will be of the rather simple, rectangular form $\prod_{i=1}^{k}[u_i, \infty)$. However, if interest lies in some combination of pollutants, then D will reflect this and will be, geometrically, as complicated as is the formula for the combination.

Examples with four dimensional parameter sets are also common: one needs only add time to examples that otherwise deal with three dimensional space. Space-time examples are important, and an example where the new theory has something to offer that the old theory did not. It is obvious that, in essentially all real scenarios, the structure of temporal dependence is quite different to that of spatial dependence. Therefore, models restricted to assumptions such as isotropy become problematic in space-time. Indeed, even stationarity can be a problem in dealing with processes where, perhaps because of external factors, there is no intrinsic temporal stationarity, despite there being spatial stationarity, or vice versa. We shall close the introduction with a (perhaps surprisingly) simple example of an eight-dimensional random field, which highlights the need for a theory that does away with classical assumptions on the simplicity of the parameter space and directionally homogeneous stochastic behaviour of the random field.

However, since we have not described what the new theory does, we shall do this first. Given that excursion sets are of intrinsic interest, it would be nice to understand a little of their properties, such as how many components they have, how large they are, what is the size of their boundaries, etc. There are a number of ways to quantify these properties, and for N dimensional sets geometers have developed $N + 1$ numerical quantifies, known under a variety of names, including Quermassintegrales, Minkowski or Steiner functionals, integral curvatures, intrinsic volumes, and *Lipschitz–Killing curvatures*, hereafter LKCs. The differences between them are only of ordering and scaling, and although it is the longest of all names, we shall choose LKCs, for consistency with our basic reference (Adler and Taylor 2007). We shall denote the LKCs of a N-dimensional set A by $\mathscr{L}_0(A), \ldots, \mathscr{L}_N(A)$ and, in many ways, they can be thought of as measures of the 'j-dimensional sizes' of A. For example, when $N = 2$ they have the following meanings:

- $\mathscr{L}_2(A)$ is the two dimensional area of A.
- $\mathscr{L}_1(A)$ is half the boundary length of A.
- $\mathscr{L}_0(A)$ is the Euler characteristic of A, which in two dimensions is given by

$$\begin{aligned}\mathscr{L}_0(A) = {} & \#\{\text{connected components in } A\} \\ & - \#\{\text{'holes' in } A\}\end{aligned}. \tag{3}$$

When $N = 3$,

- $\mathscr{L}_3(A)$ is the three dimensional volume of A.
- $\mathscr{L}_2(A)$ is half the surface area of A.
- $\mathscr{L}_1(A)$ is twice the *caliper diameter* of A, where the caliper diameter of a convex A is defined by placing the solid between two parallel planes (or calipers), measuring the distance between the planes, and averaging over all rotations of A. Thus the caliper diameter of a sphere is just the usual diameter, while for a rectangular box of size $a \times b \times c$ it is $(a + b + c)/2$, half the 'volume' used by airlines to measure the size of luggage.
- $\mathscr{L}_0(A)$ is again the Euler characteristic of A, which in three dimensions is given by

$$\begin{aligned}\mathscr{L}_0(A) = {} & \#\{\text{connected components in } A\} \\ & - \#\{\text{'handles' in } A\} + \#\{\text{'holes' in } A\}.\end{aligned} \tag{4}$$

We shall give further definitions for general dimensions in Sect. 3.

What the new theory provides is explicit formulae for the expectations of the LKC's of excursion sets—that is for

$$\mathbb{E}\{\mathscr{L}_j(A_D(f,M))\}, \tag{5}$$

—for a wide class of random fields f, over a large class of parameter sets M, and for a large class of sets D.

Without yet being explicit about the precise structure of these formulae, we discuss some of their uses.

1.1 Signal detection and thresholding

Consider the classical signal + noise paradigm, which considers the model

$$f(x) = s(x) + \eta(x). \tag{6}$$

The signal s is deterministic, and, purely for reasons of exposition, is assumed to take primarily positive values if, in fact, it is present. The noise η is a mean zero random field of background noise, the statistical properties of which are assumed to be known. The first problem here is determining whether or not the signal is at all present, and, in view of the assumed positivity of s, a useful statistic for testing this is the threshold statistic

$$\sup_{x\in M} f(x).$$

(In fact, the supremum is the maximum likelihood statistics for this test, if η is Gaussian white noise smoothed with a filter with shape matching that of the signal. This is the 'matched filter theorem' of signal processing.)

In order to use this statistic, or, indeed, to perform any statistical thresholding technique, one needs to know how to compute, at least for large values of u, the *excursion probabilities*

$$\mathbb{P}\left\{\sup_{x\in M} f(x) \geq u\right\}, \tag{7}$$

when no signal is present (By 'large' we mean large enough for this probability to be less that about 0.10, the level at which one usually begins to talk about statistical significance).

It has long been acknowledged that, for many random fields,

$$\left|\mathbb{P}\left\{\sup_{x\in M} f(x) \geq u\right\} - \mathbb{E}\{\mathscr{L}_0(A_u(f,M))\}\right| \leq \text{error}\,(u), \tag{8}$$

where the *error* function is of a smaller order than both of the other terms. This means that the excursion probabilities (Eq. 7), which are intrinsically uncomputable, can effectively be replaced by an expectation which might be more accessible.

Two of the contributions of the recent theory are the computation of this expectation in wide generality, and a full proof that, at least for Gaussian random fields, not only finally rigorously established (Eq. 8) but even identified the structure of the function *error* (cf. Eq. 35 below).

1.2 Parameter estimation

Assume that we are dealing with a random field, the general distributional structure of which is known, but for which values of parameters are unknown. There are standard statistical techniques for estimating the means and variances of random fields, as well as for estimating covariance functions, which we shall denote as

$$C(x,y) = \mathbb{E}\{f(x)f(y)\},$$

assuming, from now on, and for notational convenience, that any non-zero mean has already by removed from f [When f is stationary, we shall adopt the usual notational inconsistency of writing $C(x-y) \equiv C(x, y)$]. However, estimates of covariance functions are often not well behaved, particularly in the non-stationary case, when they are notoriously unreliable. Furthermore, in many situations, one needs to know only certain aspects of the covariance function, and not the entire function itself. For example, if one is concerned with excursion probabilities (Eq. 7) for large u, and if f is stationary and Gaussian, then it turns out that one needs only to know certain derivatives of C in the neighbourhood of the origin, or, equivalently, the second spectral moments of f (cf. Eqs. 23 and 30).

As opposed to simple means and variances, which describe static properties of f (i.e., properties of f at specific points) covariance and spectral parameters describe dynamic behaviour over the entire parameter space, and much can be done with them, without knowing the full covariance function. From a statistical point of view, it is, of course, more efficient to estimate a handful of parameters than an entire function.

However, it is precisely these parameters that appear in the formulae for the expected LKCs of excursion sets, and so those formulae provide a tool for estimating these parameters, by comparing expectations to empirically observed properties of excursion sets.

1.3 Model testing

In this problem, which is in some sense an extension of the parameter estimation problem, we are in the scenario of dealing with a random field whose precise structure is unknown, but it assumed to belong to some small class of models. Since the forms of the expected LKCs of excursion sets are often quite different for different models, comparing empirical data to various theoretical models can be used to choose among models. This technique has, for example, been used quite heavily in the astrophysical literature (e.g., Smoot and Davidson 1993; Torres 1994;

Vogeley et al. 1994) and we shall give an example of how to use it in Sect. 7.

Finally, we complete this introduction of a quite concrete example of a random field over a rather complicated subset of $\mathbb{R}^8$, in an attempt to convince the possibly sceptical, applied reader that there is actually is a need for the practitioners' tool kit to include random fields over high dimensional parameter spaces.

1.4 An eight dimensional example in scale-space

We shall take an example from fMRI brain imaging, since we can then refer you to Shafie et al. (2003) for further details, where the approach is notationally quite consistent with this review. The problem, however, is quite generic.

In fMRI procedures, measurements are taken at a large number of voxels in the (three-dimensional) brain. For simplicity, we shall assume one such measurement per voxel. However, these are not actually point measurements, but each measurement is rather a local average, taken over a small ellipsoidal neighbourhood of the voxel. Each such ellipsoid has three principle axes, the lengths of each of which can be chosen by the technician operating the scanner, giving three more dimensions to the problem. The final two dimensions come from possible rotations of the ellipsoid. Thus, what seemed originally to be an observation taken at a point in the three dimensional brain, is really an observation taken in the eight dimensional space

$$\text{brain} \times \text{axes} \times \text{rotation}. \tag{9}$$

Why is this important? Suppose we are interested in either a signal detection or thresholding problem as described above, looking for signs of unusual behaviour somewhere in the brain. The technician, as he plays with the sizes of the axes and their rotations, is clearly playing with the resolution of the image, and the classic 'uncertainty principle' comes into play. The higher the resolution (i.e., the smaller the lengths of the principle axes of the ellipsoids) the lower is the statistical reliability of the measurements at individual voxels. Thus, in changing these parameters, the technician is implementing a control procedure which can impact quite strongly on the excursion probability (Eq. 7) and so on any thresholding and risk assessment that comes from it, unless this probability is calculated in the full, eight dimensional, scenario.

Note that the covariance function of this random field will be neither stationary nor isotropic. Even if one treats the basic, three-dimensional structure of the brain as isotropic (and even this is suspect) the additional dimensions added here behave very differently. Consequently, handling this example calls for the development of a theory of non-stationary random fields.

Note also that while we chose the fMRI example to make our point, the problem itself is generic to any situation in which data collection involves a compromise between resolution and smoothness. To put things more bluntly, the problem arises whenever raw spatial data is made up of some sort of local average, which is an almost ubiquitous situation.

2 Gaussian and related random fields

The basic random fields with which we shall work are Gaussian ones, so that both their univariate *and multivariate* distributions are all normal. In the real valued case, this means that the random function $f : M \to \mathbb{R}$ has the property that all collections

$$f(x_1), \ldots, f(x_k),$$

of random variables have multivariate Gaussian distributions, for all $k \geq 1$ and all $x_1, \ldots, x_k \in M$. A similar definition holds for vector valued random fields. We have emphasised the words 'and multivariate' above on purpose. Throughout the literature one finds examples in which authors make a pointwise transformation to normality, and then claim normality for their process. The argument, for real valued random fields, usually goes as follows:

Collect data points of the random field f at points $x_i \in M$, $i = 1, \ldots, n$, and form the empirical distribution function

$$\widehat{F}_n(u) = \frac{\#\{i : f(x_i) \leq u\}}{n}.$$

Letting $\mathbf{\Phi}$ denote the distribution function of a standard normal random variable, define the 'Gaussianised', or sometimes just 'standardised', data to be

$$\widetilde{f}(x_j) \triangleq \mathbf{\Phi}^{-1}\left(\widehat{F}_n(f(x_j))\right), \quad j = 1, \ldots, n. \tag{10}$$

If the random field is stationary, and the sample size is large enough, then it certainly is true that the random field $\widetilde{f}$ has univariate distributions that are close to standard normal. However, no claim at all can be made about the *multivariate* distributions, and so, a fortiori, about $\widetilde{f}$ as a *process*, and the application of Gaussian process theory can lead to quite significant errors. (The same is true even if the empirical distribution function in Eq. 10 is replaced with the true, theoretical one, as we shall see in Sect. 7.)

2.1 Gaussian related fields

Leaving the Gaussian scenario is, however, not all that easy to do, and the new theory leaves it in a fashion that, while somewhat limited, turns out to be broad enough to cover many, if not most, statistical applications of random fields.

To be more precise, we shall call a random field $f : M \to \mathbb{R}^d$ a *Gaussian related* field if we can find a vector valued Gaussian random field,

$$g(x) = (g_1(x), \ldots, g_k(x)) : M \to \mathbb{R}^k,$$

with independent, identically distributed components having zero means and unit variances, and a function

$$F : \mathbb{R}^k \to \mathbb{R}^d, \tag{11}$$

such that f has the same multivariate distributions as $F(g)$.

When $k = 1$, or, in general $k = d$ and F is invertible, then the corresponding Gaussian related process is not much harder to study than the original Gaussian one, since what happens at the level u for f is precisely what happens at the uniquely defined level $F^{-1}(u)$ for g. In fact, this actually falls into the class of transformations that we warned against above, but in this case the transformation argument ifs completely valid.

In the more interesting cases in which F is not invertible, $f = F(g)$ can provide a process that is qualitatively different to g. For example, consider the following three choices for F, where in the third we set $k = n + m$.

$$\sum_1^k x_i^2, \quad \frac{x_1\sqrt{k-1}}{(\sum_2^k x_i^2)^{1/2}}, \quad \frac{m\sum_1^n x_i^2}{n\sum_{n+1}^{n+m} x_i^2}. \tag{12}$$

The corresponding random fields are known as χ^2 fields with k degrees of freedom, the T field with $k - 1$ degrees of freedom, and the F field with n and m degrees of freedom. These three random fields all have very different spatial behaviour, and each is as fundamental to the statistical applications of random field theory as is its corresponding univariate distribution to standard statistical theory. In each of these three cases, as in general for a Gaussian related random field, there is no simple pointwise transformation which will transform it to a real valued Gaussian field.

Note that for a Gaussian related field f the excursion sets A_D can be rewritten as

$$\begin{aligned} A_D(f, M) &= A_D(F(g), M) \\ &= \{x \in M : (F \circ g)(t) \in D\} \\ &= \{x \in M : g(x) \in F^{-1}(D)\} \\ &= A_{F^{-1}(D)}(g, M). \end{aligned} \tag{13}$$

Thus, for example, the excursion set of a *real valued* non-Gaussian $f = F \circ g$ above a level u is equivalent to the excursion set for a *vector valued* Gaussian g in $F^{-1}([u, \infty)) \in \mathbb{R}^k$.

We shall return to this important point later.

2.2 Regularity assumptions

Although we do not want to get into detailed technicalities here, we do need to put some restrictions on the central objects of this review. We start with the parameter spaces.

2.2.1 Parameter spaces

The general theory allows M to be what is known as a Whitney stratified manifold, satisfying some mild side conditions. Rather than define what these are, we shall give a list of examples which should cover most of the situations found in environmental engineering.

(i) Finite simplicial complexes. This includes sets such as N-dimensional rectangles of the form

$$T = \prod_{i=1}^{M} [0, T_i], \tag{14}$$

or unions of a finite number of such rectangles, as well as most sets with flat boundaries.

(ii) Reimannian manifolds, with or without boundary. This includes sets such as N-dimensional balls and spheres, and smooth deformations of them.

(iii) Sets that can be written in the form

$$M = \bigcup_{i=1}^{N} \partial_i M, \tag{15}$$

where $\partial_i M$ is an i-dimensional set (the i-dimensional 'boundary' of M) that fits into the one of the first two categories. For example, the eight dimensional example of Eq. 9, which fits into neither of the first two categories, requires this level of generality.

We place on the hitting sets D appearing in the definition (1) of excursion sets one out of the above possible assumptions on M.

2.2.2 Random fields

The basic building blocks of all the random fields we consider are smooth Gaussian random fields, $f : M \in \mathbb{R}^N \to \mathbb{R}^k$. The first assumption is a minor one, and only for notational convenience, that all means be fixed at zero. The second is more serious, and it is that f have constant variance throughout M. Note that this is a much weaker assumption than stationarity, and is achievable in general by replacing a random field that does not have constant variance by

$$f(x) \to f_{\text{standardised}}(x) = \frac{f(x)}{(\mathbb{E}\{f^2(x)\})^{1/2}}.$$

This transformation is often useful and involves no loss of information. For example, the excursion sets above zero are the same for f and $f_{\text{standardised}}$.

However, the fact that the new theory requires only constant variance and not full stationarity is extremely useful, since in general there is no simple transformation which will take a non-stationary field to a stationarity one. (See, however, Sampson and Guttorp (1992), which shows that if one is prepared to pay the price of moving to a higher dimensional scenario than the one inherent to the problem, there is a way around this.)

The final assumption of consequence on the Gaussian processes, and the most important one, is that each of the k components of f is twice differentiable, and that these derivatives are themselves continuous. Some additional minor assumptions of non-degeneracy also need to be made, but since these are of little practical importance we direct the interested reader to Chapter 11 of Adler and Taylor (2007) or Chapter 4 of Adler et al. (2008) for details.

2.2.3 *Gaussian related fields*

Recall that Gaussian related fields were defined as pointwise transformations of vector valued Gaussian fields. We have already assumed that the components of these fields are independent and identically distributed, and now add the assumption that the function F of Eq. 11 be twice continuously differentiable. As a consequence, all our Gaussian related fields are also twice continuously differentiable.

3 Tube formulae

In Sect. 1 we described the meanings of the Lipschitz–Killing curvatures $\mathscr{L}_j$ for two and three dimensional sets, but we also need a definition in higher dimensions as well. We shall also soon need another set of related geometric quantifiers, which we call Gaussian Minkowski functionals, and all of these will be defined in this section.

3.1 Lipschitz–Killing curvatures

The quickest way to define Lipschitz–Killing curvatures is via a basic result of integral geometry known as *Steiner's formula*. Steiner's formula deals with the N-dimensional volume of enlargements of sets, where the enlargement, or *tube*, of diameter $\rho > 0$ built around a set $A \in \mathbb{R}^N$ is the set

$$\text{Tube}(A, \rho) = \left\{ x \in \mathbb{R}^N : \min_{y \in A} \|x - y\| \leq \rho \right\}. \tag{16}$$

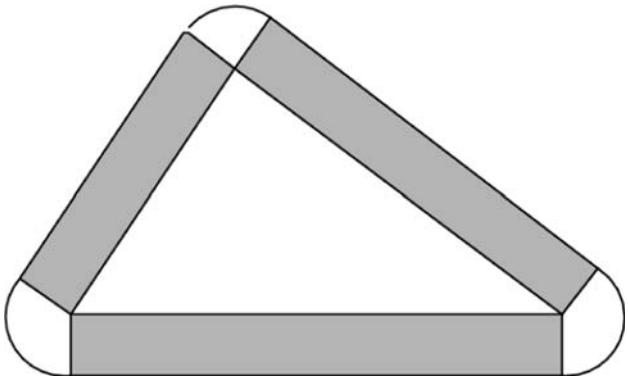

Fig. 1 The tube around a triangle

An example is given in Fig. 1, in which A is a solid triangle and Tube(A, ρ) is the larger triangular region with rounded corners. Note that Tube(A, ρ) always includes A itself. To see from where the terminology comes, think of A as being a curve in $\mathbb{R}^3$, in which case Tube(A, ρ) really does look like a (solid) physical tube.

Steiner's theorem says that if λ_N denotes volume in $\mathbb{R}^N$, then, for convex A of dimension dim (A),

$$\lambda_N(\text{Tube}(A, \rho)) = \sum_{j=0}^{\dim(A)} \omega_{N-j} \rho^{N-j} \mathscr{L}_j(A), \tag{17}$$

where $\omega_j = \pi^{j/2}/\Gamma(1 + j/2)$ is the volume of the unit ball in $\mathbb{R}^j$. This (finite) power series in ρ defines the LKCs, and by looking at examples, such as squares in the plane and cubes in $\mathbb{R}^3$, you should find is easy to convince yourself that the descriptions given earlier of the LKCs in two and three dimensions are reasonable.

There are versions of Steiner's formula for far more general sets than convex ones, in which case they are usually called 'tube formulae' and generally associated with the name of Hermann Weyl. For us it will suffice to know that, for small enough ρ, Eq. 17 holds for all the sets that we shall consider in this review.

It is quite easy to see from Eq. 17 that the LKCs satisfy a basic scaling relationship, in that

$$\mathscr{L}_j(\lambda A) = \lambda^j \mathscr{L}_j(A),$$

for $0 \leq j \leq \dim(A)$ and $\lambda > 0$, where $\lambda A = \{x : x = \lambda y, \text{ for some } y \in A\}$. Thus one can think of $\mathscr{L}_j(A)$ as some sort of measure of the 'j-dimensional volume' of A.

3.2 Gaussian Minkowski functionals

We need one more set of geometrical tools, closely related to LKCs, before we can describe the new results that are at the core of this review, and these are related to the hitting sets D of Eq. 1.

Rather than measuring the basic Euclidean sizes of the hitting sets, we want to measure aspects of their (Gaussian) probability content. In particular, let X be a vector of k independent, identically distributed, standard Gaussian random variables, and, for a nice set $A \in \mathbb{R}^k$, write

$$\gamma_k(A) = \mathbb{P}\{X \in A\} = \int_A \frac{e^{-\|x\|^2/2}}{(2\pi)^{k/2}} dx.$$

Defining tubes exactly as before, it turns out that there is a Taylor expansion for the probability content of tubes (due to Taylor 2006), which we shall write as

$$\gamma_k(\text{Tube}(A, \rho)) = \sum_{j=0}^{\infty} \frac{\rho^j}{j!} \mathscr{M}_j^k(A), \tag{18}$$

for small enough ρ. We call the coefficients $\mathscr{M}_j^k(A)$ the *Gaussian Minkowski functionals* (GMFs) of A. Clearly $\mathscr{M}_0^k(A) = \gamma_k(A)$. Note that, unlike the expansion (Eq. 17) in Steiner's theorem, this expansion is an infinite one.

To see how this expansion works in the simplest of cases, take $k = 1$ and $A = [u, \infty)$. Then, since

$$\gamma_1(\text{Tube}([u, \infty), \rho)) = \gamma_1([u - \rho, \infty))) = \Psi(u - \rho),$$

where

$$\Psi(u) = (2\pi)^{-1/2} \int_u^{\infty} e^{-x^2/2} dx$$

is the tail probability function for a standard Gaussian variable, a standard Taylor expansion along with properties of Hermite polynomials shows that

$$\mathscr{M}_j^k([u, \infty)) = H_{j-1}(u) \frac{e^{-u^2/2}}{\sqrt{2\pi}}. \tag{19}$$

Here H_n, $n \geq 0$, is the nth Hermite polynomial

$$H_n(x) = n! \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j x^{n-2j}}{j!(n-2j)!2^j}, \tag{20}$$

where $\lfloor a \rfloor$ is the largest integer less than or equal to a. When $n = -1$ we set

$$H_{-1}(x) = \sqrt{2\pi} \Psi(x) e^{x^2/2}. \tag{21}$$

In fact, it is not much harder to define the GMFs of more general sets that arise in the study of Gaussian related random fields. With the notation of Sect. 2.1 let $f = F(g)$ be such a field. Then, as we already noted above, it follows from Eq. 13, the excursion sets $A_u(f, M)$ are those of g in the hitting sets $F^{-1}([u, \infty))$, and so these are going to be of interest to us. If we now write $p_F(u)$ for the probability density of the random variable $F(g(x))$ (which for simplicity we shall assume does not depend on x) then

$$\mathscr{M}_j^k\big(F^{-1}([u, \infty))\big) = (-1)^{j-1} \frac{d^{j-1} p_F(y)}{dy^{j-1}}\Big|_{y=u},$$

for $j \geq 1$, while

$$\mathscr{M}_0^k\big(F^{-1}([u, \infty))\big) = \gamma_k\big(F^{-1}([u, \infty))\big).$$

Although these formulae seem simple, they often involve considerable delicate algebra to compute. For example, if $F(x) = \|x\|^2$, so that the Gaussian related field is χ^2 with k degrees of freedom, then, for $j \geq 1$,

$$\begin{aligned}
&\mathscr{M}_j^k\big(F^{-1}([u, \infty))\big) \\
&= \frac{u^{k-j} e^{-u^2/2}}{\Gamma(k/2) 2^{(k-2)/2}} \sum_{l=0}^{\lfloor \frac{j-1}{2} \rfloor} \sum_{m=0}^{j-1-2l} 1_{\{k \geq j-m-2l\}} \\
&\quad \times \binom{k-1}{j-1-m-2l} \frac{(-1)^{j-1+m+l}(j-1)!}{m!l!2^l} u^{2m+2l}.
\end{aligned} \tag{22}$$

For details of this and other cases see Chapter 15 of Adler and Taylor (2007] or Chapter 4 of Adler et al. (2008).

4 The main result: an expectation formula

We now come to the main result of the new theory, the fundamental starting point of which was in Jonathan Taylor's 2001 McGill PhD thesis (cf. Taylor 2001; Taylor 2006).

We present it first in the isotropic case. That is, we assume that $f : M \to \mathbb{R}^k$ is a Gaussian random field, the components of which are independent and identically distributed, have mean zero and unit variance, and are isotropic, with second spectral moment λ_2 defined by

$$\lambda_2 = \mathbb{E}\left\{ \left(\frac{\partial f(x)}{\partial x_i} \right)^2 \right\} = -\frac{\partial^2 C(x)}{\partial x_i^2}\Big|_{x=0}. \tag{23}$$

(By isotropy, this definition does not depend on i.)

Then, for parameter spaces M of dimension N, and hitting sets D of dimension k, and under the regularity assumptions of Sect. 2.2, we have

$$\begin{aligned}
&\mathbb{E}\{\mathscr{L}_i(A_D(f, M))\} \\
&= \sum_{j=0}^{N-i} \begin{bmatrix} i+j \\ j \end{bmatrix} \frac{\lambda_2^{i+j}}{(2\pi)^{j/2}} \mathscr{L}_{i+j}(M) \mathscr{M}_j^k(D),
\end{aligned} \tag{24}$$

where the so-called combinatorial 'flag coefficients' are defined by

$$\begin{bmatrix} n \\ j \end{bmatrix} = \binom{n}{j} \frac{\omega_n}{\omega_{n-j}\omega_j}.$$

Remaining for the moment in the isotropic case, we shall devote some space to explaining the meaning of Eq. 24 by considering some examples.

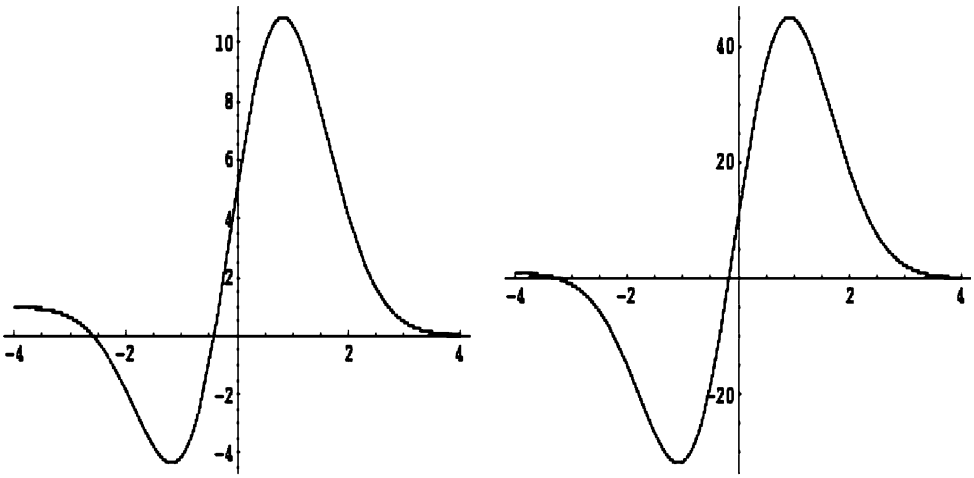

Fig. 2 Two EEC curves $\mathbb{E}\{\varphi(A_u)\}$ for Gaussian fields in two dimensions for different values of the second spectral moment λ_2. The horizontal axis gives values of u and the vertical axis the EEC

First of all, we take $k = 1$, so that we are dealing with a real valued random field, and also allow its variance to be σ^2, which is not necessarily 1. To make life even easier, we shall take the parameter space to be the N-dimensional rectangle T of Eq. 14, and shall concentrate on the case $i = 0$ in Eq. 24, so that we are looking only at Euler characteristics, rather than general LKCs. Finally, we shall take $D = [u, \infty)$. Thus, in summary, we are looking at the expected Euler characteristic of the excursion set, in a rectangle, of f above the level u.

To substitute into the right hand side of Eq. 24 we need to know how to compute the $\mathscr{L}_j(T)$ and $\mathscr{M}_j^1(D)$. The latter, however, are given by Eq. 19 and it is not hard to compute the former from Steiner's formula. However, to describe them, we need some notation.

In particular, we define a *face* J of T, of dimension k, by fixing $N-k$ coordinates, each of which is fixed at either the top or bottom of the rectangle. Thus, taking the three dimensional cube as a simple example, it has one three dimensional 'face', this being the full cube itself. It has six two dimensional faces, being its sides. The 12 edges are one dimensional faces, and the eight corners are zero dimensional faces. Now let $\mathscr{O}_k$ denote the $\binom{N}{k}$ k-dimensional faces which include the origin. Then it is not to hard to show, from Steiner's formula, that

$$\mathscr{L}_j\left(\prod_{i=1}^{N}[0, T_i]\right) = \sum_{J\in\mathscr{O}_j}|J|,$$

where $|J|$ is the (j-dimensional) volume of J.

Putting the pieces together now gives that, for isotropic random fields, $\mathbb{E}\{\varphi(A_u(f,T))\}$, or $\mathbb{E}\{\mathscr{L}_0(A_u(f,T))\}$, is given by

$$e^{\frac{-u^2}{2\sigma^2}}\sum_{k=1}^{N}\sum_{J\in\mathscr{O}_k}\frac{|J|\lambda_2^{k/2}}{(2\pi)^{(k+1)/2}\sigma^k}H_{k-1}\left(\frac{u}{\sigma}\right)+\Psi\left(\frac{u}{\sigma}\right), \tag{25}$$

or, if T is a cube of side length T rather than a rectangle, by

$$e^{\frac{-u^2}{2\sigma^2}}\sum_{k=1}^{N}\frac{\binom{N}{k}T^k\lambda_2^{k/2}}{(2\pi)^{(k+1)/2}\sigma^k}H_{k-1}\left(\frac{u}{\sigma}\right)+\Psi\left(\frac{u}{\sigma}\right). \tag{26}$$

To get a better feel for this equation, let us look at the cases $N = 2$ and $N = 3$, taking $\sigma^2 = 1$ for simplicity. In the two dimensionsal case, we obtain

$$\mathbb{E}\{\varphi(A_u)\} = \left[\frac{T^2\lambda_2}{(2\pi)^{3/2}}u+\frac{2T\lambda_2^{1/2}}{2\pi}\right]e^{\frac{-u^2}{2}}+\Psi(u). \tag{27}$$

Figure 2 gives two examples, both over the unit square, one with $\lambda_2 = 200$ and one with $\lambda_2 = 1{,}000$.

There are at least four general points that come out of from this example:

(i) You should note, for later reference, how the expression before the exponential term can be thought of as one of a number of different power series; one in T, one in u, and one in $\sqrt{\lambda_2}$.

(ii) The geometric meaning of the negative values of Eq. 27 are worth understanding. They are due to the excursion sets having, in the mean, more holes than connected components for (most) negative values of u.

(iii) Note the impact of the value of the second spectral moment. Large spectral moments lead to local fluctuations generating large numbers of small islands (or lakes, depending on the level at which the excursion set is taken) and this leads to larger variation in the values of $\mathbb{E}\{\varphi(A_u)\}$.

(iv) Note that, as $u \to \infty, \mathbb{E}\{\varphi(A_u)\} \to 0$. This is reasonable, since at high levels we expect that the excursion set will be empty, and so have zero Euler characteristic. On the other hand as $u \to -\infty$, $\mathbb{E}\{\varphi(A_u)\} \to 1$. The excursion set geometry behind this is simple. Once $u < \inf_T f(t)$ we have $A_u \equiv T$, and so $\varphi(A_u) = \varphi(T) = 1$.

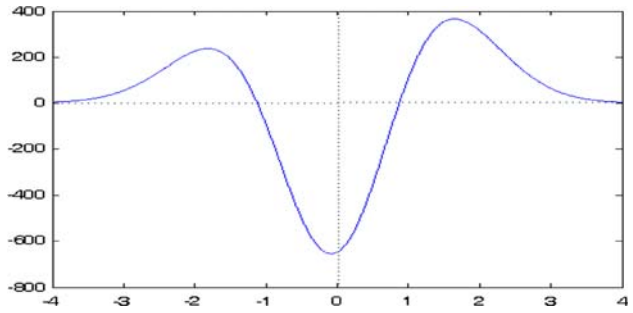

Fig. 3 EEC for a Gaussian field in three dimensions. Axes as for Fig. 2

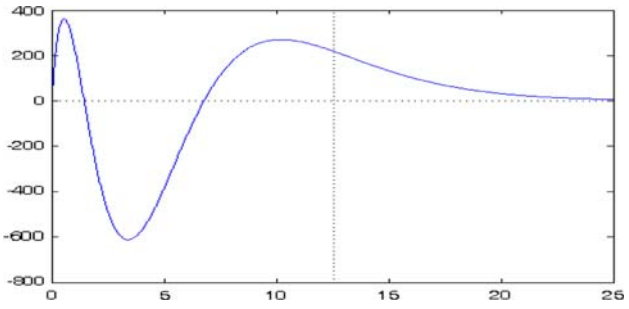

Fig. 4 EEC for a χ^2_5 field in three dimensions. Axes as for Fig. 2

In three dimensions, Eq. 26 becomes, again for $\sigma^2 = 1$,

$$\mathbb{E}\{\varphi(A_u)\} = \Psi(u) + e^{\frac{-u^2}{2}}\left[\frac{T^3\lambda_2^{3/2}}{(2\pi)^2}u^2 + \frac{3T^2\lambda_2}{(2\pi)^{3/2}}u + \frac{3T\lambda_2^{1/2}}{2\pi} - \frac{T^3\lambda_2^{3/2}}{(2\pi)^2}\right].$$

Figure 3 gives an example, over the unit cube, with $\lambda_2 = 880$.

Note that once again there are a number of different power series appearing here, although now, as opposed to the two dimensional case, there is no longer a simple correspondence between the powers of $T, \sqrt{\lambda_2}$ and u.

The two positive peaks of the curve are due to A_u being primarily composed of a number of simply connected components for large u and primarily of simple holes for negative u (Recall that in the three dimensional case the Euler characteristic of a set is given by the number of components minus the number of handles plus the number of holes). The negative values of $\mathbb{E}\{\varphi(A_u)\}$ for u near zero are due to the fact that A_u, at those levels, is composed mainly of a number of interconnected, tubular like regions; i.e., of handles.

Although we have discussed only the real valued case in detail above, and for rather simple hitting sets, the extension of the isotropic case to vector valued and to Gaussian related random fields should now be clear. Essentially, the argument is the same, but the computation of the GMFs $\mathscr{M}_j^k$ becomes more complicated. Consider the case of a real valued χ^2 random field, given by

$$f(x) = g_1^2(x) + \cdots + g_k^2(x),$$

where the g_j are independent, identically distributed, isotropic, mean zero, Gaussian random fields with second spectral moment λ_2. Then the expectations $\mathbb{E}\{\mathscr{L}_i(A_u(f,M))\}$ follow immediately from Eq. 24. The $\mathscr{M}_j^k$ are as given in Eq. 22, and the $\mathscr{L}_j$ are precisely as we just calculated for the simple Gaussian case.

An example of what the corresponding expected Euler characteristic curve looks like is given in Fig. 4, for which we have taken $k = 5$ and λ_2 (the second spectral moment of the underlying Gaussian processes) to be 20 (The reason for this particular choice of parameters will be given in Sect. 7).

The χ^2 example exemplifies one of the most important properties of Eq. 24, which is a *separation of parameters*, by which we mean that the parameters relating to the random field, to the parameter set, and to the hitting set do not interact in other than a multiplicative fashion.

Unfortunately, this is not quite true when we leave the isotropic situation. In the general case—with assumptions as above but with isotropy (and the implied stationarity) replaced by an assumption only of constant unit variance—Eq. 24 becomes

$$\mathbb{E}\{\mathscr{L}_i(A_D(f,M))\} = \sum_{j=0}^{N-i}\begin{bmatrix} i+j \\ j \end{bmatrix}(2\pi)^{-j/2}\mathscr{L}_{i+j}(M)\mathscr{M}_j^k(D). \tag{28}$$

However, in this equation the LKCs no longer have the simple meanings that they have had up until now. We shall try, in one short paragraph, to explain how they change, but the details are technical and so we refer the interested reader to Adler and Taylor (2007) for details.

Consider either a real valued Gaussian random field or a single component of a vector valued field. In either case, denote this by f. At a point $x \in M$, take two unit vectors, say V_1 and V_2, with base at x. Let $V_i f$ denote the directional derivative of f in direction V_i. Now define a function g_x on all such vectors by setting

$$g_x(V_1, V_2) = \mathbb{E}\{V_1 f(x) V_2 f(x)\}. \tag{29}$$

Then g_x defines what is known as a Riemannian metric on M, turning it into a Riemannian manifold, and with a corresponding notion of volume. It turns out that tube formulae such as Eq. 17 still hold for such volumes, and the coefficients in the corresponding expansion are the new LKCs. Note that the LKCs (except for the Euler characteristic $\mathscr{L}_0$) now incorporate information about the random field f, specifically information about the covariances of partial derivatives. The fact that the Euler characteristic does not change is a deep result known as the Gauss-

Bonnet theorem, and, as we shall see in the following section, is extremely convenient when looking at excursion probabilities.

Despite the entrance of Riemannian geometry into our story, there are at least two cases in which life is not too bad, and it is possible to compute the LKCs quite simply. For all the other cases we shall describe, in Sect. 6, how to estimate them from data. In fact, this estimation procedure is so simple that it enables one to use formulae like Eq. 28, at least for the important case $i = 0$, without even having to know the definition of the LKCs appearing on the right hand side of the equation.

Now, however, we consider the two simple cases. The first occurs under stationarity, for which we can define a collection of second order spectral moments by

$$\lambda_{ij} = \mathbb{E}\left\{\frac{\partial f(x)}{\partial x_i}\frac{\partial f(x)}{\partial x_j}\right\} = -\frac{\partial^2 C(x)}{\partial x_i \partial x_j}|_{x=0}. \tag{30}$$

Return now to the case Eq. 14 of a N-dimensional rectangle T. Let $(e_1,...,e_N)$ be the usual set of axes in $\mathbb{R}^N$ and for a k-dimensional face J let Λ_J be the matrix of second spectral moments λ_{ij} for which i and j satisfy the requirement that both e_i and e_j intersect J at points other than the origin (If $\dim(J) = 0$, then we define Λ_J to be the constant 1). Then the LKCs of (28) become

$$\sum_{J\in\mathcal{C}_k} |J|[\ \det(\Lambda_J)]^{1/2},$$

something which is quite simple to evaluate in practice. Unfortunately, however, the computation is a little more complicated when the parameter space is not a rectangle. However, the GMFs, as always, are independent of the properties of both f and M.

Another helpful situation arises when one is interested in high level excursion sets of real valued random fields. Looking back at results such as Eqs. 25 and 26, we note that they have the form of a power series expansion in u, so that, if u is large, the first term in the expansion should dominate all the others. Putting this together with Eq. 28 now tells us that, at least for real valued Gaussian fields of unit variance, we should have

$$\mathbb{E}\{\mathscr{L}_i(A_u(f,M))\} \approx \begin{bmatrix} N-i+j \\ N-i \end{bmatrix} (2\pi)^{-(N-i)/2}\mathscr{L}_N(M)\mathscr{M}^1_{N-i}([u,\infty)), \tag{31}$$

where we interpret $a(u) \approx b(u)$ to mean that $\lim_{u\to\infty} a(u)/b(u) = 1$. Thus, all we need, at least for asymptotics, is to know how to compute $\mathscr{L}_N(M)$. However, although the other LKCs are hard to compute in general, there is a simple representation for this one. In fact, if we let $\Lambda(x)$ be the $N \times N$ matrix with elements

$$\lambda_{ij}(x) = \mathbb{E}\left\{\frac{\partial f(x)}{\partial x_i}\frac{\partial f(x)}{\partial x_j}\right\} = \frac{\partial^2 C(u,v)}{\partial u_i \partial v_j}|_{u=v=x},$$

then one can show that

$$\mathscr{L}_N(M) = \int_M [\det(\Lambda(x))]^{1/2} dx,$$

so that Eq. 31 becomes a very general, and useful, result.

5 Excursion probabilities

We now want to see how the formulae of the previous section can be applied to evaluating excursion probabilities, something which we have already mentioned is one of their main applications.

Actually, most of what we had to say was already said in Eq. 8, where we pointed out that, for many random fields,

$$\left|\mathbb{P}\left\{\sup_{x\in M} f(x) \geq u\right\} - \mathbb{E}\{\varphi(A_u(f,M))\}\right| \leq \text{error}(u), \tag{32}$$

where $\varphi = \mathscr{L}_0$ is the Euler characteristic. The previous section was devoted to ways of computing the expectation here for a wide class of random fields, and so all we have left to do is to say something about the error function and offer a heuristic explanation as to why such a result might hold. We start with the heuristics.

Note firstly that

$$\begin{aligned}\sup_{x\in M} f(x) \geq u &\Leftrightarrow A_u(f,M) \neq \emptyset \\ &\Leftrightarrow \#\{\text{connected components of } A_u\} \geq 1.\end{aligned} \tag{33}$$

Consider the Euler characteristic of A_u. Recall, that we already saw, in two and three dimensions, that $\varphi(A_u)$ is equal to the number of connected components of A_u, plus or minus other factors, such as the numbers of holes and handles. In fact, it is true in any dimension that the Euler characteristic of a set is given by the number of its connected components, plus an alternating sum of what might be called 'higher order' functionals, which, as the dimension grows, describe more and more complicated geometric structures. However, it can be shown that, as u becomes large, the structure of an excursion set A_u becomes rather simple, containing only simply connected components, without holes, handles, etc. In fact, if u is large enough, A_u will contain only one (small) simply connected component, and so will have Euler characteristic of one. In other words, for large u we can replace Eq. 33 by

$$\sup_{x\in M} f(x) \geq u \text{“} \Leftrightarrow \text{”} \varphi(A_u(f,M)) = 1, \tag{34}$$

where we interpret "⇔" to mean "asymptotically if and only if", and now $\varphi(A_u)$ can, with high probability, only take the values 0 or 1.

Taking probabilities on both sides of this equation, and recalling that if X is a 0–1 random variable then $\mathbb{P}\{X = 1\} = \mathbb{E}\{X\}$, gives that

$$\mathbb{P}\left\{\sup_{x\in M} f(x) \geq u\right\} \approx \mathbb{E}\{\varphi(A_u(f, M))\}$$

for large u, or, equivalently, that the error term in Eq. 32 should be small for large u.

This heuristic argument works well for most random fields, and you can read more about it and other heuristic arguments for random field extrema in Adler (2000). However, a rigorous result has been proven only for constant variance Gaussian fields (Unfortunately, though, the above heuristics have little to do with the proof). We shall give just one example of it.

Suppose f is Gaussian, isotropic, and with constant unit variance. Then the error function in Eq. 32) is of the form

$$error(u) = e^{-u^2(1+\sigma_c^{-2})/2}, \tag{35}$$

where

$$\sigma_c^2 = \mathrm{Var}\left(\frac{\partial^2 f(x)}{\partial x_1^2}\,|f(x)\right) = \left.\frac{\partial^4 C(x)}{\partial x_1^4}\right|_{x=0} - 1.$$

This particular result is an example of a very general phenomenon for unit variance Gaussian random fields, of basically the same form, but in which the exponent on the right hand side of Eq. 35 is replaced by $-u^2(1+\eta)/2$, for a $\eta > 0$ that is not always easy to identify. However, even when η cannot be identified, this is a remarkable result.

To see why, consider, for example, the unit variance stationary case with parameter space a N-dimensional rectangle (although any set would do) in which case we now know from Eq. 25 that we can rewrite the above as

$$\mathbb{P}\left\{\sup_{x\in T} f(x) \geq u\right\} = C_0\Psi(u) + u^N e^{-u^2/2}\sum_{j=1}^{N} C_j u^{-j},$$

plus an error of order $o(e^{-u^2(1+\eta)/2})$ where the C_j are constants depending on the LKCs of T and the second order spectral moments of f.

Why is this remarkable? If we think of the right hand side above as a terminated version of an infinite power series expansion, then it would be natural to expect that the error term would be of the order of the 'next' term in the expansion, and so of order $u^{-1}e^{-u^2/2}$. However, we have seen that this is *not* the case, and that the error is actually *exponentially smaller* than this. Presumably this theoretical result is behind the fact, well established in practice, that the approximation of excursion probabilities by expected Euler characteristics of excursion sets works remarkably well. In general, numerical and other evidence suggests that at levels at which the true excursion probability is less than 0.10, the error in the Euler characteristic approximation is no more than 5–10% of the true probability.

6 On measuring, calculating and estimating LKCs and Minkowski functionals

We now know that there are three main components that appear in the new theory that has been discussed above. The first are the Lipschitz–Killing curvatures of excursion sets, which are random variables and can be computed from data. We have not gone into detail as to how to do this, and it is not always easy. However, it is easy for the Euler characteristic φ ($\equiv \mathscr{L}$) and we shall see now how to do this.

Suppose, which is almost always that case, that we see our data not on a continuum, but rather on a rectangular lattice L_δ, of edge size δ, so that rather than seeing $\{f(x)\}_{x\in M}$ we see only $\{f(x)\}_{x\in L_\delta\cap M}$. Then we also cannot really see the true excursion sets $A_D(f, M)$, but rather rectangular approximations to them of the form

$$A_D^\delta(f, M) \triangleq \{x \in M \cap L_\delta : f(x) \in D\}.$$

If we join each point in $A_D^\delta(f, M)$ to those distance δ from it (there are a maximum of $2N$ such points) we construct an N-dimensional object, made up of N-dimensional squares, $(N-1)$-dimensional faces, and so on, down to 0-dimensional vertices (which are actually the points of $A_D^\delta(f,M)$. Denoting the number of such k-dimensional faces by N_k^δ, $k = N, N-1, \ldots, 0$, it turns out that

$$\varphi\left(A_D^\delta(f, M)\right) = \sum_{k=0}^{N}(-1)^{N-k}N_k^\delta.$$

Not surprisingly, under mild regularity conditions, it also turns out that $\varphi(A_D^\delta(f, M)) \to \varphi(A_D(f, M))$ as $\delta \to 0$. This is how one computes, or approximates, the Euler characteristic of excursion sets in practice. The remaining LKCs, particularly in the non-isotropic case, are somewhat more complex. In the isotropic case, however, one can use the definitions given in Sect. 1 to compute them.

The remaining two components of the theory are the LKCs of the parameter space, measured with respect to the Riemannian metric Eq. 29 as described at the end of Sect. 4, and the Gaussian Minkowski functionals of the hitting set D, described in Sect. 3.2. Computing either of these can involve considerable knowledge of differential geometry. The good news is that there are ways around this.

First of all, the Gaussian Minkowski functionals depend only on the hitting set D, and there are only a limited

number of hitting sets that are in common usage. Thus one can generally find what one needs already done somewhere. An exhaustive list of known examples can be found in Chapter 4 of Adler et al. (2008), along with techniques for developing any additional cases that might arise.

The case of the Lipschitz–Killing curvatures is somewhat different. If the random field is stationary, and the parameter space T is a rectangle, then we are covered by Eq. 25. A few other examples are given in Chapter 4 of Adler et al. (2008). All of these examples require estimation of the parameters λ_{ij}, or, at least, of determinants of matrices made up by them. A more direct way of estimating LKCs has been proposed in Taylor and Worsley (2007). The strength of this approach is that there is actually no need for the practitioner to have any serious depth of understanding of the geometrical meaning of the $\mathscr{L}_j$. Rather, they can be treated almost as nuisance parameters to be estimated from data via a rather straightforward algorithm.

7 An application: model identification

To conclude this review, I want to describe one rather interesting and powerful application of the new formulae of Sect. 4, and rather than describing the general problem, I shall do it via a particular example.

7.1 The example

Suppose that we are given a real valued random field f on $\mathbb{R}^3$, which we observe on the unit cube $[0, 1]^3$. Our problem is that we are not certain what the distribution of the field is. For simplicity, we shall assume that the choices are Gaussian and χ^2, where both are assumed to be stationary and isotropic.

Then one way to choose between the two is to calculate, from the data, an empirical Euler characteristic curve, given by

$$\widehat{\varphi}(u) = \varphi\Big(A_u\Big(f, [0, 1]^3\Big)\Big).$$

If the data is Gaussian, then the graph of $\widehat{\varphi}$ should be of the general form of Fig. 3, where the actual height and width of the graph will depend on the variance and second spectral moment of f.

If the data is χ^2, then the graph of $\widehat{\varphi}$ will have a different form, closer to that of Fig. 4, for example, which recall is the expected Euler characteristic graph for a χ^2 random field with five degrees of freedom, and with variance and spectral moment chosen to match those of the Gaussian random field of Fig. 3.

The graphs of Figs. 3, 4 are quite different, and thus one should be able to see the same qualitative differences in graphs of $\widehat{\varphi}$ in the two cases. This indeed is generally the case, and so we see that the expected Euler characteristic (EEC) curve provides a diagnostic tool for differentiating between models.

One caveat should be made here, however, before we continue. The significant differences between the Gaussian and χ^2 graphs would, of course, weaken considerably were we to take the latter with a far higher degree of freedom, say 30 or more. This, however, is far from surprising, since in such a regime the central limit theorem tells us that the not only are the differences between Gaussian and χ^2 random variables rather small, the same is true for Gaussian and χ^2 random fields.

Returning to the χ^2_5 example, suppose we were to adopt what we claimed at the beginning of Sect. 2 was the erroneous approach of 'transforming to normality' via a transformation as in Eq. 10. In fact, let us go a little further now, assuming that we know our data is sampled from a χ^2_5 random field f. Let F_5 to be the (known) distribution function of a χ^2_5 random variable, and create a new random field by setting

$$\widetilde{f}(x) \triangleq \Phi^{-1}\Big(\widehat{F}_5(f(x))\Big). \tag{36}$$

In other words, we perform the 'Gaussianisation' of Eq. 10 which, because of our assumptions, we can do without data.

Then the univariate distributions of $\widetilde{f}$ are standard Gaussian, and so it would be impossible, looking merely at these, to distinguish between $\widetilde{f}$ and a Gaussian random field.

Figure 5 gives the EEC curve for $\widetilde{f}$, with parameters chosen so that $\widetilde{f}$ has the same variance and second spectral moment of the Gaussian field of Fig. 3. This curve is quite different to the corresponding Gaussian one, both in terms of its general shape and the rate at which it reaches its asymptotes. Thus, while univariate marginal distributions of the Gaussianised random field offer no indication of inherent non-Gaussianity, the EEC curve does.

This fact has often been used, particularly in the astrophysical literature, not only as a way to distinguish between two models, but to search among models for the correct one (e.g., Gott et al. 2008; Torres 1994; Vogeley et al. 1994).

7.2 What we can learn from the example

After reading the above 'application', a referee asked a number of pertinent questions that would probably be of interest to any reader who works with spatially correlated data. Thus, it seems appropriate to close with a

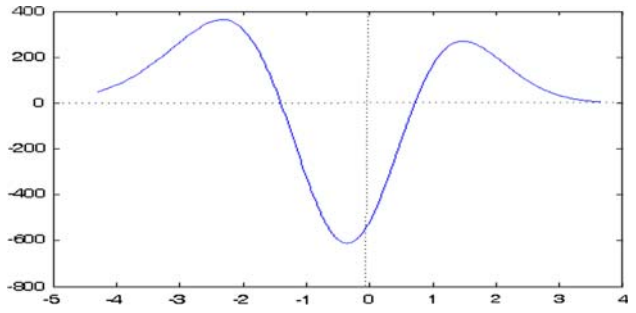

Fig. 5 EEC for a 'normalised' χ^2_5 field in three dimensions. Axes as for Fig. 2

paraphrased, concatenated version of his questions, along with the best answers that I can provide to them.

1. What are the advantages of the above inferential procedure to one based on standard distributional tests?

In a certain sense, a test for normality based on Euler characteristics or, indeed, any of the LKCs, is a distributional test. The basic expectation formulae depend only the variance of f and the variances and covariances of its first order partial derivatives. If, for example, the random field is assumed to be isotropic, as in the example above, then the test described there is really only a test of the joint normality or otherwise of the distribution of the Gaussianised field and its first order derivatives.

What makes it somewhat different from standard tests is how it tests for non-normality, by placing emphasis on the geometric structures that are generated by the random field. Whether or not this is a good thing will depend very much on where one's interest in the data is concentrated. If it is concentrated on geometry, we believe that EEC techniques provide a more powerful approach to testing for non-normality that standard tests. This belief certainly seems to be shared, and put into practice, by large parts of the brain mapping and astrophysics communities for whom the analysis of shape has become an important part of their inferential procedures.

Note, however, that if assumptions of neither isotropy nor stationarity are made, then straightforward distributional tests are impossible to apply, since there are an uncountable infinity of different distributions to be considered. In these cases a summary statistic of some sort must be employed. The geometric approach suggested above seems to be particularly suited to these scenarios.

2. How does the theory presented here convey information about the long range properties of the random field?

I have left this point to last, since it is an extremely important one. The fact is that none of the formulae of this paper are affected by, or descriptive of, long range properties of the random field. Of course, all the formulae for the means $\mathbb{E}\{\mathscr{L}_j(A_u(M,f))\}$ depend on the size and shape of M, and, for example, they grow, in an additive fashion, as M grows. However, this growth does not take into account the rate of decay, for example, of the covariance function of f.

At first sight, this is rather surprising, but, in fact, it is not something particularly new. Consider, for example, the excursion probability (Eq. 7). From the asymptotic equivalence (Eq. 8) [or any other technique for computing (Eq. 7)] and what we now know about the mean Euler characteristic of excursions sets, it follows that, at least at high levels, the excursion probability also does not depend on the long range behaviour of the covariance function. There are many other examples in the theory of Gaussian random fields for which this kind of behaviour occurs, and the geometric theory of this paper is just one more.

Is this a strength or a weakness? The answer, of course, depends on what one is interested in. If long range behaviour of covariance functions is important for some other reason, then the theory of this paper has little to offer to help in understanding or estimating them. On the other hand, it is nice to know that there are many interesting aspects of random fields that do not depend on the decay rates of covariance functions. Indeed, it may well be rather comforting, since such decay rates are one of the hardest phenomena to accurately measure, in practice, from real data.

Acknowledgments Throughout this review I have given few references and very little detail as to the long history of this subject. Two names stand out. One is Jonathan Taylor, whose fundamental contributions to the theory have already been mentioned, and the second is Keith Worsley, who has made more contributions to the application of the Euler characteristic approach to random fields than anyone else I know. However, there are many other names that deserved to be mentioned, and were not. I acknowledge that I owe an apology to both these authors and the reader for not mentioning them by name. My justification for this oversight is twofold. Firstly, in a brief review article like this one it would be impossible to even attempt to do credit to the history of the subject. Secondly, there are sources that you can turn to for the missing details, including Adler (2000), Adler and Taylor (2007) and Adler et al. (2008). Finally, a debt of gratitude is due to Einat Engelhart, a Technion MSc student, who helped with some of the numerical work.

References

Adler RJ (2000) On excursion sets, tube formulae, and maxima of random fields. Ann Appl Probab 10:1–74

Adler RJ, Taylor JE (2007) Random fields and geometry. Springer, Boston

Adler RJ, Muller P, Rozovskii B (eds) (1996) Stochastic modelling in physical oceanography (1996). Birkhaüser, Boston

Adler RJ, Taylor JE, Worsley KJ (2008) Applications of random fields and geometry: foundations and case studies. Some chapters avaliable at http://ie.technion.ac.il/Adler.phtml (in preparation)

Christakos G (1991) Random field models in earth sciences. Academic Press, San Diego
Christakos G (2000) Modern spatiotemporal geostatistics. Oxford University Press, Oxford
Daley R (1991) Atmospheric data analysis. Cambridge University Press, Cambridge
Evans AC, Marret S, Neelin P, Collins L, Worsley KJ, Dai W, Milot S, Meyer E, Bub D (1992) Anatomical mapping of functional activation in stereotactic coordinate space. Neuroimage 1:43–53
Gott JR, Hambrick CD, Vogeley MS, Kim J, Park C, Choi Y-Y, Cen R, Ostriker JP, Nagamine K (2008) Genus topology of structure in the Sloan digital sky survey: model testing. Astrophys J 675:16–28
Matérn B (1960) Spatial variation: Stochastic models and their application to some problems in forest surveys and other sampling investigations, Meddelanden Fran Statens Skogsforskningsinstitut, Band 49, Nr. 5, Stockholm
Olea RA (1999) Geostatistics for engineers and earth scientists. Kluwer, Boston
Rubin Y (2002) Applied stochastic hydrogeology. Oxford University Press, Oxford
Sampson PD, Guttorp P (1992) Nonparametric estimation of nonstationary spatial covariance structure. J Am Stat Assoc 87:108–119
Shafie K, Sigal B, Siegmund D, Worsley KJ (2003) Rotation space random fields with an application to fMRI data. Ann Stat 31:1732–1771
Smoot G, Davidson K (1993) Wrinkles in time. William Morrow, NY
Taylor JE (2001) Euler characteristics for Gaussian fields on manifolds, Ph.D. dissertation, McGill
Taylor JE (2006) A Gaussian kinematic formula. Ann Probab 34:122–158
Taylor JE, Worsley KJ (2007) Detecting sparse signals in random fields, with an application to brain mapping. J Am Stat Assoc 479:913–928
Taylor JE, Takemura A, Adler RJ (2005) Validity of the expected Euler characteristic heuristic. Ann Probab 33:1362–1396
Torres S (1994) Topological analysis of COBE-DMR cosmic microwave background maps. Astrophys J 423:L9–L12
Vogeley MS, Park C, Geller MJ, Huchin JP, Gott JR (1994) Topological analysis of the CfA redshift survey. Astrophys J 420:525–544