Set-level threshold-free tests on the intrinsic volumes of SPMs

Gareth R. Barnes *, Gerard R. Ridgway, Guillaume Flandin, Mark Woolrich, Karl Friston
The Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

A B S T R A C T

Conventionally, set-level inference on statistical parametric maps (SPMs) is based on the topological features of an excursion set above some threshold—for example, the number of clusters or Euler characteristic. The expected Euler characteristic—under the null hypothesis—can be predicted from an intrinsic measure or volume of the SPM, such as the resel counts or the Lipschitz–Killing curvatures (LKC). We propose a new approach that performs a null hypothesis omnibus test on an SPM, by testing whether its intrinsic volume (described by LKC coefficients) is different from the volume of the underlying residual fields: intuitively, whether the number of peaks in the statistical field (testing for signal) and the residual fields (noise) are consistent or not. Crucially, this new test requires no arbitrary feature-defining threshold but is nevertheless sensitive to distributed or spatially extended patterns. We show the similarities between our approach and conventional topological inference—in terms of false positive rate control and sensitivity to treatment effects—in two and three dimensional simulations. The test consistently improves on classical approaches for moderate (>20) degrees of freedom. We also demonstrate the application to real data and illustrate the comparison of the expected and observed Euler characteristics over the complete threshold range.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Random field theory is used in neuroimaging to account for the intrinsic smoothness of images, when making inferences on the basis of statistical parametric maps (Kilner and Friston, 2010; Worsley et al., 1996, 2004). Statistical parametric maps (SPMs) are realisations of random fields, whose topological characteristics—under the null hypothesis—can be predicted using random field theory. This provides a powerful framework for topological inference on data that are collected over one or more dimensions, such as images or time–frequency responses. In particular, random field theory allows one to make predictions about the Euler characteristic (EC) of the excursion set produced by thresholding a random field (intuitively, the number of blobs minus the number of holes of the excursion set). This prediction is useful, because—at high threshold—holes disappear and the expected EC approximates the number of maxima one would expect under the null hypothesis. This use of random field theory requires a threshold to create excursion sets that can then be assessed at various levels of inference (Friston et al., 1996); for example, on the basis of the extent or number of maxima above some threshold. In this work we introduce an approach that exploits random field theory without the need for a particular threshold.

The Euler characteristic of an SPM—at any threshold—can be predicted through the Gaussian Kinematic Formula (Taylor and Worsley, 2007). This formula expresses the expected EC as the sum over products of an EC density function and a measure of intrinsic volume. The EC density function depends only on the type of statistic that constitutes the SPM, and the intrinsic volume is a measure (curvature or count) of the multidimensional extent of the random field. In this work we use the Lipschitz–Killing curvatures (LKC) as our measure of intrinsic volume (Taylor and Worsley, 2007). However, it should be noted that LKC is interchangeable with the more traditionally used resel count through a scale factor.

One can gain some intuition about the form of the Gaussian Kinematic Formula by considering a one-dimensional statistical field or process (Fig. 1). For a given (high) threshold, the number of supra-threshold segments (the Euler characteristic) depends on the underlying statistic (determined by the density function) and a measure of the smoothness and length of the process (determined by the LKC). Smoother or shorter processes will have a smaller LKC and a smaller EC for a given threshold. Fig. 1 shows two t-statistical processes with the same intrinsic volume or LKCs: one is five times longer than the other but is also five times smoother. The intrinsic volume of a random field is the volume it would occupy when statistically flattened—so that it has unit smoothness everywhere. When flattened, the expected Euler characteristic is simply the EC density—per unit of intrinsic volume—times the intrinsic volume. In our one-dimensional example, both processes have the same intrinsic volume (the same underlying LKCs) and both are t-fields. Therefore, they have the same expected EC. In this case, the expected EC at this threshold (υ = 0.8) is...
an excursion set (like the number of maxima above a threshold),

In other words, instead of using the LKC to assess the signiﬁcance of the excursions between the observed and null distributions, one can simply solve a linear regression problem using the column design matrix \( X \) multiplied by a vector of regression coefﬁcients \( \beta \) plus a vector of normal errors \( e \).

\[
y = X\beta + e.
\]

Ignoring relationships between elements of \( y \) over trials for clarity (see Discussion), the least squares estimator of the regression coefﬁcients is

\[
\hat{\beta} = \left( X^T X \right)^{-1} X^T y.
\]

Leaving a vector of residuals \( \tilde{e} \) at each voxel

\[
\tilde{e} = y - X\hat{\beta}
\]

where the standardised residuals \( r \) are given by

\[
r = \frac{\sqrt{N-1} \tilde{e}}{\sqrt{\tilde{e}^T \tilde{e}}}.
\]

In the next section we will use the result that adjacent voxels will have similar standardised residuals.

Random ﬁeld theory

This section describes random ﬁeld theory as typically used to quantify the number of maxima above a threshold one would expect by chance. This material is reviewed comprehensively in a number of other texts (Kiebel et al., 1999; Kilner and Friston, 2010; Worsley et al., 1996, 1999). For a \( D \) dimensional space \( S \), let \( A_u \) be the excursion set at threshold \( u \).

\[
A_u = \{ s \in S : A(s) \geq u \}.
\]

For example, \( S \) might constitute a triangular mesh spanning the cortical surface with values for some statistical test at each vertex. Alternatively, it could be a regular lattice approximation to a three dimensional search space. \( A_u \) would then constitute all the vertices above threshold \( u \).
The expected value of the Euler characteristic of this excursion set (intuitively, the number of blobs minus the number of holes) is given by the Gaussian Kinematic Formula (Taylor and Worsley, 2007)

\[
E[\phi(A_u)] = \sum_{d=0}^{D} L_d(S) \rho_d(u)
\]

(6)

where \( L_d \) are the Lipschitz–Killing curvatures (LKC) and depend on the smoothness and shape of \( S \). Here, \( \phi(A_u) \) is the Euler characteristic of the excursion set \( A_u \) and \( \rho_d \) are EC density functions, determined purely by the statistic in question (usually a \( t \) or \( F \) statistic). \( D \) is the dimensionality of the image (e.g., 2 for a surface), where \( D = 0 \) corresponds to a single point. \( L_0(S) \) is simply given by the Euler characteristic of the space under test (e.g., \( L_0(S) = 2 \) for two separate hemispheric volumes; \( L_0(S) = 4 \) for two separate surfaces with spherical topology).

Having estimated the LKC for some space, one can infer (for example) whether two clusters (EC = 2) at threshold \( u = 3 \) would be expected by chance. Fig. 2 shows the correspondence between the empirical and predicted ECs from Eq. (6) using simulated data (see below).

If the space is not homogeneously smooth, it can be readily transformed by expressing distances (between vertices) in terms of the correlations between residuals (at each vertex) over observations (Worsley et al., 1999). This new space becomes uniformly smooth (statistically flat), where vertices with similar residuals are closer together. For a given dimension \( d \in \{0, \ldots, D\} \), each component (say triangle on a two dimensional mesh) will have \( d + 1 \) vertices. Now, indexing by component \( (j) \) let the \( j \)th component (say a triangle) have \( d + 1 \) vertices, each with a vector of standardised residuals \( \{r_{j1}, r_{j2}, \ldots, r_{jd}\} \) (Eq. (4)). As we are only interested in the intrinsic volume occupied by this component (triangle), we can take one vertex \( \{r_{j1}\} \) as a reference to define a \( d \) dimensional solid. Each component (triangle) is then defined by

\[
\Delta R_j = [r_{j1} - r_{j1}, \ldots, r_{jd} - r_{jd}]
\]

(7)

where \( \Delta R_j \in \mathbb{R}^{N \times d} \). The volume (or area for a surface) of this \( d \) dimensional component is then simply given by

\[
\text{Volume}_j = \frac{1}{d!} |\Delta R_j \Delta R_j^T|^{1/2}.
\]

(8)

For a triangle, this is half base multiplied by height. In the three dimensional case 8 adjacent voxel corners can be broken down into 5 tetrahedral components (Worsley et al., 1999). The Lipschitz–Killing curvature (Taylor and Worsley, 2007) of the whole space \( L_d(S) \) is then simply the sum of the volumes occupied by the \( j \) individual components (triangles or tetrahedra).

\[
L_d(S) = \frac{1}{d!} \sum_{j=1}^{N} |\Delta R_j \Delta R_j^T|^{1/2}.
\]

(9)

The equivalent resel counts are simply \((4\log(2))^{-d/2}L_d(S)\). Finally, at high thresholds, when there are no holes in the image, the LKC given in Eq. (9) can be used in Eq. (6) to estimate the probability (or the expected value of the Euler characteristic) of a global maximum \( M \) of value greater than or equal to \( u \)

\[
P(M \geq u) \approx E[\phi(A_u)].
\]

(10)

Extending the above to calculate the probability that the number of clusters \( (c_{\text{max}}) \) should be greater than or equal to \( C \) clusters observed we can use the Poisson clumping heuristic (Friston et al., 1996)

\[
P(c_{\text{max}} \geq C) = 1 - \sum_{c=0}^{C-1} \frac{e^{-\lambda \Lambda}}{c!}
\]

(11)

where the term on the right is the cumulative Poisson distribution with mean \( \lambda = E[\phi(A_u)] \), the number of clusters expected at threshold \( u \).

When estimating the probability of a single cluster above threshold \( (C=1) \), the approximation is consistent with Eq. (10) and \( P(c_{\text{max}} \geq 1) = E[\phi(A_u)] \).

There are two key points that arise in this use of the LKC. We have to specify a threshold to make use of the Gaussian Kinematic prediction (Eq. (6)). Second, in order to estimate the LKC, we need to know the local covariance structure of differences in residuals over observations and, implicitly, the topology or connectivity that defines these differences (Eq. (8)) (see Discussion).

**Estimating LKC through regression**

In this section, we address the problem of estimating the intrinsic volume of a single realisation of a random field. This will allow us to

---

**Fig. 2.** A. The average Euler characteristic as a function of threshold, for the standardised residual fields (black) with error bars showing the standard deviation over realisations. The green circles show the estimate of the Euler characteristic of the underlying random field—here predicted from the Gaussian Kinematic Formula (Eq. (6)) using the standard smoothness (resel) estimator (blue dotted line) and the regression estimator (green circles). This confirms that the curves in Fig. 2A are almost identical at high threshold. The dashed lines are binomial 95% confidence intervals around the anticipated false positive rate (black, dotted).
decide whether the LKC of the SPM are consistent with those of the residual fields (in a group study there will be one per subject) that we know conform to the null hypothesis. By definition, the standardised residual at each vertex has a mean of zero and a variance of unity (Eq. (4)).

Each residual field corresponds to a Z-field, which has a predictable Euler characteristic over a range of \(H\) thresholds, \(u_h\) (we used \(H = 81\) thresholds from \(-4\) to \(+4\) with steps of \(0.1\)). Following Bartz et al. (in press), we can then estimate the LKC using the following single general linear model:

\[
\varphi_{n,h} = \sum_{d} h_d(S) \rho_d(u_h) + \varepsilon_{n,h}
\]  
(12)

where \(\varphi_{n,h}\) is the measured EC at threshold \(u_h\) of the \(n\)-th residual field (in a group study, \(n\) would index subjects). The unknown LKC of this residual field for dimension \(d\) is \(l_{n,d}\) and the superscript \(z\) in \(\rho_d(u_h)\) signifies that this is the EC density for a Gaussian (rather than \(t\) or \(F\)) field at threshold \(u_h\).

At dimension zero, the LKC is simply the Euler Characteristic of the space under test and so this coefficient need not be estimated, giving

\[
\varphi_{n,h} - l_{n,0}(S) \rho_0(u_h) = \sum_{d} l_{n,d}(S) \rho_d(u_h) + \varepsilon_{n,h}
\]  
(13)

or

\[
\bar{\varphi}_{n,h} = \sum_{d} h_d(S) \rho_d(u_h) + \varepsilon_{n,h}
\]  
(14)

where

\[
\varphi_{n,h} = \varphi_{n,h} - l_{n,0}(S) \rho_0(u_h).
\]  
(15)

In order to simplify this notation, we can define the (data independent) \(D\) dimensional EC density (for a Gaussian) in Eq. (14) for all \(H\) threshold values as the matrix \(\Gamma^2 \in \mathbb{R}^{H \times D}\) and \(l_{n} \in \mathbb{R}^{D \times 1}\) as a vector of the (unknown) \(D\) Lipschitz–Killing curvatures (Eq. (9)) for residual field \(n\).

We can re-write Eq. (14) in matrix form:

\[
\varphi_n = \Gamma^2 l_n + \varepsilon_n
\]  
(16)

where \(\varphi_n \in \mathbb{R}^{H \times 1}\) is the matrix form of Eq. (15) measured over the \(H\) threshold values for residual field (subject) \(n\) and \(\varepsilon_n \in \mathbb{R}^{H \times 1}\) is a vector of errors.

Eq. (16) has a familiar form but there are some concerns about the assumptions required to solve this GLM. For example, the error terms are heteroscedastic (less variance at high thresholds) and the LKCs themselves are correlated (Bartz et al., in press). These authors examined a number of different covariance estimators (smoothed diagonal to account for heteroscedasticity, smoothed covariance to account for correlation, etc.) and found the simpler covariance models (ordinary least squares, smoothed diagonal) to be more robust (in terms of bias and variance).

Defaulting to the simplest regression model, here we estimate the LKC using ordinary least squares

\[
\hat{l}_n = [\Gamma^2]^+ \varphi_n
\]  
(17)

where \([\Gamma^2]^+ \in \mathbb{R}^{D \times H}\) is the pseudo-inverse of the Gaussian \((Z)\) EC density and the estimate of the \(D\) unknown LKC coefficients of the \(n\)th \((n = 1\) to \(N\)) residual field is the vector \(l_n \in \mathbb{R}^{D \times 1}\).

We can make a similar estimate of the LKCs underlying the test statistic image (which in the null case should be identical to those from the residual images)

\[
\hat{l}_{test} = [\Gamma^2] \varphi_{test}
\]  
(18)

where \(\varphi_{test} \in \mathbb{R}^{H \times 1}\) has the same form as \(\varphi_n\) but is instead based on the measured EC of the SPM at over the threshold range. The LKC estimate for the test SPM is \(\hat{l}_{test}\). Note again that the superscript in \(\Gamma^2\) identifies this as the matrix of EC densities for the Subject \(t\)-statistic.

We now can test whether the LKC estimated over the \(N\) residual fields are significantly different from those estimated from the SPM. Note that the use of the appropriate density function in Eq. (12) factors out any dependence on the statistic in question (\(t\) or \(F\)). We test for these differences using a standard multivariate general linear model:

\[
\begin{bmatrix}
\tilde{l}_{test} \\
\tilde{l}_1 \\
\vdots \\
\tilde{l}_N
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix} \beta_M + E
\]  
(19)

This provides regression parameters, \(\beta_M \in \mathbb{R}^{2 \times D}\) in which the first row is the LKC estimates for the test SPM, and the second row is the mean of the LKC estimates for the residual fields averaged over all samples or observations of the residuals. We then test the multivariate hypothesis that the two rows of \(\beta_M \in \mathbb{R}^{2 \times D}\) are the same to provide a classical \(p\)-value. The test hinges on Wilks’ lambda statistic, which is effectively a (marginal) likelihood ratio test comparing the full model to the reduced (null) model without the different estimates for residual test LKCs under Gaussian assumptions about the errors. For large \(N\) the log-likelihood ratio has a scaled Chi-squared distribution (Chatfield and Collins, 1980). In this case, we use Rao’s F approximation to Wilks’ lambda statistic (Anderson, 2003), to test whether the (intrinsic volume of the) SPM was sampled from the null distribution.

The main assumption behind this test is that of multivariate normality. We used Mardia’s test for multivariate normality based on the 3rd and 4th order moments of the distribution to test for this in the real data example (see Discussion); we also verified that the false positive rate (FPR) for this test was well controlled (Fig. 3D).

This concludes the description of our procedure that furnishes a simple test of the null hypothesis that the intrinsic volume of an SPM is the same as the intrinsic volume of its constituent residual fields—fields that contain no treatment effects. In the next section, we turn to numerical simulations to establish the accuracy and sensitivity of this scheme. All of this software is available from the authors on request and will be made part of SPM12.

**Simulations**

We ran simulations using Gaussian white noise data for both two and three dimensional cases over volumes (or observations) of \(100 \times 100\) pixels and (predominantly) \(30 \times 30 \times 30\) voxels respectively. For the two dimensional data, we used tests based on \(N = 100\) observations, and for the three dimensional case we looked at \(N = 100, 50, 20\) and 10 observations per test. We ran tests in batches of 400 random realisations to evaluate false positive rates (e.g., 400 tests with \(N = 100\) random observations per test). To estimate receiver operating characteristic (ROC) curves, synthetic signals or treatment effects were created by adding a number (between 1 and 25) of impulse responses at random vertices (the same vertices across all realisations) (see Discussion); we also verified that the false positive rate (FPR) for this test was well controlled (Fig. 3D).
Results

Fig. 2A shows the mean and standard deviation of the observed Euler characteristics (black) of the standardised residual fields over thresholds. Also shown are the Euler characteristics predicted from the Gaussian Kinematic Formula using the standard LKC estimates based on smoothness (Eq. (9)) and the LKC estimates based on regression (Eq. (17)). Usually, one uses the high threshold region of these curves to evaluate the probability of a maximum occurring by chance. Fig. 2B shows the false positive rate when testing for one or more of these curves to evaluate the probability of a maximum occurring by chance. Fig. 2B shows the false positive rate when testing for one or more clusters above threshold using the standard (local smoothness based) measure (Eq. (9)), alongside the false positive rate based on the residual fields—shown in panels 3C (blue dots) alongside the LKC estimates (red star) of this non-central SPM. This deviation of the trial and test LKCs is the basis of the multivariate test (Eq. (19)). In panel D we show that, over the 400 null realisations, the multivariate test has good control over the false positive rate.

To evaluate the sensitivity and specificity of multivariate tests on the LKC we compared its performance with standard (based on local smoothness, Eq. (9)) set-level tests on the number of peaks above feature inducing thresholds corresponding to uncorrected p-values of 0.01 and 0.001 and a third test for one or more clusters above p ≤ 0.05 FWE (Eq. (10)). We also were interested to find out how sensitive the approach would be for different subject numbers. Fig. 4 shows an ROC curve quantifying the performance of the multivariate test in comparison to the standard approaches. Panels A, B, C show the performance with N = 10, 20 and 100 observations respectively for volumes of size 30 voxels. The performance of the classical and multivariate approaches are similar for moderate N (= 20) with deterioration of the multivariate performance at low N (= 10). At N = 100 the multivariate test outperformed all chosen feature inducing thresholds for this volume. Note that although it is sometimes possible to find a feature-inducing threshold that outperforms the multivariate test (e.g. N = 20, feature threshold p < 0.001); in practice, searching for the best feature-defining threshold would incur a multiple comparison penalty (by inducing another dimension of the search space).

Panel D shows the area under the ROC curve (for false positive rates ≤ 0.1) as a function of both the number of observations (different lines) and image volume (x axis) for the (local smoothness based) classical test for one or more clusters above p ≤ 0.05 FWE (dotted) and the multivariate method (solid). For small volumes or small numbers of trials the local smoothness based methods are more powerful; as the volume (or number of observations) increases, the EC estimates (over realisations) become less variable and the multivariate tests outperform the local smoothness based tests.

Existing parametric RFT relies on normality assumptions, and the new method is no different. In order to investigate dependence on Gaussian assumptions about error terms in more detail, we added an unmodelled (and un-physiological) step function to the simulated data of the same magnitude as the noise. The resulting non-Gaussian (bimodal) residual distribution was relatively benign for the smoothness based methods, resulting in a more conservative false positive rate; but for our method the effect was to give rise to capricious behaviour (see Supplemental Fig. S1A). Due to the step function, two dominant LKC clusters (each characterising half of the observations) emerged and their combination rather poorly described the EC of the final t-statistic field. We were able to address this effect, thanks to one of our reviewers, by calculating LKC estimates based on a rotated set of residuals (by multiplying the residuals from each of the N observations by the singular vectors of a random N × N matrix). This meant that the rotated (orthonormally mixed) residuals were not only more Gaussian (due to the central limit theorem) but also the variance due to this single step event was spread over the sequence of residuals, affording a better estimate of the true variance of the LKC coefficients (see Supplemental Figs. S2, S3). In the case of stationary data, rotation had no effect and could be applied to data where parametric
Panel B of Fig. 5 shows the exceedances at a threshold corresponding to \( p < 0.01 \) uncorrected (green dots), number of clusters above \( p < 0.001 \) uncorrected (red dashed) and our multivariate test (blue diamonds). Panels A–C show the performance of these tests, in a volume of cube side 30, on SPM(data based on 10, 20 and 100 observations respectively. As the number of observations increases the multivariate test comes to outperform all the feature defined tests. Note that for 10 subjects, although the multivariate test is less sensitive than threshold-dependent tests, it has the fundamental advantage that it requires no search over thresholds (and corresponding statistical correction). Panel D shows the area under the ROC curve (false positive rate 0 to 0.1) against the side length of the cubic volume under test (x axis) for different numbers of observations (N = 10, 20, 50, 100 in different colours) for the multivariate test (solid) and for (the classical) one or more clusters above \( p < 0.05 \) FWE corrected test (dotted). The performance in panels A, B and C correspond to the blue, green and cyan curves for a volume of side 30 voxels (FWHM = 4). The larger the number of observations the greater the potential improvement of the multivariate over the classical test, but for small volumes (VoSize = 20) the classical test (based on local smoothness) outperforms the multivariate test (based on the global EC count).

Discussion

In this work we have used the Gaussian Kinematic Formula to estimate the LKC of a single realisation of a random field. This allows us to estimate the LKC of an SPM assuming the null hypothesis and compare this estimate of its intrinsic volume to equivalent estimates of residual fields. In contrast to standard approaches this way of assessing the departure of the SPM from the null hypothesis, does not depend upon any threshold—rather it gives a complete characterisation of whether the Euler characteristic of the SPM, as a function of threshold, is consistent with the LKC estimated from component residual images. We note that other threshold free approaches exist—most notably permutation testing with threshold free cluster-enhancement (Smith and Nichols, 2009), which is however a non-parametric method. However, to our knowledge, this is the first parametric solution to this problem, and considers a complete characterisation of the field across all thresholds. One could consider this test to be the most general topological (set-level) inference; after which more specific post-hoc questions could be posed. The method has the same form irrespective of whether one wishes to make inferences on peak voxel intensity or some more diffuse global change in signal (Worsley et al., 1995). Notice that the procedures described in this paper can be applied to any SPM, including those based on serially correlated data (provided temporal correlations are modelled appropriately). This is because (standard) implementations of SPM use maximum likelihood estimators of effect sizes and therefore whiten the data (and residual fields) to render them approximately independent.

We note that the non-parametric TFCE (Smith and Nichols, 2009) procedure was developed with the same motivation in mind (to reduce the number of user specified parameters). Initial tests, show that the non-parametric TFCE also outperforms the individual feature defined tests with very similar performance to our parametric method (see Supplemental Fig. S4). The main strength of TFCE is that it is relatively
immune to violations of Gaussian assumptions (see below); whereas the main advantage of our multivariate method (besides computational efficiency) is that there are no user-defined parameters at all (no matter what the statistical test, as long as the associated random field has known EC densities). Our multivariate approach will suffer when the intrinsic volumes under test are small (see Fig. 4D) or the residuals are non-Gaussian and therefore there is some work needed to characterise the trade off between the non-parametric methods (such as TFCE) and our parametric multivariate method. Non-Gaussian residuals are generally not a problem in fMRI data due to the nature of image reconstruction and haemodynamic convolution, which render the data Gaussian by the central limit theorem. However, there could be circumstances (see Fig. S1) or other applications where departures from Gaussian behaviour may be more evident. Indeed multivariate methods may be particularly sensitive to violations of Gaussian assumptions and entail the additional assumption of multivariate normality. In the real data example, Mardia’s test was not significant (p = 0.1863, 0.3016; corresponding to tests for multivariate skewness and kurtosis respectively). We also re-ran the real-data analysis using permutation testing (randomly labelling the tests and observations) and found a similar significant effect (p < 0.0205). This is clearly an area for future validation, with different data sets (VBM, MEG etc.), which can be easily verified through non-parametric methods.

One parameter—that could improve computational efficiency—is the step size and spacing of threshold levels over which to evaluate the empirical EC function. In this work, we used a fixed step size of 0.1 (Z or t), increasing linearly over the threshold range. The more threshold levels the better but this comes at some computational cost. This parameter was examined comprehensively by Bartz et al. (in press) who looked at the theoretical variance of the regression estimated LKCs as a function of step-size. They varied the number of threshold levels from 5 to 200 and found that the estimated variance plateaued at around 50. These authors also looked at different distributions of these levels (equal spacing, quantile spacing etc.). Ultimately these authors used 50 threshold levels over a linear range from −3 to 3, which gives a comparable step size of 0.12. The authors used the same procedure to select a robust covariance estimator and ultimately selected a smoothed diagonal estimator, although they found that the ordinary least squares approach gave comparable performance. In this work we opted for the least complex model (the OLS) but the use of a smoothed diagonal estimator would be an interesting avenue for further work.

In this work, we make use of theoretical properties of a Gaussian random field to make a compact prediction of the LKC coefficients under the null hypothesis. An alternative, in the absence of expressions for the EC density of the test in question, might be to make an inference on the same residual and statistical fields (e.g. Z-fields) by comparing curve descriptions, using a method such as functional data analysis (Ramsay and Silverman, 1997) or summary measures of observation-wise estimates of extent and height similar to TFCE (Smith and Nichols, 2009). Indeed, initial tests using a simple univariate difference measure (between EC in test and EC in residuals over threshold) proved quite effective. The advantage of the parametric approach presented here is that the EC density allows us to predict the EC of any field (e.g. F or t fields) using standard topological theory.

As noted above this eschews an arbitrary feature defining threshold and therefore no search over thresholds (and implicit adjustment for multiple comparisons) are required for set level inference. Given that the set level test is significant one might then consider post-hoc tests with more localising power; for example, using standard tests on peaks or thresholded clusters. In this context, a significant effect at the set level provides protection for localised tests. In other words, having established a significant effect using threshold-free inference, one can then report local tests without further correction for multiple post-hoc tests. This is because the false positive rate of post-hoc (localised) tests can never exceed the nominal false positive rate, provided one does not proceed to post-hoc testing in the absence of a significant test at the set level. Clearly, false positive rates for local tests are only controlled in a (technically) weak sense. On this note, some localising power is possible with the algorithm if a specific anatomical region is specified a priori. From Fig. 4D it is clear that the main advantages of using the multivariate framework become apparent at intrinsic volumes of around 125 resels (20 voxel cube side at 4 mm smoothing).
The ability to estimate the LKC coefficients of a single statistical field allows one to test for signal-induced changes in the apparent intrinsic volume (as estimated through regression). For example, individuals with more extensive horizontal connections in V1 (Schwarzkopf et al., 2012) might give rise to a quantitatively distinct noise field. The LKC coefficients may be a principled and compact parameterisation of the spatial correlations in neuronal fluctuations that could be used to test for such effects.

As pointed out by Bartz and colleagues, the regression method provides a computationally and conceptually simple LKC estimator. It avoids the rather complicated calculation of neighbourhoods (Worsley et al., 1996) and, importantly, is easily generalised to any number of dimensions. As no geometric knowledge of the field is necessary (but simply its topology), exactly the same methodology can be applied directly to tests on two dimensional cortical surfaces or high dimensional connectivity images (by simply changing D in Eq. (6)). For example, in MEG, source orientation as well as Euclidean distance determines the covariance between voxels (Barnes et al., 2011) and the true topology is not necessarily that of nearest neighbours. The use of persistent homology (Adams and Carlsson, 2009) allows one to estimate the topology—and Euler characteristic—of sets of arbitrary data points. This means that one can exploit the simplicity of the regression approach to estimate the LKC of fields with unknown dimension or topology. We will explore this important application in future work on topological inference in MEG data that can show a complicated correlation structure and statistical topology.

Acknowledgments

We would like to thank Rik Henson and Dan Wakeman for providing us with the example dataset; Christian Gaser and Tom Nichols for help with TFCE; and our anonymous referees for their helpful suggestions. This work was supported by the Wellcome Trust [grant number 091593/Z/10/Z]; and the Medical Research Council [grant number MR/J014257/1]. The WTCN is supported by core funding from the Wellcome Trust [grant number 091593/Z/10/Z].

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.11.046.

References

Wakeman, D.G., Henson, R., 2010. Effective connectivity in face processing, as inferred from MEG, EEG and fMRI data. Biomag 2010—17th International Conference on Biomagnetism.