ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 23 (2004) S189–S195

Unified univariate and multivariate random field theory

Keith J. Worsley,[a,b,*] Jonathan E. Taylor,[c] Francesco Tomaiuolo,[d] and Jason Lerch[b]

[a]*Department of Mathematics and Statistics, McGill University, Montreal, Canada H3A 2K6*
[b]*Montreal Neurological Institute, McGill University, Montreal, Canada*
[c]*Department of Statistics, Stanford University, Roma, Italy*
[d]*IRCCS Fondazione 'Santa Lucia', Roma, Italy*

Available online 12 September 2004

We report new random field theory *P* values for peaks of canonical correlation SPMs for detecting multiple contrasts in a linear model for multivariate image data. This completes results for all types of univariate and multivariate image data analysis. All other known univariate and multivariate random field theory results are now special cases, so these new results present a true unification of all currently known results. As an illustration, we use these results in a deformation-based morphometry (DBM) analysis to look for regions of the brain where vector deformations of nonmissile trauma patients are related to several verbal memory scores, to detect regions of changes in anatomical effective connectivity between the trauma patients and a group of age- and sex-matched controls, and to look for anatomical connectivity in cortical thickness.

Keywords: SPMs; Deformation-based morphometry; Random field theory

Introduction

Multivariate image data are now collected routinely; that is, several measures of brain function or anatomy are available at each voxel. Examples are vector deformations to warp an MRI image to an atlas standard, diffusion in several different directions (the precursor to the diffusion tensor), and the HRF sampled at several time points. Researchers often want to relate these data to several explanatory variables such as task, stimulus, performance measures, age, gender, or disease state. The simplest approach is to use a multivariate multiple regression model at each voxel.

For a single contrast in the explanatory variables, the natural test statistic is Hotelling's T^2. Random field theory *P* values for the peaks of a Hotelling's T^2 SPM have been available for some time (Cao and Worsley, 1999a). For multiple contrasts, there are several multivariate analogues of the *F*-statistic, such as Wilks's Λ and the Lawley–Hotelling trace. There are as yet no known random field theory *P* values for these SPMs. In this paper, we report new random field theory results for another common multivariate analogue of the *F*-statistic, Roy's maximum root, equivalent to the maximum canonical correlation (Taylor and Worsley, 2004). This completes the random field theory for multivariate linear models (see Table 1).

Since all other SPMs are special cases of Roy's maximum root, then our new results encompass all the random field theory results known so far, including those in the 'Unified' paper (Worsley et al., 1996). More satisfying still, the computer code to actually implement this (the `stat_threshold` function of `fmristat`, available at http://www.math.mcgill.ca/keith/fmristat) is now shorter because separate code for each special type of random field is no longer needed.

Multivariate linear model

Our data set consists of N independent observations, taken for example on N different subjects. Each observation is an image in D dimensions, and at each voxel in the image, there is a set of q random measurements, such as vector deformations, diffusions, or HRF values. We wish to relate these to a set of p contrasts in explanatory variables, such as task, stimulus, performance measures, age, gender, or disease state, common to every voxel (we shall relax this assumption later). At one voxel, we can arrange the observations into an $N \times q$ matrix Y, and the contrasted explanatory variables into an $N \times p$ matrix X. Nuisance variables or variables unaffected by the contrasts (such as a constant term) can be arranged into an $N \times r$ matrix Z. Our problem is to relate Y to X, allowing for Z. This can be formalized by a multivariate linear model, whose theoretical development can be found in any book on multivariate statistics, such as Anderson (1984). The model is

$$Y = XB + ZG + \Xi\Sigma^{1/2} \quad (1)$$

where B and G are $p \times q$ and $r \times q$ matrices of unknown coefficients, Ξ is an $N \times q$ matrix of independent, zero mean, unit

* Corresponding author. Department of Mathematics and Statistics, McGill University, Montreal, Canada H3A 2K6. Fax: +1 514 398 3899.
E-mail address: keith.worsely@mcgill.ca (K.J. Worsley).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.07.026

Table 1
Multivariate linear model SPMs for which random field theory P values for local maxima are now available

		Number of contrasts p	
		1	>1
Number of variates, q	1	T	F
	>1	Hotelling's T^2	Roy's maximum root

variance Gaussian errors, and $\Sigma = \Sigma^{1/2\prime}\Sigma^{1/2}$ is an unknown $q \times q$ covariance matrix of the q random variates. Formally, we wish to test the null hypothesis that $B = 0$.

Union–intersection principle

One way of doing this is Roy's union–intersection principle (Roy, 1953), which is the key to the random field theory in the next section. The idea is as follows: if the alternative hypothesis is a union of (simpler) hypotheses, or equivalently, if the null hypothesis to be rejected is an intersection of (simpler) hypotheses, then test each of the simpler hypotheses and reject only if each of the simpler hypotheses is rejected at some fixed level α^*. This determines the test procedure, but not the level α^* of the simpler tests needed to achieve level α (=0.05, say) for the overall test. In the case of multivariate data, Roy implemented this principle as follows:

1. Take a linear combination of the multivariate image data, creating (simpler) univariate image data.
2. Work out the F-statistic for relating the univariate image data to the multiple contrasts.
3. Maximize the F-statistic over all such linear combinations. The result is Roy's maximum root, R.

If there is at least one such set of univariate data that shows an effect in the population (the union), then there is an effect; if none of them have any effect (the intersection), then there is none. By the union–intersection principle, the overall null hypothesis of no effect is rejected if Roy's maximum root exceeds a threshold (to be determined to control the false-positive rate).

Formally, let v be a $q \times 1$ vector of arbitrary constants and $y = Yv$ be the corresponding linear combination of the components of the data. Then from (1)

$$y = Xb + Zg + \xi\sigma$$

where $b = Bv$, $g = Gv$, ξ is a vector of N independent zero mean, unit variance Gaussian random variables, and $\sigma^2 = v'\Sigma v$. This is just a standard univariate linear model, for which we can evaluate the usual F-statistic F_v for testing $b = 0$ in the usual way (Friston et al., 1995). The degrees of freedom are p and $m = N - p - r$. Then Roy's maximum root statistic can be defined as

$$R = \max_v F_v.$$

Another way of defining Roy's maximum root, more useful in practice, is given in Appendix A. When the number of contrasts is $p = 1$, R is the same as Hotelling's T^2, which equals the maximum of the square of the T statistic for relating y to X, hence the name.

Another approach is via maximal canonical correlations. Let u be a $p \times 1$ vector of constants and let $x = Xu$. Then let $c_{u,v}$ be the squared correlation between x and y removing the effect of (allowing for) Z. To do this, let $R_Z = I - Z(Z'Z)^{-1}Z'$, where I is the $N \times N$ identity matrix, and $x^* = R_Z x$, $y^* = R_Z y$, then

$$c_{u,v} = \frac{(x^{*\prime}y^*)^2}{x^{*\prime}x^* y^{*\prime}y^*}.$$

The maximum canonical correlation is

$$C = \max_{u,v} c_{u,v},$$

which is related to Roy's maximum root by

$$R = \frac{C/p}{(1-C)/m},$$

so that R and C are equivalent SPMs.

Further generalizations are possible. We could extend the above effective connectivity analysis by searching over all reference voxels, as well as over all target voxels. This leads to the cross-correlation SPM introduced by Cao and Worsley (1999b). This is defined in the same way as above, except that X is now random image data sampled at a voxel in an E-dimensional search region T. The number of dimensions of the cross-correlation SPM now becomes $D + E$. In the case of the autocorrelation SPM, the images are the same, and X and Y come from different well-separated voxels. Obviously neighboring voxels should be excluded from the autocorrelation SPM because they are correlated due to the smoothness of the data. It is possible that the images are different, and we are only interested in cross-correlations at the same voxel. This is known as the homologous-correlation SPM (Cao and Worsley, 1999b). Examples include correlations of two different tasks at the same voxel.

Another generalization is to restrict the linear combinations to a part of the sphere, such as a cone (Taylor and Worsley, 2004). Applications include searching over temporal shifts in the HRF, as suggested by Friman et al. (2003).

Random field theory

The key is that the maximum of F over all voxels is the maximum of F_v over all voxels and all linear combinations v. In other words, it is the maximum of an F-statistic SPM over a larger $(D + q)$-dimensional search region: the D dimensions of the voxel and the q dimensions of the linear combinations v. Although it looks like we have made the problem more difficult by adding extra dimensions, in fact, we have made it simpler: all we need to do is to extend random field theory to a larger number of dimensions. The theoretical details are given in Taylor and Worsley (2004), but we report the main ideas and the final result that they lead to, in this section.

The corrected P value of local maxima of a smooth SPM in a D-dimensional search region S is well approximated by

$$P\left(\max_S \mathrm{SPM} > t\right) \approx \sum_{d=0}^{D} \mathrm{Resels}_d(S)\mathrm{EC}_d(t), \tag{2}$$

where Resels_d and EC_d are the resels of the search region and the Euler characteristic (EC) density of the SPM in d dimensions. Expressions for both of these for Gaussian, χ^2, T and F SPMs, and $d \le 3$ can be found in Worsley et al. (1996). Table 2 gives results for the most common case, a T-statistic SPM with m degrees of freedom inside a convex search region S in $D = 3$ dimensions; for arbitrary dimensions, see Appendix B.

Proceeding with the above heuristic, we first note that the search region for v can be restricted to the $q-1$-dimensional unit sphere V consisting of all unit vectors v with $v'v = 1$. We only need to find the resels for the product of S and V. Fortunately, this is quite simple: it only depends on sums of products of the resels of S and the resels of V. This in turn implies that the P value approximation for Roy's maximum root is once again a linear combination of resels of S. In other words, we have the elegant result that the P value approximation to the maximum of the Roy's maximum root SPM is of the same form as (2) but with different EC densities.

There are simple formulas for the resels of V, and putting these all together, we obtain the following very simple expression for the 'EC density' of the Roy's maximum root random field:

$$\mathrm{EC}_d^R(t) = \sum_{i=0}^{q-1} a_i^q \mathrm{EC}_{d+i}^F(t), \tag{3}$$

where $\mathrm{EC}_d^F(t)$ is the EC density of the F-statistic SPM with p and m degrees of freedom and

$$a_i^q = \left(\frac{\pi}{\log 2}\right)^{i/2} \frac{\Gamma\left(\frac{q+1}{2}\right)}{i!\left(\frac{q-1-i}{2}\right)!} \tag{4}$$

if $q-1-i$ is even, and 0 otherwise, $i = 0,\ldots,q-1$. To use this result, we need EC densities beyond the $d = 3$ dimensions given in Worsley et al. (1996). A general expression can be found in Worsley (1994), and for convenience, it is reproduced in Appendix B. The EC density of Hotelling's T^2 (Cao and Worsley, 1999a) now falls out as a special case when $p = 1$.

The reason for the quotes around 'EC density' is that $\mathrm{EC}_d^R(t)$ is not quite the density of the Euler characteristic of the excursion set of voxels where R exceeds t. It is an alternating sum of the EC density of the maximum root (or eigenvalue—see Appendix A), minus the EC density of the second largest root, plus the EC density of the third largest root, and so on. In the case of Hotelling's T^2 ($p = 1$), there is only one nonzero root, and so $\mathrm{EC}_d^R(t)$ is the EC density of the Hotelling's T^2 SPM. For $p > 1$, the presence of the extra terms makes $\mathrm{EC}_d^R(t)$ slightly too small (anticonservative) for the true EC density of R. However, for large thresholds, those usually encountered in practice, the chance that the second largest root exceeds t is much smaller than the chance that the largest root exceeds t, so the extra terms are negligible.

Table 2
Resels for a convex search region S and EC densities for a T-statistic SPM with m degrees of freedom in d dimensions

D	$\mathrm{Resels}_d(S)$	$\mathrm{EC}_d(t)$
0	1	$\int_t^\infty \frac{\Gamma\left(\frac{m+1}{2}\right)}{(m\pi)^{1/2}\Gamma\left(\frac{m}{2}\right)}\left(1+\frac{u^2}{m}\right)^{-(m+1)/2} du$
1	$\frac{2\ \mathrm{Diameter}(S)}{\mathrm{FWHM}}$	$\frac{(4\log_e 2)^{1/2}}{2\pi}\left(1+\frac{t^2}{m}\right)^{-(m-1)/2}$
2	$\frac{\frac{1}{2}\mathrm{Surface\ area}\ (S)}{\mathrm{FWHM}^2}$	$\frac{(4\log_e 2)}{(2\pi)^{3/2}}\frac{\Gamma\left(\frac{m+1}{2}\right)}{\left(\frac{m}{2}\right)^{1/2}\Gamma\left(\frac{m}{2}\right)}\left(1+\frac{t^2}{m}\right)^{-(m-1)/2} t$
3	$\frac{\mathrm{Volume}(S)}{\mathrm{FWHM}^3}$	$\frac{(4\log_e 2)^{3/2}}{(2\pi)^2}\left(1+\frac{t^2}{m}\right)^{-(m-1)/2}\left(\frac{m-1}{m}t^2-1\right)$

FWHM is the effective full width at half maximum of a Gaussian kernel used to smooth the white noise errors ξ in the image data y. The diameter of a convex 3D set is the average distance between all parallel planes tangent to the set. For a ball, this is the diameter; for a box, it is half the sum of the sides. For a 2D set, it is a quarter of the perimeter length. Note also that a 2D set has two sides, so the surface area should be doubled, and its volume is zero.

P values for local maxima of the maximum canonical auto- or cross-correlation SPM C can then be found by a formula very similar to (3):

$$P\left(\max_{S,T} C > t\right) \approx \sum_{d=0}^{D} \mathrm{Resels}_d(S) \sum_{e=0}^{E} \mathrm{Resels}_e(T) \times \sum_{i=0}^{q-1} a_i^q \sum_{j=0}^{p-1} a_j^p \mathrm{EC}_{d+i,e+j}^C(t). \tag{5}$$

Details are given in Taylor and Worsley (2004). The EC densities $\mathrm{EC}_{d,e}^C(t)$ can be found in Cao and Worsley (1999b), and for convenience, they are reproduced in Appendix C.

In fact the above theory can be unified still further. All the other known results can be obtained from the cross-correlation field via (3) and by setting T, the search region for X, to a single point, so its 0-dimensional resels is 1, and the others are 0. For example, to get the FSPM results, set $q = 1$; to get Hotelling's T^2 SPM results, set $p = 1$; in both cases transform C to F or Hotelling's T^2 by $(C/p)/((1\text{-}C)/m)$.

Illustrative applications

As an illustration of the methods, we apply the P values for Roy's maximum root and maximum canonical correlation (but not cross-correlation) to a data set on nonmissile trauma (Tomaiuolo et al., 2004). The subjects were 17 patients with nonmissile brain trauma who were in a coma for 3–14 days. MRI images were taken after the trauma, and the multivariate data were $q = 3$ component vector deformations needed to warp the MRI images to an atlas standard (Collins et al., 1995; Chung et al., 2001) sampled on a 2-mm voxel lattice. The same data were also collected on a group of 19 age- and sex-matched controls.

Damage is expected in white mater areas, so the search region S was defined as the voxels where smoothed average control subject white matter density exceeded 5%. For calculating the resels, this was approximated by a sphere with the same volume, 1.31 litres, which is slightly liberal for a nonspherical search region. The effective full width at half maximum (FWHM), averaged over the search region, was 13.3 mm.

The first analysis was to look for brain damage by comparing the deformations of the 17 trauma patients with the 19 controls, so the sample size is $N = 36$. We are looking at a single contrast, the difference between trauma and controls, so $p = 1$ and the residual degrees of freedom is $m = 34$. In this case, Roy's maximum root is Hotelling's T^2. The $P = 0.05$ threshold, found by equating (2) to 0.05 and solving for t, was $t = 54.0$ (the Bonferroni threshold was 60.3). The thresholded data, together with the estimated contrast (mean trauma-control deformations), are shown in Fig. 1a. A large region near the corpus callosum seems to be damaged. The nature of the damage, judged by the direction of the arrows, is away from the center (see Fig. 1b). This can be interpreted as expansion of the ventricles, or more likely, atrophy of the surrounding white matter, which causes the ventricle/white matter boundary to move outwards.

The second analysis was to try to relate the damage to six measures of memory taken on the 17 trauma patients: immediate and delayed recall of words, short story, and Rey figure (7% were missing, so for purposes of illustration, these were imputed from the rest of the data by fitting a simple two-way ANOVA model). The

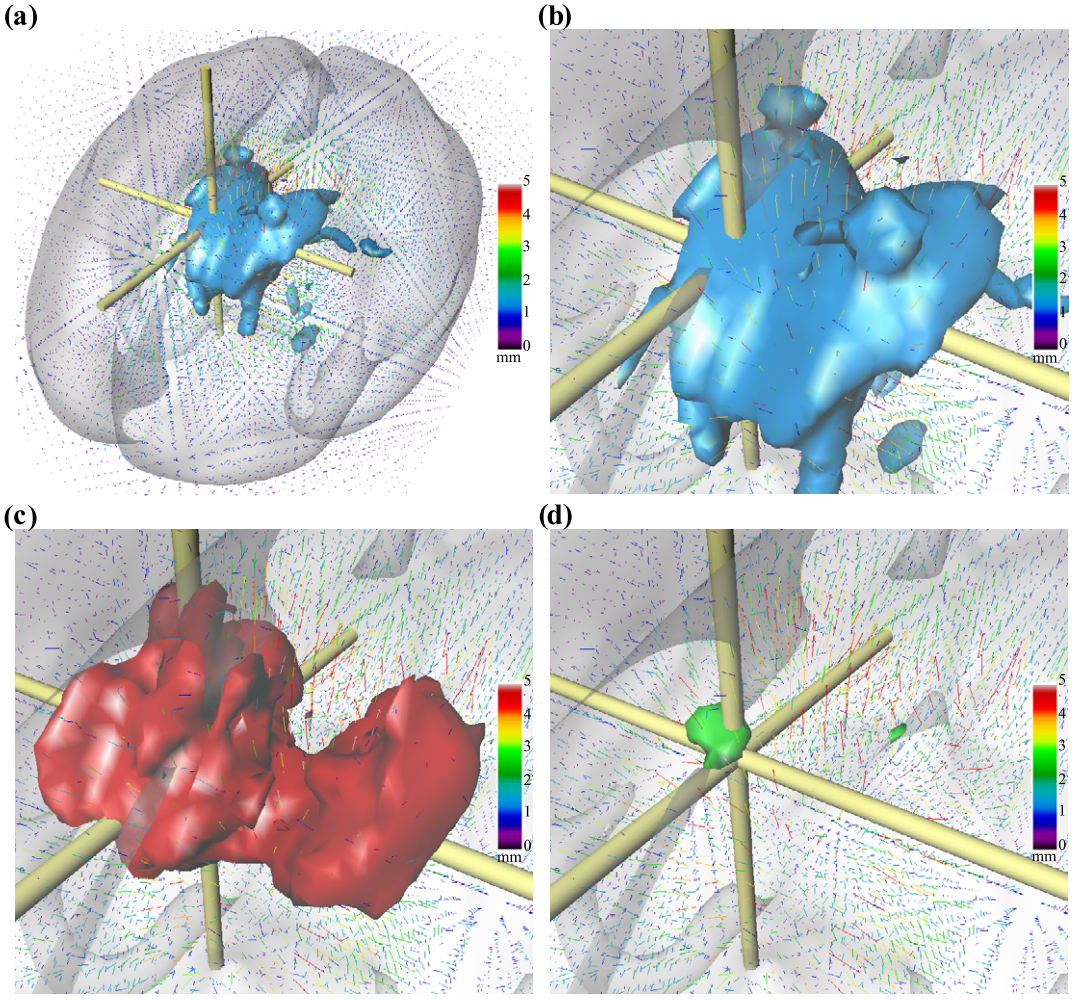

Fig. 1. Deformation-based morphometry of nonmissile trauma data. (a) Trauma minus control average deformations (arrows and color bar), sampled every 6 mm, with Hotelling's T^2 statistic for significant differences (threshold $t = 54.0$, $P = 0.05$, corrected). The reference voxel of maximum Hotelling's T^2 is marked by the intersection of the three axes. (b) Closeup of (a) showing that the damage is an outward movement of the anatomy, either due to swelling of the ventricles or atrophy of the surrounding white matter. (c) Regions of effective anatomical connectivity with the reference voxel, assessed by the maximum canonical correlation (threshold $t = 0.746$, $P = 0.05$, corrected). The reference voxel is connected with its neighbors (due to smoothness) and with contralateral regions (due to symmetry). (d) Regions where the connectivity is different between trauma and control groups, assessed by Roy's maximum root (threshold $t = 30.3$, $P = 0.05$, corrected). The small region in the contralateral hemisphere is more correlated in the trauma group than the control group.

hypothesis was that only damage in certain areas might be related to these memory measures. The sample size is now $N = 17$, with $p = 6$ contrasts and $m = 10$ residual degrees of freedom. The $P = 0.05$ threshold for Roy's maximum root is 712.6, but the largest observed value was 66.8, so there is not enough evidence to relate brain damage to memory scores. This is hardly surprising since the number of contrasts is so high and the number of residual degrees of freedom is so low, which produce a Bonferroni threshold of 283.6 that is in fact much lower than the random field threshold.

We might also be interested in effective anatomical connectivity: are there any regions of the brain whose shape (as measured by the deformations) is correlated with shape at a reference voxel? In other words, the explanatory variables are the deformations at a preselected reference voxel, and the test statistic is the maximum canonical correlation SPM, or equivalently, the Roy's maximum root SPM with $p = 3$. We chose as the reference the voxel with maximum Hotelling's T^2 for damage, marked by axis lines in Fig. 1. Fig. 1c shows the resulting canonical correlation SPM above the $P = 0.05$ threshold of 0.746 for the combined trauma and control data sets removing separate means for both groups ($N = 36$, $p = 3$, $m = 31$). Obviously, there is strong correlation with voxels near the reference, due to the smoothness of the data. The main feature is the strong correlation with contralateral voxels, indicating that brain anatomy tends to be symmetric.

A more interesting question is whether the correlations observed in the control subjects are modified by the trauma (Friston et al., 1997). In other words, is there any evidence for an interaction between group and reference vector deformations? To do this, we simply add another three covariates to the linear model whose values are the reference vector deformations for the trauma patients and the negative of the reference vector deformations for the control subjects. The resulting Roy's maximum root SPM for testing for these three extra covariates, thresholded at 30.3 ($P = 0.05$, $N = 36$, $p = 3$, $m = 28$), is shown in Fig. 1d. Apart from changes in the neighborhood of the reference voxel, there is some evidence of a change in correlation at a location in the contraleteral side, slightly anterior. Looking at the

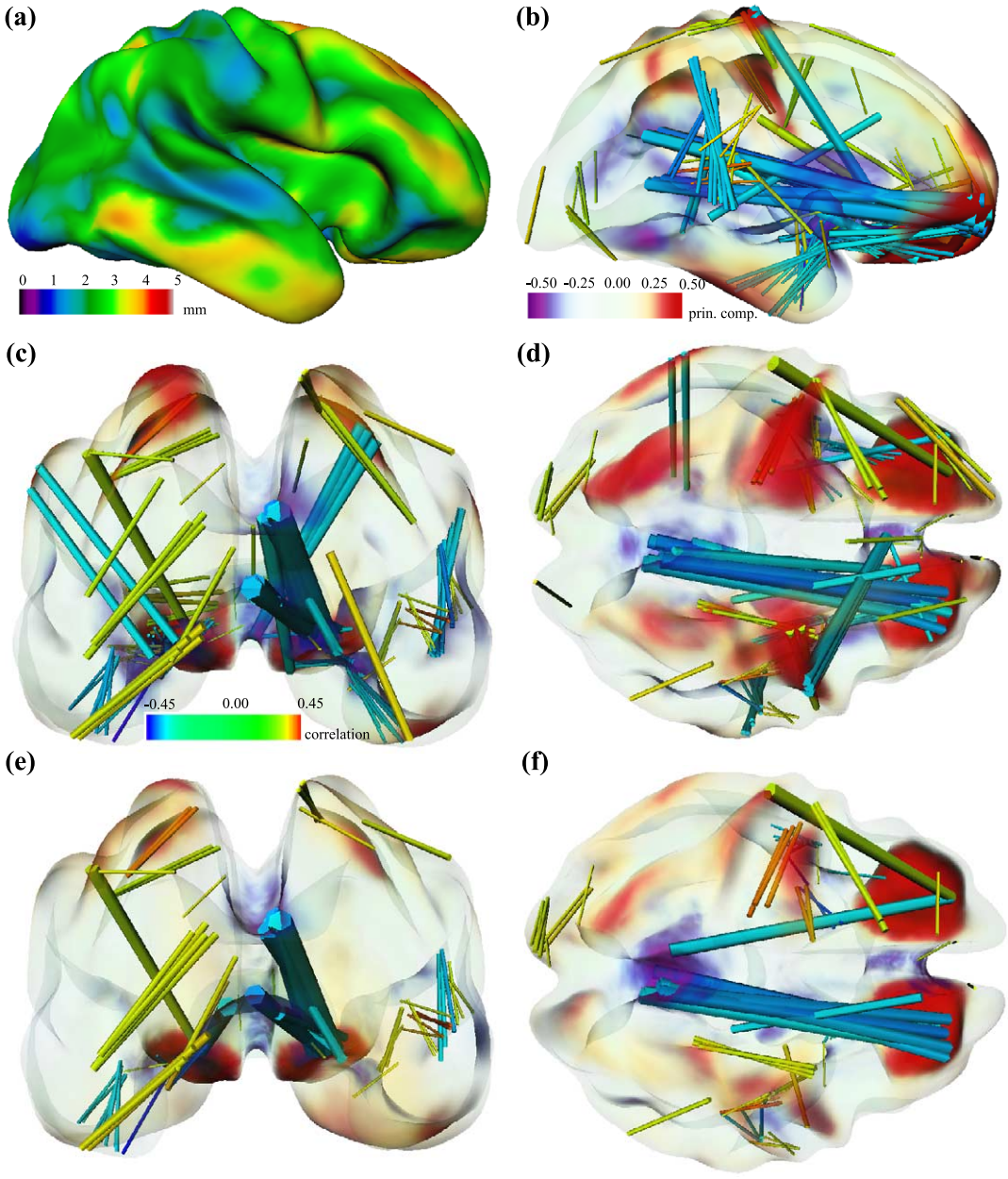

Fig. 2. Connectivity of cortical thickness. (a) Cortical thickness of one subject, smoothed by 20 mm, plotted on the average of the $N = 321$ midcortical surfaces. (b) First principal component of the subject × node matrix of residuals removing a gender effect. The ends of the rods join nodes where the autocorrelation SPM $\sqrt{C}$ of cortical thickness exceeded $t = \pm 0.338$ ($n = 319$ null degrees of freedom, $P = 0.05$, corrected). Only 4D local maxima inside the same hemisphere are shown. (c) Back view and color bar: yellow to red rods indicate positively correlated nodes; blue rods indicate negatively correlated nodes. (d) Top view. Note that red rods tend to join similarly colored principal component regions, whereas blue rods tend to join differently colored principal component regions. (e, f) same as (c, d) but removing an age effect and age–gender interaction ($n = 317$), which also removes some of the effective connectivity.

maximum canonical correlations in the two groups separately, we find that correlation has increased at this location from 0.719 to 0.927, perhaps indicating that the damage is strongly bilateral.

If we chose to search over all reference voxels as well as all target voxels, this would lead to a maximum canonical autocorrelation SPM, for which the P value (5) could be used to set a threshold. This is obviously computationally very expensive, and unless the number of subjects is very large, it is unlikely to reveal much beyond the obvious symmetry reported above. Moreover, it is probable that the Bonferroni bound may be better than the random field theory P value. Furthermore, the theory cannot be applied to the more interesting interaction analysis that we tried above, that is, to look for changes in connectivity between groups. The reason is that the nuisance variables (the main effect or common deformations) are now also random fields, and the theory we have presented here assumes that the nuisance variables are common to all voxels (such as a constant term, age, or gender).

Nevertheless, we illustrate the method on the cortical thickness of $N = 321$ normal adult subjects aged 20–70 years, smoothed by 20-mm FWHM, part of a much larger data set fully described in Goto et

al. (2001). The data are univariate, so $p = q = 1$ (Fig. 2a). We removed a gender effect, then calculated the square root of the autocorrelation SPM, $\sqrt{C}$, for all pairs of the 40,962 triangular mesh nodes, ignoring pairs of voxels that were too close. The search regions S and T are the whole cortical surface, with $D = E = 2$, $\text{Resels}_0(S) = 2$, $\text{Resels}_1(S) = 0$ (since a closed surface has no boundaries), and $\text{Resels}_2(S) = 759$ (see Worsley et al., 1999). From (5) with $n = 319$ null degrees of freedom, the $P = 0.05$ threshold for $\sqrt{C}$ is $t = \pm 0.338$. Figs. 2(b–d) show only 4D local maxima above t inside the same hemisphere. The cortical surface is color-coded by the first principal component of the subject × node matrix of residuals removing a gender effect. This also reveals regions that are similarly correlated. In fact, the positive autocorrelations tend to join regions with either high (red) or low (blue) principal component, and negative autocorrelations tend to join regions with different principal component values. Figs. 2e,f show the same analysis but removing a linear age effect and an age–gender interaction, so that $n = 317$ null degrees of freedom. Note that many of the correlations now disappear, showing that they were induced by age effects. For example, if two regions become thinner over time, than this will induce an apparent positive correlation between these two regions. This illustrates the importance of removing all fixed explanatory variables before carrying out an effective connectivity analysis.

Discussion

The results announced in this paper complete the random field theory for multivariate linear models: we now have results available for P values of local maxima for all types of SPMs. They are incorporated into the `stat_threshold` function of `fmristat`, available at http://www.math.mcgill.ca/keith/fmristat. However, these results are only for local maxima; a common alternative is to detect activation by the size of clusters (connected components) of suprathreshold voxels (Friston et al., 1994). Results are only available for univariate (T and F) SPMs (Cao, 1999), and so far, nothing is known about the size of clusters of multivariate SPMs.

It should be noted that the random field theory P values are only accurate when the FWHM of the SPM is large compared to the voxel size, otherwise the Bonferroni approximation is lower and better (since it is an upper bound). This is particularly so if the degrees of freedom is small, as we saw in the illustrative example of a relation between verbal memory scores and deformations within the trauma group. As a result, `stat_threshold` always outputs the minimum of the random field theory and Bonferroni P values.

Even so, the above results are useful for calculating the Bonferroni bound. Unlike the T, F, and Hotelling's T^2, there is no closed form expression for the probability density function of Roy's maximum root. Instead, the random field P value approximation (2) with $D = 0$ is a very accurate approximation for the P value at a single voxel.

There is one restriction on the spatial correlation of the components. The spatial correlation of every linear combination y must be identical. This rules out the application of our results to a type of canonical correlation analysis proposed by Friman et al. (2002). This paper proposes to use the data at a central voxel and the $3^D - 1$ neighboring voxels as the multivariate observations in a multivariate linear model. If the errors are stationary, the central voxel and all its neighbors have identical spatial correlation structure, but this is not so for every linear combination. In particular, the numerical derivatives have a different spatial correlation structure from that of the voxels. Unfortunately, this also rules out further generalizations that replace neighboring voxels by steerable filters (Friman et al., 2003). Note that correlations between components at the same voxel can be arbitrary and different for different voxels. Aside from this, allowances can be made for nonisotropic errors in the same way as for univariate data—see Worsley et al. (1999), Taylor and Adler (2003), and Hayasaka et al. (2004).

Appendix A. Multivariate test statistics

We wish to test the null hypothesis that $B = 0$. This can be assessed by comparing the $q \times q$ mean hypothesis sum of squares matrix H, to the $q \times q$ mean error sum of squares matrix W, defined as

$$H = Y^{*\prime}X^*(X^{*\prime}X^*)^{-1}X^{*\prime}Y^*/p,$$

$$W = Y^{*\prime}\left[I - X^*(X^{*\prime}X^*)^{-1}X^{*\prime}\right]Y^*/m,$$

where $X^* = R_Z X$, $Y^* = R_Z Y$, and $m = N - p - r$ is the residual degrees of freedom. In the univariate case ($q = 1$), the appropriate test statistic is the F statistic, defined as

$$F = H/W,$$

but in the multivariate case, we cannot simply divide two *matrices* to get a scalar test statistic. Instead, several ways of comparing H to W have been proposed in the statistics literature. All are monotonic functions of the eigenvalues (roots) $f_1,\ldots,f_q$ of $W^{-1}H$, and all are equivalent to the F-statistic in the case $q = 1$. The most natural is the likelihood ratio, equivalent to Wilks's $\Lambda = 1/\prod_{i=1}^{q}(1+f_i p/m)$; another is the Lawley–Hotelling trace $\sum_{i=1}^{q} f_i$. The statistic we shall concentrate on here is Roy's maximum root,

$$R = \max_i f_i.$$

In the case of one contrast ($p = 1$), there is only one nonzero eigenvalue, and so all test statistics are equivalent, and equivalent to Hotelling's T^2. Different test statistics are more sensitive to different departures from the null hypothesis. If the effect of the contrasts is mainly expressed in one variate or in one linear combination of variates, then Roy's maximum root is the most sensitive.

The canonical correlations between X and Y removing the effect of (allowing for) Z are the eigenvalues $c_1,\ldots,c_q$ of

$$(Y^{*\prime}Y^*)^{-1}Y^{*\prime}X^*(X^{*\prime}X^*)^{-1}X^{*\prime}Y^* = (Wm + Hp)^{-1}Hp. \qquad (6)$$

In the univariate case, the (single) canonical correlation is just the square of the usual correlation coefficient between X and Y removing the effect of Z. The canonical correlations all lie between 0 and 1 and are related to the previous eigenvalues (if they are both ordered) by

$$c_i = \frac{f_i p}{m + f_i p}.$$

This implies that all the previous multivariate test statistics are functions of the canonical correlations. In particular, the maximum canonical correlation is a function of Roy's maximum root:

$$C = \frac{Rp}{m + Rp}.$$

Appendix B. EC density of the *T*- and *F*-statistic SPM

The EC density of the *F*-statistic SPM in $d = 0$ dimensions is just the upper tail probability of the *F*-statistic with p and m degrees of freedom, $P(F > t)$. For $d \geq 1$ dimensions it is

$$\begin{aligned}EC_d^F(t) = &\left(\frac{\log 2}{\pi}\right)^{d/2} \frac{2(d-1)!\Gamma\left(\frac{p+m-d}{2}\right)}{m^{(p-d)/2}\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(1+\frac{pt}{m}\right)^{-(p+m-2)/2} \\ &\times \sum_{i=0}^{d-1} (-1)^{d-1-i}(pt)^{i+(p-d)/2} \sum_{j=0}^{\min\{i,d-1-i\}} \\ &\times \binom{\frac{p+m-d}{2}+j-1}{j}\binom{m-1}{i-j}\binom{p-1}{d-1-i-j} m^{-i},\end{aligned}$$

where $\binom{b}{a}$ is $\Gamma(b+1)/(\Gamma(a+1)\Gamma(b-a+1))$ if $0 \leq b \leq a$ and zero otherwise. For the *T*-statistic SPM ($p = 1$),

$$EC_d^T(t) = EC_d^F(t^2)\,\mathrm{sign}(t)^{d+1}/2 + (t < 0)(d = 0).$$

Note that the EC density in Worsley (1994) differs from that given here and in Worsley et al. (1996) by a factor of $(4 \log 2)^{d/2}$. This is because the *P* value approximation in Worsley (1994) is expressed in terms of Minkowski functionals (intrinsic volumes) of *S*, whereas in Worsley et al. (1996), the *P* value is expressed in resels of *S*. The summations have also been rearranged for easier numerical evaluation. The terms in the above expression have been carefully grouped so that they remain numerically stable for large values of m, provided the gamma functions are evaluated on the log scale.

Appendix C. EC density of the cross-correlation SPM

The EC density of the cross-correlation SPM in $d = 0$, $e = 0$ dimensions is just the upper tail probability of the beta distribution with parameters 1/2, $(n-1)/2$, where $n = N - q$ is the null degrees of freedom. For $d > 0$, $e \geq 0$ and $n > d + e$, it is

$$\begin{aligned}EC_{d,e}^C(t) = &\left(\frac{\log 2}{\pi}\right)^{(d+e)/2} \frac{2^{n-1}}{\pi} \sum_{k=0}^{\min\{d-1,e\}} (-1)^k t^{(d+e-1)/2-k}(1-t)^{(n-1-d-e)/2+k} \\ &\times \sum_{i=0}^{k}\sum_{j=0}^{k} \frac{\Gamma\left(\frac{n-d}{2}+i\right)\Gamma\left(\frac{n-e}{2}+j\right)(d-1)!e!}{i!j!(k-i-j)!(n-1-d-e+i+j+k)!(d-1-k-i+j)!(e-k-j+i)!},\end{aligned}$$

where terms with negative factorials are ignored, and $EC_{e,d}^C(t) = EC_{d,e}^C(t)$. Note that the EC density in Cao and Worsley (1999b) differs from that given here by a factor of $(4 \log 2)^{(d+e)/2}$. This is because the *P* value approximation in Cao and Worsley (1999b) is expressed in terms of Minkowski functionals (intrinsic volumes), whereas here, the *P* value (5) is expressed in resels. The summations have also been rearranged for easier numerical evaluation.

References

Anderson, T.W., 1984. An Introduction to Multivariate Statistical Analysis. Second ed. Wiley, New York.

Cao, J., 1999. The size of the connected components of excursion sets of χ^2, *t* and *F* fields. Adv. Appl. Probab. 31, 579–595.

Cao, J., Worsley, K.J., 1999a. The detection of local shape changes via the geometry of Hotelling's T^2 fields. Ann. Stat. 27, 925–942.

Cao, J., Worsley, K.J., 1999b. The geometry of correlation fields, with an application to functional connectivity of the brain. Ann. Appl. Probab. 9, 1021–1057.

Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Hum. Brain Mapp. 3, 190–208.

Chung, M.K., Worsley, K.J., Paus, T., Cherif, C., Collins, D.L., Giedd, J.N., Rapoport, J.L., Evans, A.C., 2001. A unified statistical approach for deformation based morphometry. NeuroImage 14, 595–606.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., 2002. Detecting neural activity in fMRI using maximum correlation modeling. NeuroImage 15, 386–395.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., 2003. Adaptive analysis of fMRI data. NeuroImage 19, 837–845.

Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. Hum. Brain Mapp. 1, 214–220.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-B., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functionaliimaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Büchel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6, 218–229.

Goto, R., Kawashima, R., Zijdenbos, A., Neelin, P., Lerch, J., Sato, K., One, S., Nakagawa, M., Taki, Y., Sugiura, M., Watanabe, J., Fukuda, H., Evans, A.C., 2001. Normal aging and sexual dimorphism of Japanese brain. NeuroImage 13, S794.

Hayasaka, S., Luan-Phan, K., Liberzon, I., Worsley, K.J., Nichols, T.E., 2004. Non-stationary cluster-size inference with random field and permutation methods. NeuroImage 22, 676–687.

Roy, S.N., 1953. On a heuristic method of test construction and its use in multivariate analysis. Ann. Math. Stat. 24, 220–238.

Taylor, J.E., Adler, R.J., 2003. Euler characteristics for Gaussian fields on manifolds. Ann. Probab. 31, 533–563.

Taylor, J.E., Worsley, K.J., 2004. Random Fields of Multivariate Test Statistics, with an Application to Shape Analysis. In preparation.

Tomaiuolo, F., Worsley, K.J., Lerch, J., Di Paulo, M., Carlesimo, G.A., Bonanni, R., Caltagirone, C., Paus, T., 2004. Change in white matter in long-term survivors of severe non-missile traumatic brain injury: a computational analysis of magnetic resonance images. J. Neurotrauma (submitted).

Worsley, K.J., 1994. Local maxima and the expected Euler characteristic of excursion sets of χ^2, *F* and *t* fields. Adv. Appl. Probab. 26, 13–42.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73.

Worsley, K.J., Andermann, M., Koulis, T., MacDonald, D., Evans, A.C., 1999. Detecting changes in nonisotropic images. Hum. Brain Mapp. 8, 98–101.