Connectivity of anatomical and functional MRI data

Keith J. Worsley[12], Arnaud Charil[2], Jason Lerch[2], Alan C. Evans[2]
[1]Department of Mathematics and Statistics, and
[2]Montreal Neurological Institute
McGill University
Montreal, Canada
http://www.math.mcgill.ca/keith

***Abstract*— We are all familiar with the correlation coefficient between two sets of numbers. Now suppose we replace the numbers by vector-valued images in any number of dimensions. The correlation random field is the 'image' of correlations at all possible pairs of points in the two images. We use random field theory to set a threshold on the correlations so that those above the threshold are statistically significant, corrected for searching over all pairs of points. We apply this idea to resting state networks of fMRI images of brain activity, and networks of connectivity in cortical thickness.**

I. INTRODUCTION

Suppose we have vector-valued data at each point in an image, for example scalp EEG data at 3 different frequency bands, so that the image space is the 2D manifold of the scalp embedded in 3D. Suppose we also have a second image of vector-valued data, for example fMRI BOLD response at say 4 different time points after presentation of a stimulus. This is repeated several times, so we now have pairs of images, each with vector valued data at every image point. We are interested in finding those pairs of points that are "connected", that is, pairs of points (one on the scalp, one inside the brain) which have high correlation of their vector-valued data. Our main interest is to find a threshold for this correlation that controls the false positive rate of ever finding such connectivity when in fact no connectivity exists.

The measure of correlation that we shall use is the maximum canonical correlation, defined as follows. Let $X(s)$, $s \in S \subset \Re^M$, and $Y(t)$, $t \in T \subset \Re^N$ be matrices of the vector-valued data at image points s, t, with p and q columns and the same number of rows. The vector-valued observations are the rows, and the columns are the repetitions. We assume that nuisance effects such as a constant term have been removed by fitting a common linear model to the data, and the columns of $X(s)$ and $Y(t)$ are the residuals from this linear model. The residual degrees of freedom is ν. Define the maximum canonical correlation random field as

$$C(s,t) = \max_{u,v} \tilde{C}(s,t,u,v)$$

where

$$\tilde{C}(s,t,u,v) = \frac{u'X(s)'Y(t)v}{(u'X(s)'X(s)u \; v'Y(t)'Y(t)v)^{1/2}}.$$

Note that C is the maximum of the canonical correlations between X and Y, defined as the singular values of $(X'X)^{-1/2}X'Y(Y'Y)^{-1/2}$.

We model the data as smooth Gaussian random fields with zero mean. The P-value of the maximum canonical correlation maximized over searching all pairs of points in R, S is well approximated by

$$\mathrm{P}\left(\max_{s,t} C(s,t) \geq c\right) \approx \tfrac{1}{2}\sum_{i=0}^{M}\mu_i(S)\sum_{j=0}^{N}\mu_j(T) \times \sum_{k=0}^{p}\mu_k(U)\sum_{l=0}^{q}\mu_l(V)\rho_{i+k,j+l}(c) \quad (1)$$

where U is the unit sphere in $\Re^p$, V is the unit sphere in $\Re^q$, $\mu_i(S)$ is the i-dimensional *intrinsic volume* of S [1,4] and ρ is the *EC density* of the correlation random field. We now define both these, and give examples.

The intrinsic volume of a set is, roughly speaking, a measure of its i-dimensional content relative to the smoothness of the random field. The intrinsic volume of a search region $S \in \Re^M$ of smoothed image data modelled as Gaussian white noise smoothed with a Gaussian-shaped isotropic filter with Full Width at Half Maximum $FWHM$ is related to the resels by

$$\mu_i(S) = (4\log 2)^{i/2} Resels_i(S).$$

Resels of typical search regions are given in Table I. The intrinsic volume of the sphere U (likewise V) is

$$\mu_k(U) = \frac{2^{k+1}\pi^{\frac{k}{2}}\Gamma\left(\frac{p+1}{2}\right)}{k!\left(\frac{p-1-k}{2}\right)!}$$

if $p-1-k$ is even, and zero otherwise, $k = 0, \ldots, p-1$.

The EC density $\rho_{i,j}$ of the (cross) correlation random field $\tilde{C}$ for fixed u, v is given by

$$\begin{aligned}\rho_{M,N}(c) &= \frac{(M-1)!N!2^{\nu-2-D}}{\pi^{\frac{D}{2}+1}} \\ &\times \sum_{k=0}^{\lfloor (D-1)/2 \rfloor} (-1)^k c^{D-1-2k}(1-c^2)^{\frac{\nu-1-D}{2}+k} \\ &\times \sum_{i=0}^{k}\sum_{j=0}^{k}\frac{\Gamma(\frac{\nu-M}{2}+i)\Gamma(\frac{\nu-N}{2}+j)}{i!j!(k-i-j)!} \\ &\times \frac{1}{(M-1-k-i+j)!(N-k-j+i)!} \\ &\times \frac{1}{(\nu-1-D+i+j+k)!}.\end{aligned}$$

$Resels_i(S)$ M	i 0	1	2	3
0	1			
1	1	$\frac{\text{Length}(S)}{FWHM}$		
2	1	$\frac{\frac{1}{2}\text{ Perimeter length}(S)}{FWHM}$	$\frac{\text{Area}(S)}{FWHM^2}$	
3	1	$\frac{2\text{ Diameter}(S)}{FWHM}$	$\frac{\frac{1}{2}\text{ Surface area}(S)}{FWHM^2}$	$\frac{\text{Volume}(S)}{FWHM^3}$

TABLE I

$Resels_i(S)$ FOR A CONVEX SEARCH REGION S IN M DIMENSIONS. $FWHM$ IS THE EFFECTIVE FULL WIDTH AT HALF MAXIMUM OF A GAUSSIAN KERNEL USED TO SMOOTH THE WHITE NOISE ERRORS IN THE IMAGE DATA X. 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.

where $D = M + N$.

If there is only one set of images, so that $R = S$ and $X = Y$, then we are only interested in auto-correlations at voxels sufficiently far apart to avoid smoothing effects. There are now only half as many possible pairs of correlations, so the resulting P-value from (1) should be halved.

II. Application

We apply the above methods to auto-correlations of functional data (fMRI resting state networks), and anatomical data (cortical thickness).

A. fMRI resting state network

A subject was given a 9s painful heat stimulus, followed by 9s rest, then 9s warm (neutral) stimulus, followed by 9s rest, repeated 10 times, fully described in [3]. 120 frames were acquired at TR=3s; the first 3 were discarded. A linear model was fitted to account for the hot and warm block stimuli (convolved with an HRF), and drift was modelled as a cubic in the acquisition time. The residuals from this linear model, whitened to remove temporal correlation [3], were used for further analysis, leaving $\nu = 111$ null df.

The search region $R = S$ is the brain and we have $p = q = 1$ measurement per voxel. From (1) with $\nu = 111$ null degrees of freedom, the $P = 0.05$ two-sided threshold for C is $c = \pm 0.563$ (since we are interested in both positive and negative correlations). Only points separated by at least 20mm were considered. There were too many correlations above this threshold to display, so Figure 1 shows only 6D local maxima above $c = \pm 0.7$ ($P < 5 \times 10^{-9}$).

There are a large number of right-left correlations between lateral regions in opposite hemispheres, indicating that these regions are synchronised or 'connected' in resting state. There are some long-range correlations on the midline linking large blood vessels. A number of short-range in-slice correlations on the right outer cortex are probably an artefact of imprecise motion correction. Thus would displace the entire cortical boundary in some slices, creating the illusion that boundary voxels are correlated.

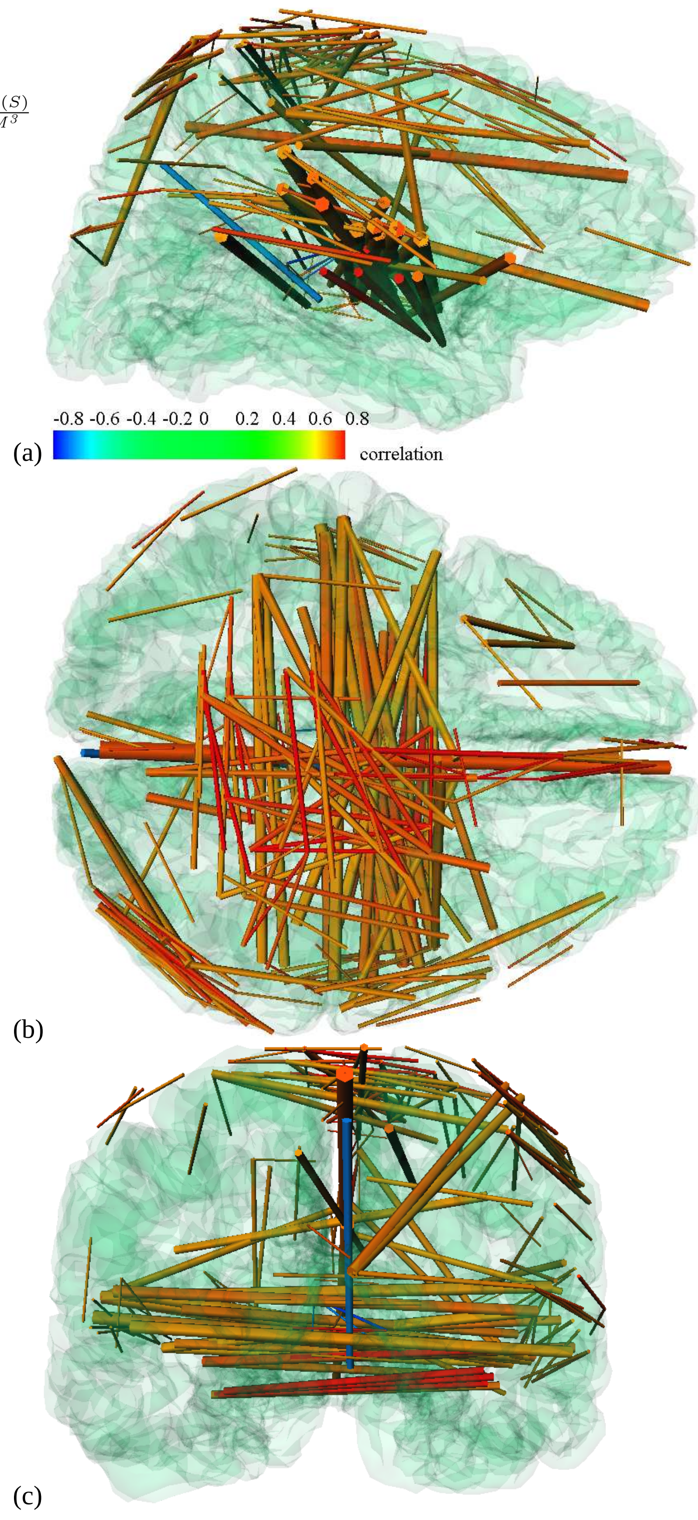


Fig. 1. fMRI resting state network. Inside the mid-cortical surface (transparent), the ends of the rods join voxels where the auto-correlation C of fMRI residuals exceeded $c = \pm 0.7$ (higher than the $P = 0.05$ threshold of $c = \pm 0.563$). Only 6D local maxima are shown. Red rods indicate positively correlated voxels; blue rods indicate negatively correlated voxels (there is only one). (a) side; (b) top; (c) back.

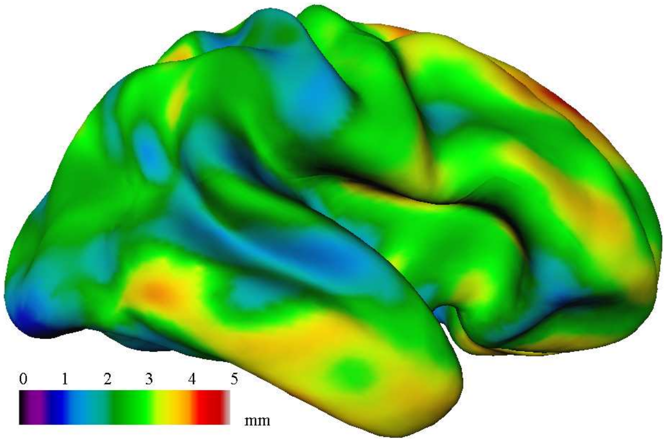
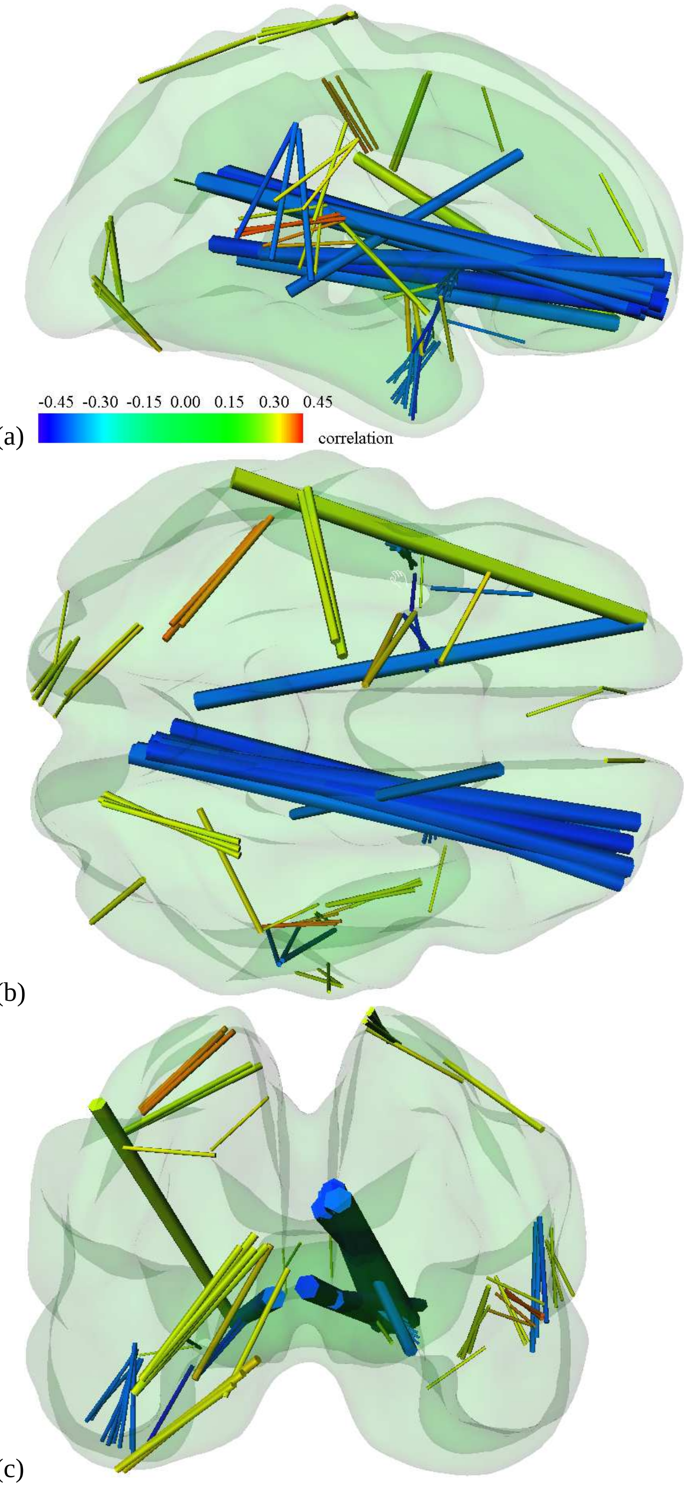


Fig. 2. Cortical thickness of one subject, smoothed by 20mm, plotted on the average of the 321 mid-cortical surfaces.

B. Cortical thickness

We illustrate the method on the cortical thickness of 321 normal adult subjects aged 20-70 years, smoothed by 20mm FWHWM, part of a much larger data set fully described in [2] and also analyzed in [4]. The data on one subject is shown in Figure 2. We removed a linear age, gender and age-gender interaction effects, then calculated the auto-correlation $C(s,t)$, for all pairs of the 40962 triangular mesh nodes, ignoring pairs of nodes that were too close. The search region $R = S$ is the whole cortical surface, with $M = N = 2$, $Resels_0(S) = 2$, $Resels_1(S) = 0$ (since a closed surface has no boundaries), and $Resels_2(S) = 842$. Again there is only one measure at each node, so $p = q = 1$. From (1) with $\nu = 319 - 4 = 317$ null degrees of freedom, the $P = 0.05$ two-sided threshold for C is $c = \pm 0.340$ (since we are interested in both positive and negative correlations). Figure 3 shows only 4D local maxima above c inside the same hemisphere.

The most interesting correlations are the long-range negative correlations linking occipital regions with frontal regions. These suggest that those individuals with thicker occipital cortex have thinner cortex in frontal regions, and vice-versa. There are some interesting positive correlations within the occipital region itself and between right parietal regions and the frontal lobe.



(a)

(b)

(c)

Fig. 3. Connectivity of cortical thickness. The ends of the rods join nodes where the auto-correlation C of cortical thickness exceeded $c = \pm 0.340$ ($P = 0.05$, corrected), removing a linear age, gender and age-gender interaction ($\nu = 321 - 4 = 317$ null degrees of freedom). Only 4D local maxima inside the same hemisphere are shown. Yellow to red rods indicate positively correlated nodes; blue rods indicate negatively correlated nodes. (a) side; (b) top; (c) back.

References

[1] Cao, J. & Worsley, K.J. (1999). The geometry of correlation fields, with an application to functional connectivity of the brain. *Annals of Applied Probability*, **9**:1021-1057.

[2] 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.

[3] Worsley, K.J., Liao, C., Aston, J.A.D., Petre, V., Duncan, G.H., Morales, F. & Evans, A.C. (2002). A general statistical analysis for fMRI data. *NeuroImage*, **15**:1-15.

[4] Worsley, K.J., Taylor, J.E., Tomaiuolo, F. & Lerch, J. (2004). Unified univariate and multivariate random field theory. *NeuroImage*, **23**:S189-195.