An improved theoretical P-value for SPMs based on discrete local maxima

K.J. Worsley

August 9, 2005

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. West, Montreal, Québec, Canada H3A 2K6, and McConnell Brain Imaging Centre, Montreal Neurological Institute, 3801 University Street, Montreal, Québec, Canada H3A 2B4.
tel: 1-514-398-3842, fax: -3899, web: http://www.math.mcgill.ca/keith

We present a new continuity correction to the P-value for local maxima of a statistical parametric map that bridges the gap between small $FWHM$, when the Bonferroni correction is accurate, and large $FWHM$, when random field theory is accurate. The new method, based on discrete local maxima, is always an upper bound (like the Bonferroni), but lower and hence more accurate for large $FWHM$, without increasing false positives. It resulted in P-values that were $\sim$43% lower than the best of Bonferroni or random field theory methods when applied to a typical fMRI data set.

1 Introduction

The last step in the statistical analysis of a statistical parametric map Z is to assign P-values P to local maxima of height t. Apart from costly simulations or permutations, the choice is between a Bonferroni correction (BON):

$$P \le P_{\rm BON} = N \times P(Z > t) \tag{1}$$

where N is the number of voxels in the search region, or random field theory (RFT), which for large search regions has the form:

$$P \approx P_{\rm RFT} = R \times {\rm EC}(t) \tag{2}$$

where R is the number of resels and EC is the Euler characteristic density of the SPM (Worsley *et al.*, 1992, 1996). The resels of a search region in D dimensions is

$$R = V/FWHM^D \tag{3}$$

where V is the volume and $FWHM$ is the Full Width at Half Maximum of a Gaussian kernel applied to white noise that would give the same smoothness as the noise component of the

data. Note that constant $FWHM$ across the SPM is no longer necessary - see Worsley *et al.* (1999) and Taylor & Adler (2003). For a comparative review see Nichols and Hayasaka (2003).

Which method to use? BON is an upper bound, too high if the data is smooth, very accurate if the data is not smooth; RFT is an approximation that is only accurate if the data is smooth. In between, for moderate smoothing, there is a gap where BON is too high and RFT is not accurate.

Figure 1 illustrates the point for Gaussian SPMs. The Gaussian SPMs were simulated on a 32^3 lattice of voxels by smoothing white noise with an isotropic Gaussian kernel (see Section 2.4). For $FWHM < 1$ voxel BON is accurate; for $FWHM > 5$ voxels RFT is accurate.

Also shown in Figure 1 are RFT results for T-statistic local maxima with $\nu = 10, 20$ degrees of freedom, based on results in Worsley *et al.* (1996). They are considerably larger than the Gaussian thresholds, particularly for small degrees of freedom ($\nu = 10$). This implies that the SPM has to be much smoother before RFT is accurate ($FWHM > 10$ voxels for $\nu = 10$ degrees of freedom).

A simple solution is to take the best of both, i.e. the minimum of the two P's or thresholds. This will give P-values that are never more than twice the true value in the example in Figure 1. This method is currently used by the FMRISTAT (Worsley *et al*, 2002) and SPM software. But perhaps we can do better for the mid range of $FWHM$.

Since resels count 'effective' number of independent voxels in the data, why not replace the number of voxels in P_{BON} by the number of resels (perhaps adjusted by a constant)? This heuristic has occurred to many researchers, e.g. Xiong *et al.* (1995), but unfortunately it won't work - see Figure 1. For high thresholds, the Bonferroni and random field P's have different functional forms, e.g. for a Gaussian SPM:

$$\begin{aligned} P_{\mathrm{BON}} &\approx \mathrm{constant}_1 \times \exp(-t^2/2) \times t^{-1}, \\ P_{\mathrm{RFT}} &\approx \mathrm{constant}_2 \times \exp(-t^2/2) \times t^{D}, \end{aligned}$$

so that no adjustment to the constants can make the two formulas agree for all t.

We could always smooth the data, which might in fact increase detectability at the expense of losing resolution. Another alternative is to interpolate the data to a smaller voxel size so that we can apply RFT. However interpolated T-statistics do not have a T distribution, so we must interpolate the raw data, then recalculate the T SPM. Note that the new RFT P (and threshold) may be higher than the old BON P, suggesting that we might be losing sensitivity, but then this is as it should be: it is possible that local maxima of the interpolated data might be *higher* than the original local maxima, and the higher P's and thresholds now compensate for this. These two computationally expensive alternatives are efforts to make the data fit the theory. It seems preferable to make the theory fit the data, which is what we propose in this paper.

2 A continuity correction: the DLM method

We now propose a 'continuity correction' for Gaussian SPMs that bridges the gap between small $FWHM$, where BON is accurate, and large $FWHM$, where RFT is accurate. The proposed P-value approximation is the expected number of discrete local maxima (DLM) above threshold:

$$P \leq P_{DLM} = \sum_x P(Z > t \text{ and } Z > \text{neighbouring } Z'\text{s}), \tag{4}$$

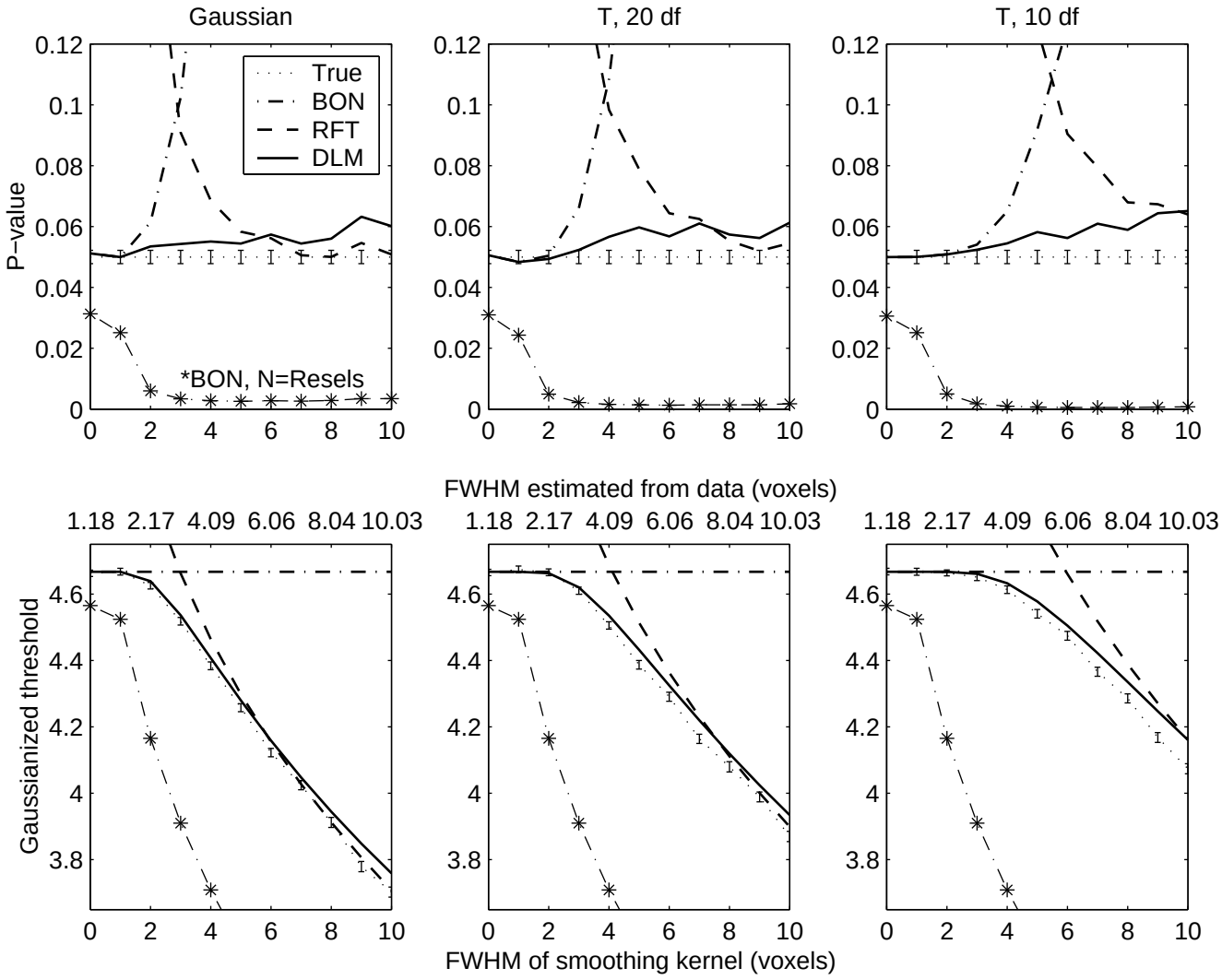

Figure 1: Comparison of Bonferroni (BON), Random Field Theory (RFT) and proposed Discrete Local Maxima (DLM) P-values and Gaussianized thresholds for a 32^3 voxel search region, as a function of kernel $FWHM$ relative to voxel size, $FWHM/v$. The top axis shows the $FWHM$ estimated from the data, $\widehat{FWHM}/v$. The thresholds used for the P-values in the top row were the P=0.05 thresholds based on 9999 simulations (True). Error bars are ± 1 Sd. The thresholds in the bottom row are at P=0.05. The lower Bonferroni values (*) are based on replacing the number of voxels by the number of resels.

where summation is over all voxels x in the search region. The $2D$ neighbours are those that differ by just one voxel in each lattice direction. In contrast, $P_{\rm BON}$ is the expected number, and $P_{\rm RFT}$ is the expected Euler characteristic, of the voxels above threshold. We shall show that, like BON, DLM is conservative, which is reassuring for practical applications, but unlike BON it is very accurate for all $FWHM$; for large $FWHM$ and thresholds DLM converges to RFT.

The calculations needed for DLM are intractable unless we make more assumptions about the spatial correlation structure of the data. We shall assume that the spatial correlation function is Gaussian at neighbouring voxels, with the principal axes aligned with the axes of the voxel lattice. That is, we assume that the spatial correlation between two voxels $x_1, x_2 \in \Re^D$ that are neighbours of $x \in \Re^D$ (differing by one voxel in one lattice direction only) is of the form

$$\exp(-(x_1 - x_2)'\Lambda(x_1 - x_2)/2) \tag{5}$$

for some positive diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_D)$ that can vary with x. In contrast, no assumptions are required for BON, and almost none for RFT.

2.1 Gaussian SPM

The result is as follows, and the derivation is in Taylor & Worsley (2005). It only involves the correlation ρ_d between adjacent voxels along each lattice axis d at voxel x. In terms of the above assumptions, $\rho_d = \exp(-\lambda_d v_d^2/2)$ where v_d is the voxel size in direction d, although it should be stressed that DLM depends directly on ρ_d rather than λ_d. First let the Gaussian density and uncorrected P-values be

$$\phi(z) = \exp(-z^2/2)/\sqrt{2\pi}, \quad \Phi(z) = \int_z^\infty \phi(u)du, \tag{6}$$

respectively. Let

$$\alpha = \sin^{-1}\left(\sqrt{(1-\rho^2)/2}\right), \qquad h = \sqrt{\frac{1-\rho}{1+\rho}},$$

and

$$\begin{aligned} q(\rho, z) &= \frac{1}{\pi}\int_0^\alpha \exp(-\tfrac{1}{2}h^2z^2/\sin^2\theta)d\theta & (7)\\ &= \frac{hz}{\pi}\int_0^{\exp(-z^2/(1+\rho)^2)} \frac{1}{-2\log y\sqrt{-2\log y - h^2z^2}}\, dy. & (8)\end{aligned}$$

Equation (8), obtained from (7) by changing variables to $y = \exp(-\frac{1}{2}h^2z^2/\sin^2\theta)$, is more convenient for numerical integration. Then define

$$Q(\rho, z) = 1 - 2\Phi(hz^+) + q(\rho, z) \tag{9}$$

where $z^+ = z$ if $z > 0$ and 0 otherwise. Then the P-value of local maxima is bounded by

$$P \le P_{\rm DLM} = \sum_x \int_t^\infty \left(\prod_{d=1}^D Q(\rho_d, z)\right)\phi(z)dz, \tag{10}$$

where the summation is over all voxels x in the search region, and the summand depends on x only through ρ_d.

A boundary correction is easily implemented. For a voxel on the boundary of the search region with just one neighbour in axis direction d, replace $Q(\rho, z)$ by $1 - \Phi(hz)$, and by 1 if it has no neighbours.

When the voxels are independent ($\rho_d = 0$), P_{DLM} is very slightly smaller than P_{BON} but slightly larger than the true P, specifically,

$$P_{\mathrm{DLM}} = \left(1 - (1 - P)^{(2D+1)/N}\right) N/(2D + 1). \tag{11}$$

In fact the difference is hardly noticeable in Figure 1 at $FWHM = 0$. For large $FWHM$ relative to v, it can be shown that P_{DLM} converges to P_{RFT} for large thresholds, specifically,

$$P_{\mathrm{DLM}} = P_{\mathrm{RFT}} \times (1 + O(1/t^2)) \tag{12}$$

where $t^2 O(1/t^2)$ converges to a positive constant as t approaches infinity (see Taylor & Worsley, 2005).

Note also that P_{DLM} does not require stationarity nor isotropy; the only requirement is the local axis-aligned Gaussian structure (5) of the spatial correlations. In the Appendix we argue that even this requirement can be relaxed and P_{DLM} appears to be still conservative under all circumstances.

2.2 Other non-Gaussian SPMs

Unfortunately a continuity correction for other statistics, such as T- or F-statistics, using the DLM method, seems to be possible but quite messy, involving many more integrals that must be evaluated numerically. As an approximation, we propose Gaussianizing the statistic, then re-adjusting the $FWHM$ so that P_{RFT} of the statistic matches P_{RFT} of a Gaussian statistic, then applying the DLM method. The steps for a T-statistic are as follows (the same procedure would be followed for other statistics, such as an F-statistic):

1. Gaussianize the T-statistic local maximum $T_{\max}$ by finding its (uncorrected) P-value, then find the Gaussian statistic $Z_{\max}$ that would have the same P-value. Specifically, $Z_{\max} = \Phi^{-1}(P(T > T_{\max}))$.

2. Find $c = \mathrm{EC_T}(T_{\max})/\mathrm{EC_Z}(Z_{\max})$, the ratio of the EC densities for T and Gaussian random fields (Worsley *et al.*, 1996). Note that $c > 1$ because the T field is rougher than the Gaussian field.

3. Calculate the usual sample correlation $\widehat{\rho}_d$ of neighbouring residuals from a linear model along lattice axis d.

4. Let $f = c^{2/D}$ and $\rho_d = |\widehat{\rho}_d|^f \mathrm{sign}(\widehat{\rho}_d)$, preserving the sign if by chance $\widehat{\rho}_d$ is negative. This adjusts $FWHM$, and hence the resels, so that P_{RFT} is identical for the T and Gaussian random fields. Note that raising correlations to a fixed power also preserves the local Gaussian correlation structure (5).

5. Now calculate P_{DLM} as above for $t = Z_{\max}$.

This method of adjusting the $FWHM$ to approximate non-Gaussian DLM is similar in spirit to the method proposed in Worsley *et al.* (1992) to approximate non-Gaussian RFT before exact RFT appeared in Worsley (1994). In fact the above method capitalizes on the exact non-Gaussian RFT to make the DLM adjustment.

2.3 Computational short cuts

The amount of computation for the DLM method is quite large - we must repeat the above procedure for each local maximum $T_{\max}$, since c, and hence $\widehat{\rho}_d$, depend on $T_{\max}$. We can cut down on computation by first evaluating $Q(\rho, z)$ for an array of values of ρ and z, then use interpolation to get $Q(\rho_d, z)$ for each ρ_d along each axis at each voxel. This cuts the number of numerical integrals from $D+1$ per voxel down to just one per voxel.

Even so it took 20 minutes to do this for the 13 local maxima in Table 1, so we suggest an approximation that cuts the number of integrals down to at most 27 per local maximum. The idea is to replace $\widehat{\rho}_d$ at each voxel by its average over the search region, evaluate the integral once, then multiply by the number of voxels. This must be repeated for each of the 3^D different types of neighbourhoods (0, 1 or 2 neighbours in each of the D axis directions). This cuts the time down to just 2 seconds for the same 13 local maxima.

The remaining question is how to choose a scale for averaging the correlations. Looking at (8) and (9) we see that for large ρ (and so small h), $Q(\rho, z) \propto h \propto \sqrt{1-\rho}$, so this suggests that we average $\sqrt{1-\rho}$. In fact this would be exact for large ρ if the D dimensional histogram of $\sqrt{1-\widehat{\rho}_1}, \ldots, \sqrt{1-\widehat{\rho}_D}$ (over voxels in the search region) was the product of its marginals, or in other words, if $\sqrt{1-\widehat{\rho}_1}, \ldots, \sqrt{1-\widehat{\rho}_D}$ were 'independent'.

To implement this, we suggest adding one more step between steps 3 and 4 to the above procedure:

3a. Replace $\widehat{\rho}_d$ by $\bar{\rho}_d$ chosen so that

$$\sqrt{1-\bar{\rho}_d} = \sum_x \sqrt{1-\widehat{\rho}_d}/N.$$

where summation is over all voxels x in the search region.

We denote this approximation by $\bar{P}_{\mathrm{DLM}}$.

2.4 Simulations

The methods were compared on simulated data that were constructed to be as close to reality as possible. In particular, when we calculated the resels R, we were careful to use values of $FWHM$ estimated from the data, not that used in the kernel to create the data. To do this, we first specify $FWHM$, then create the discrete kernel

$$k(x) = \exp(-x^2(4\log 2)/FWHM^2), \tag{13}$$

then convolve $k(x)$ with Gaussian white noise, using Fourier methods, to simulate smooth Gaussian random fields sampled on a discrete lattice. The correlation between neighbouring voxels is then:

$$\rho = \sum_x k(x)k(x+v) \;/\; \sum_x k(x)^2 \tag{14}$$

where v is the voxel size. From this we obtain the $FWHM$, to be used in the random field P-values, derived from the discrete data in the same way as standard packages such as SPM and FMRISTAT (Kiebel *et al.*, 1999). This estimate depends on the standard deviation of the numerical derivative of residuals $\epsilon(x)$ and is given by:

$$\widehat{FWHM} = \frac{\mathrm{Sd}(\epsilon(x))\sqrt{4\log 2}}{\mathrm{Sd}((\epsilon(x)-\epsilon(x+v))/v)} \to \frac{v\sqrt{2\log 2}}{\sqrt{1-\rho}}. \tag{15}$$

as the sample size becomes infinite. Then we replace resels R by the asymptotic value of $V/\widehat{FWHM}^D$ in the calculation of P_{RFT}. Note that the estimate $\widehat{FWHM}$ is slightly larger than $FWHM$. In Figure 1, we plot P-values and thresholds against the specified $FWHM$; the estimate $\widehat{FWHM}$ is shown on the top axis.

Fourier methods were used to smooth the volume, so the data is periodic. While this is not realistic, it saves computer time, especially for the T-statistics, and it simplifies the P-value calculations since there are no boundaries to the search region. More realistic simulations using a ball search region of the same volume (32^3) but embedded in a 64^3 periodic volume gave essentially similar results to those in Figure 1 (not shown).

To estimate the true $P = 0.05$ threshold, SPMs were simulated $M = 9999$ times. The $(M+1)P$th largest global maximum estimates the true threshold (unbiasedly if the simulated values were uniform). The standard deviation of this estimate was itself estimated by specifying a small width $\delta = 0.02$, then by

$$\mathrm{Sd} = \frac{(M+1)(P+\delta)\mathrm{th} - (M+1)(P-\delta)\mathrm{th}}{2\delta}\sqrt{\frac{P(1-P)}{M+2}}. \tag{16}$$

This is based on the usual linear approximation to the variance of function of a random variable. The first term is an estimate of the inverse of the probability density, the second square root is the Sd of the sample P-value.

3 Results

Results for a $P = 0.05$ threshold and $N = 32^3$ voxels (equivalent to $\sim$1 litre of 3mm voxels) are shown in Figure 1. The SPMs were isotropic Gaussian and T-statistics with 10 and 20 degrees of freedom. Note again that we plot P-values and thresholds against the specified kernel $FWHM$, not the estimate $\widehat{FWHM}$, relative to voxel size v. True thresholds were estimated by 9999 simulations (error bars show one standard deviation).

The DLM P-value is always an accurate upper (conservative) bound on the true P-value, which almost equals BON when $FWHM = 0$ and slightly overestimates RFT when $FWHM > 6$ voxels. In between DLM is better than either of them. The greatest discrepancy occurs at $FWHM = 3$ voxels, where DLM is about half either of the others.

The T-statistic thresholds are always larger than the Gaussian thresholds, even after Gaussianization, but of course BON remains the same. Our proposed DLM method transforms the $FWHM$ so that the Gaussian RFT matches the T RFT, then applies the same transformation to the Gaussian DLM. The agreement with the simulations appears to be as good as in the Gaussian case.

The methods were compared at $P = 0.01$ and the results were very similar. This is to be expected, since $P = 0.05$ results are very similar to $P = 0.01$ results for a search region of 1/5 the size (not shown).

4 Application

We compared our methods on fMRI data from one run of one subject from a study in pain perception (Chen *et al.*, 2002), fully described Worsley *et al.*, (2002). A subject received 9s hot stimulus to the right calf, 9s rest, 9s warm stimulus, 9s rest, repeated 10 times for a total

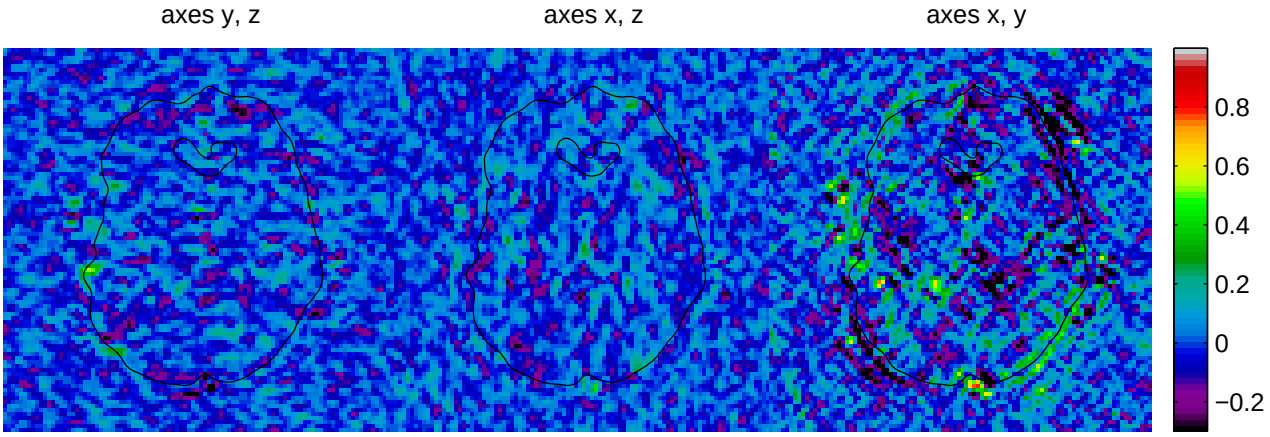

Figure 2: Real fMRI data partial correlation of whitened residuals between neighbours on different lattice axes, partialing out the central voxel. Correlations between each pair of lattice axes are shown for just one slice. The DLM method is based on the assumption that these correlations are zero. This seems to be the case, except close to the brain boundary (black contours).

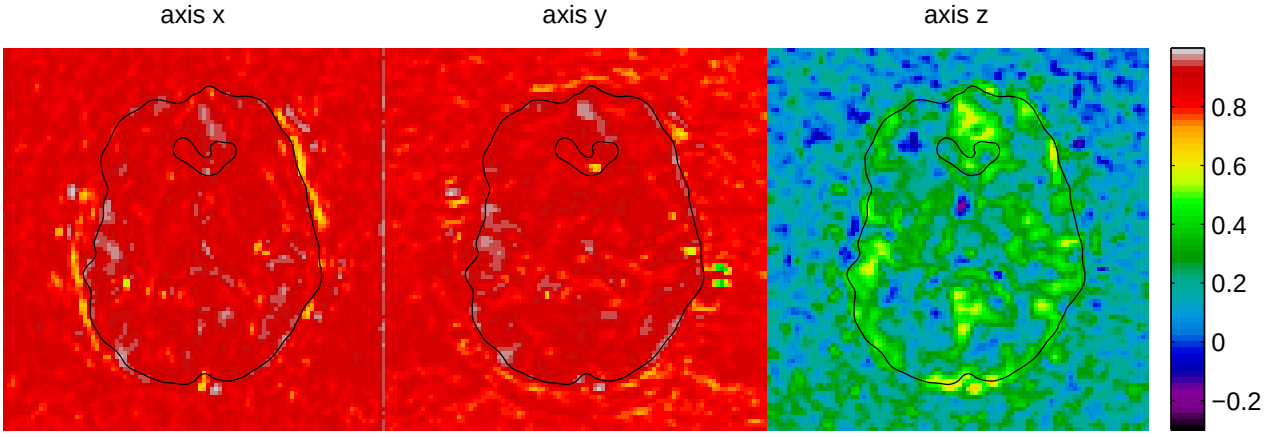

Figure 3: Real fMRI data correlation of whitened residuals between neighbours on the same lattice axes. Correlations for each lattice axis are shown for just one slice. Correlation is high in-slice due in part to 6mm in-slice smoothing during motion correction, close to zero between slices because of the 7mm slice separation, but still noticeably larger in grey matter. The colour bar is the same as for Figure 2.

of 360s (TR=3s). The voxel sizes were $2.3 \times 2.3 \times 7.0$mm, and 6mm smoothing was applied in-slice during motion correction. This is an interesting test case, because $FWHM \sim 2.55$ voxels (averaged across voxels and axes) which puts us in the zone where BON and RFT are not very accurate.

The proposed DLM method depends on the assumption of a local Gaussian correlation structure. To check this, we focused on one consequence: neighbouring voxels on different lattice axes should be independent, conditional on the central voxel. We therefore calculated the partial correlation of whitened residuals (see Worsley *et al.*, 2002) between neighbours on different lattice axes, partialing out the central voxel. This should be zero if the assumption is true. Figure 2 appears to confirm this - the partial correlations are as close to zero inside the brain as outside, except on the brain boundary.

To apply the DLM method, we first estimated the correlation between whitened residuals $\epsilon(x)$ (see Worsley *et al.*, 2002) at neighbouring voxels along each axis by

$$\widehat{\rho} = (\mathrm{Cor}(\epsilon(x), \epsilon(x+v)) + \mathrm{Cor}(\epsilon(x), \epsilon(x-v)))/2 \tag{17}$$

(dropping subscript d). Figure 3 shows the results for one slice. The correlations in the x and y directions are very high ($\bar{\rho}_1 = 0.87$, $\bar{\rho}_2 = 0.89$), particularly near the boundary of the brain, due in part to the 6mm in-slice smoothing applied during motion correction. The correlations in the z direction ($\bar{\rho}_3 = 0.27$) are small since no smoothing was applied and the voxel size is quite large. Nevertheless there is some evidence of higher correlations in grey matter.

P-values for local maxima of a T-statistic (110 df) for the effect of the hot stimulus minus the warm stimulus are shown in Table 1 and Figure 4. We first notice that BON is better than RFT, presumably because $FWHM$ is quite small (2.55 voxels, averaged across voxels and axes). However DLM is better than BON, giving P-values $\sim$43% lower. The faster DLM method using averaged correlations is just as accurate, giving essentially the same conclusions. The net result is that DLM detects one extra local maximum at $P < 0.05$, and three extra at $P < 0.01$ (see Figure 5). One of these is in the right thalamus (ipsilateral to the stimulus), complementing the larger activation already detected in the left thalamus (contralateral to the stimulus), a region implicated in pain perception in the full study (Chen *et al.*, 2002).

5 Discussion

We have developed a method of determining P-values for local maxima of SPMs, based on discrete local maxima (DLM). Under the assumptions of a local axis-aligned Gaussian spatial correlation, DLM is an upper bound (like the Bonferroni), so if it is lower than other methods, it has to be better. It depends solely on the correlation of whitened residuals between adjacent voxels. It does not depend on isotropy, and there is a simple boundary correction. It is reasonably accurate, though still conservative, over the middle range of $FWHM$ to voxel size, the values often encountered in practice. The Bonferroni (BON) method is still more accurate for very small $FWHM$, and the random field theory (RFT) method is more accurate for large $FWHM$. To cover the whole range of $FWHM$, we recommend simply taking the best of the three methods: BON, RFT and DLM.

The only possible drawback is that DLM depends on a local axis-aligned Gaussian correlation structure for the spatial correlation. This is assured if the data is white noise smoothed with an axis-aligned Gaussian filter (the most common filter shape employed in brain mapping). It appears to be a reasonable assumption for real fMRI data. This assumption is not too

$T_{\max}$	P_{BON}	P_{RFT}	P_{DLM}	$\bar{P}_{\text{DLM}}$
5.71	0.0015**	0.0058**	0.0009**	0.0009**
5.29	0.0098**	0.0307*	0.0056**	0.0057**
5.18	0.0158*	0.0466*	0.0091**	0.0092**
5.17	0.0162*	0.0476*	0.0093**	0.0095**
5.17	0.0164*	0.0481*	0.0094**	0.0096**
5.15	0.0178*	0.0516	0.0102*	0.0104*
5.13	0.0192*	0.0549	0.0109*	0.0111*
5.11	0.0211*	0.0598	0.0121*	0.0123*
5.09	0.0227*	0.0636	0.0129*	0.0132*
4.92	0.0470*	0.1190	0.0267*	0.0272*
4.77	0.0862	0.1996	0.0487*	0.0495*
4.59	0.1805	0.3727	0.1014	0.1031
4.54	0.2241	0.4468	0.1256	0.1277

Table 1: Real fMRI data local maximum T-statistics $T_{\max}$ (110 df) and their P-values for the effect of the hot stimulus minus the warm stimulus. Only local maxima between 4.5 and 6 are shown. P_{BON} = Bonferroni; P_{RFT} = random field theory; P_{DLM} = discrete local maxima; $\bar{P}_{\text{DLM}}$ = discrete local maxima using averaged correlations ($^*P < 0.05$, $^{**}P < 0.01$). The discrete local maxima method proposed in this paper is clearly the best. Using averaged correlations instead of voxel correlations is just as accurate and a lot faster to calculate.

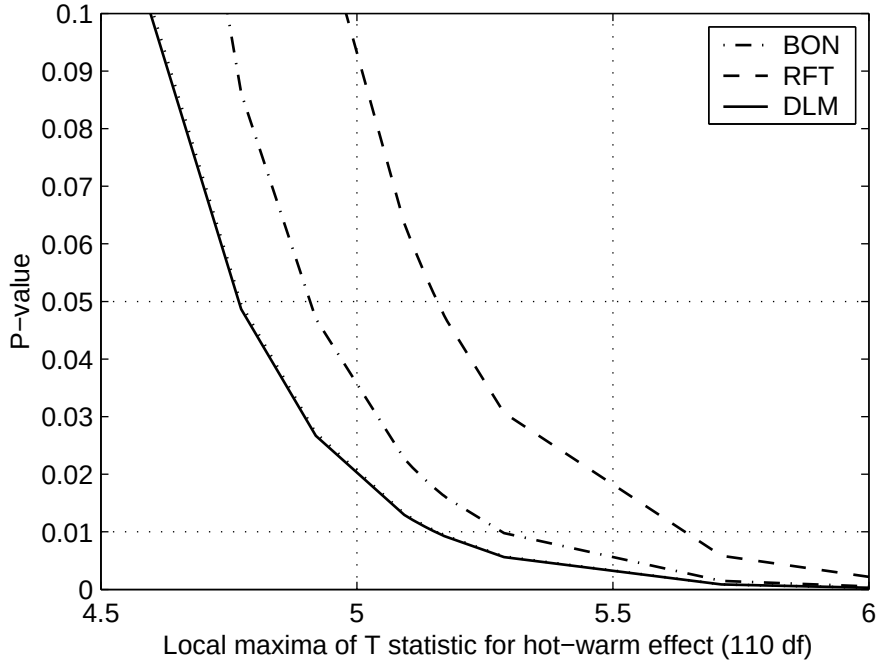

Figure 4: The same data as in Table 1. The averaged correlations $\bar{P}_{\text{DLM}}$ are shown by a dotted line, almost indistinguishable from the voxel correlations P_{DLM}.

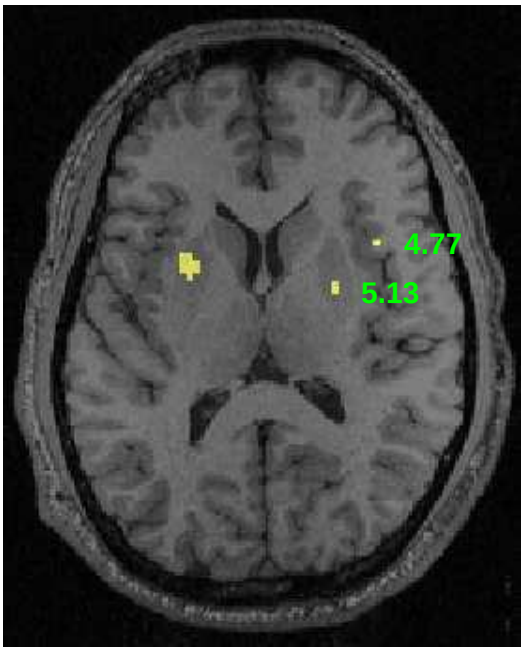

Figure 5: The same T-statistic as in Table 1 thresholded at $P < 0.05$ using the DLM method ($T > 4.76$). The slice is the same as in Figures 2 and 3. $T_{\max} = 4.77$ (labelled) would not have been detected using either BON or RFT; $T_{\max} = 5.13$ (labelled, right thalamus, ipsilateral to the pain stimulus) would have been detected using BON, but not RFT (see Table 1).

restrictive for two reasons. The first is that locally, any smooth SPM has a correlation function that resembles the Gaussian, since both are locally quadratic to a first approximation. Second, even if this quadratic approximation is not axis-aligned, the DLM P-value is always larger than RFT (see Appendix 5), suggesting that DLM is conservative even if the DLM assumptions are not met.

Unfortunately there does not seem to be a reasonable way of doing the necessary calculations for non-Gaussian SPMs. Instead we have proposed an approximation based on a "double" Gaussianization: we adjust the SPM so that it has a Gaussian distribution at each voxel, then we adjust the *FWHM* so that its RFT P-value matches the RFT P-value of a Gaussian SPM. We then apply the DLM method, then convert back. Our simulations with a T-statistic SPM show that this method is reasonably accurate.

Finally, the above method has been implemented in the FMRISTAT package, available from `http://www.math.mcgill.ca/keith/fmristat`.

Appendix Non-Gaussian spatial correlation

We now give some evidence that DLM is always conservative under all circumstances, even if the local spatial correlation structure is not axis-aligned Gaussian. Theoretical reasons are given in Taylor & Worsley (2005). Further evidence is given by the simulations of the Gaussian SPM shown in Figure 6. We kept the Gaussian spatial correlation, but aligned the major axis in the diagonal direction (1,1,1). We then increased the anisotropy, a = largest/smallest eigenvalue -

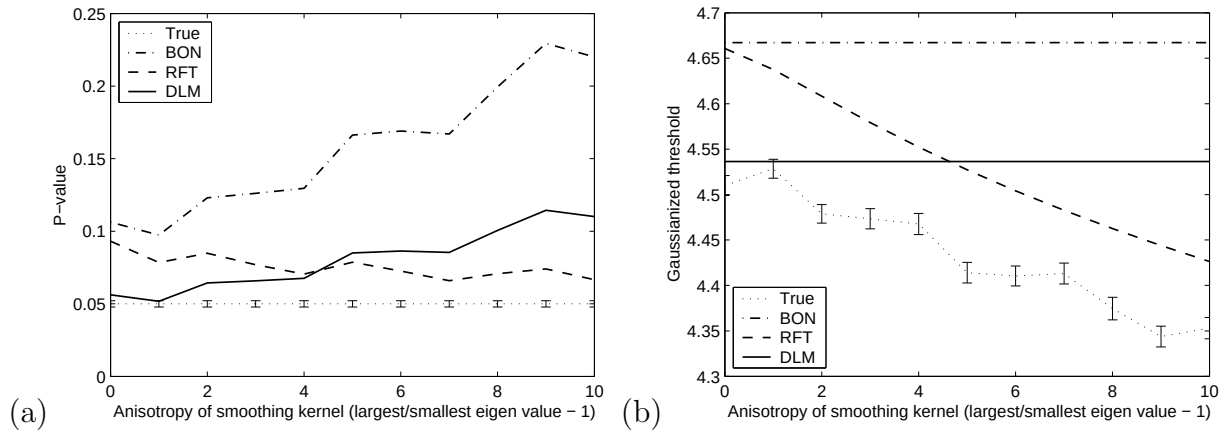

Figure 6: Simulated Gaussian SPM (a) P-value, and (b) threshold as in Figure 1, but for a smoothing kernel with major axis in the diagonal (1,1,1) direction, varying the anisotropy (largest/smallest eigen value - 1) from 0 (isotropic and so axis-aligned) to 10. The DLM appears to be always conservative, supporting the conjecture that DLM is always conservative even if the spatial correlation is not axis-aligned Gaussian.

1, from 0 (isotropic and so axis-aligned) to 10. Specifically, we chose

$$\Lambda = (-2\log\rho)\begin{bmatrix} a+3 & a & a \\ a & a+3 & a \\ a & a & a+3 \end{bmatrix}/(a+3) \tag{18}$$

so that the correlation between adjacent voxels was fixed at ρ (= 0.8572, corresponding to $FWHM = 3$ in Figure 1), and the DLM threshold is always the same. We can see from Figure 6 that the DLM P-value is always greater than the true P-value (held at 0.05), and the DLM threshold is always greater than the true threshold. This supports the conjecture above that DLM is conservative in all cases.

References

Chen J-I, Ha B, Bushnell MC, Pike B, Duncan GH (2002). Differentiating noxious- and innocuous-related activation of human somatosensory cortices using temporal analysis of fMRI. *Journal of Neurophysiology*, **88**:464-474.

Kiebel SJ, Poline J-B, Friston KJ, Holmes AP, Worsley KJ (1999). Robust smoothness estimation in statistical parametric maps using normalised residuals from the general linear model. *NeuroImage*, **10**:756-766.

Nichols T, Hayasaka S (2003). Controlling the familywise error rate in functional neuroimaging; A comparative review. *Statistical Methods in Medical Research*, **12**:419-446.

Taylor JE, Adler RJ (2003). Euler characteristics for Gaussian fields on manifolds. *Annals of Probability*, **31**:533-563.

Taylor JE, Worsley KJ (2005). Maxima of discretely sampled random fields, with an application to 'bubbles'. *Biometrika*, submitted. `http://www.math.mcgill.ca/keith/bubbles/dlm_abstract.htm`

Worsley KJ, Marrett S, Neelin P, Evans AC (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, **12**:900-918.

Worsley KJ (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2, F and t fields. *Advances in Applied Probability*, **26**:13-42.

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, **4**:58-73.

Worsley KJ, Andermann M, Koulis T, MacDonald D, Evans AC (1999). Detecting changes in nonisotropic images. *Human Brain Mapping*, **8**:98-101.

Worsley KJ, Liao C, Aston J, Petre V, Duncan GH, Morales F, Evans AC (2002). A general statistical analysis for fMRI data. *NeuroImage*, **15**:1:15.

Xiong J, Gao JH, Lancaster JL, Fox PT (1995). Clustered pixels analysis for functional MRI activation studies of the human brain. *Human Brain Mapping*, **3**:287-301.