

# Characterizing the Response of PET and fMRI Data using Multivariate Linear Models (MLM)

K.J. WORSLEY<sup>1</sup>, J-B. POLINE<sup>2</sup>, K.J. FRISTON<sup>2</sup> AND A.C. EVANS<sup>3</sup>

<sup>1</sup>*Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. West, Montreal, Québec, Canada H3A 2K6,* <sup>2</sup>*The Wellcome Department of Cognitive Neurology, Queen Square, London, UK WC1N 3BG and* <sup>3</sup>*McConnell Brain Imaging Centre, Montreal Neurological Institute, 3801 University Street, Montreal, Québec, Canada H3A 2B4.*

---

This paper presents a new method for characterizing brain responses in both PET and fMRI data. The aim is to capture the correlations between the scans of an experiment and a set of external predictor variables that are thought to affect the scans, such as type, intensity, or shape of stimulus response. Its main feature is a Canonical Variates Analysis (CVA) of the estimated effects of the predictors from a multivariate linear model (MLM). The advantage of this over current methods is that temporal correlations can be incorporated into the model, making the MLM method suitable for fMRI as well as PET data. Moreover, tests for the presence of any correlation, and inference about the number of canonical variates needed to capture that correlation, can be based on standard multivariate statistics, rather than simulations. When applied to an fMRI data set previously analyzed by another CVA method, the MLM method reveals a pattern of responses that is closer to that detected in an earlier non-CVA analysis.

---

**Running title:** CHARACTERIZING THE RESPONSE OF PET AND fMRI DATA

**Addresses of corresponding author:**

e-mail: [worsley@math.mcgill.ca](mailto:worsley@math.mcgill.ca)

web: <http://www.math.mcgill.ca/~keith>

ph: 1-514-398-3842

fax: 1-514-398-3899.

# 1 Introduction

A number of methods derived from Canonical Variates Analysis (CVA) have recently been proposed to analyze brain functional images. Some of these are dedicated to assessing the regional significance of experimentally induced effects while others are meant to characterize the brain response as a global pattern. The aim of methods of characterizing global brain response is two-fold. The first is to provide a sensitive global test of non-focal activation in blood flow images in response to a set of predictor variables (predictors) measured for each scan. The second is to capture, summarize, or explain, the correlation between predictors and voxel blood flows by a small number of *latent* variables, one set for predictors and one set for voxels. These latent variables take the form of *weights*; if the characterization is successful, then the weighted average of the voxel blood flows should be highly correlated with the corresponding weighted average of the predictors, over scans.

Two recent papers have characterized brain responses in PET data (McIntosh *et al.*, 1996) and fMRI data (Friston *et al.*, 1995b) using CVA methods. Both papers have their drawbacks. The inference for the Partial Least Squares (PLS) method of McIntosh *et al.* (1996) is based on a non-parametric permutation test that cannot be applied to temporally correlated data such as fMRI. The method of Friston *et al.* (1995b) uses a heuristic argument to correct for temporal correlation, but this gives a biased test that increases the false positive rate. In this paper we propose an alternative method based on multivariate linear models (MLM) that overcomes these drawbacks, and which can be applied to both PET and fMRI data. The MLM method uses CVA of corrected least squares estimators from a multivariate linear model, but unlike PLS, our inference depends on (parametric) multivariate linear models theory rather than simulations. The MLM method is now implemented in the SPM package.

This paper is structured as follows. Current CVA methods, and the proposed MLM method, are summarized in Section 2. We introduce the multivariate linear model in Section 3, and generalize the results of Worsley & Friston (1995) to testing for the effects of several predictors simultaneously at a single voxel. In Section 4 we define a global test that generalizes that of Worsley *et al.* (1995) to several predictors simultaneously. We then show in Section 5 how to characterize the effect of the predictors using CVA. Section 6 summarizes the MLM procedure, and Section 7 discusses the underlying assumptions. Procedures are validated in Section 8 on both null and simulated signal data, then applied to the same data set as Friston *et al.* (1995b) in Section 9. Technical details are relegated to an extensive Appendix.

## 2 Summary of CVA methods

Generic CVA methods are summarized in Figure A.6, in which  $\mathbf{X}$  and  $\mathbf{Y}$  represent the scans  $\times$  predictors and scans  $\times$  voxel blood flow matrices, respectively, possibly standardized to have zero mean and unit standard deviation. Their complex covariance structure is  $\mathbf{X}'\mathbf{Y}$ .  $\mathbf{U}$  and  $\mathbf{V}$  are the (orthonormal) latent variables  $\times$  predictors and latent variables  $\times$  voxels matrices, respectively. The weighted averages, or predicted and observed temporal responses, are  $\mathbf{X}\mathbf{U}'$  and  $\mathbf{Y}\mathbf{V}'$  respectively. Their simpler covariance structure is the diagonal matrix of

latent roots  $\Lambda$ . Particular methods for doing this are as follows.

## 2.1 SVD (Singular Value Decomposition)

This method was first used for estimating functional connectivity in PET data by Friston *et al.* (1993). It approximates the blood flow matrix  $\mathbf{Y}$  but does not take into account any information in  $\mathbf{X}$ :

$$\mathbf{Y} \approx \mathbf{U}'\Lambda\mathbf{V}.$$

In practice,  $\mathbf{U}$  is first found by a Principal Components Analysis (PCA) of the smaller scans  $\times$  scans matrix  $\mathbf{Y}\mathbf{Y}'$ :

$$\mathbf{Y}\mathbf{Y}' \approx \mathbf{U}'\Lambda^2\mathbf{U},$$

then the larger  $\mathbf{V}$  is found by

$$\mathbf{V} = \Lambda^{-1}\mathbf{U}\mathbf{Y}.$$

Note that the way this approach is implemented in the current SPM package is not in fact independent of  $\mathbf{X}$  since the effects of no interest are removed from the data  $\mathbf{Y}$  before the SVD. It is however of interest to investigate the structure of the covariance matrix of  $\mathbf{X}$  and  $\mathbf{Y}$  (see below). Similar techniques have been used recently for fMRI data (see Bullmore *et al.*, 1996).

## 2.2 PLS (Partial Least Squares)

This method, introduced by McIntosh *et al.* (1996), approximates the covariance matrix of  $\mathbf{X}$  and  $\mathbf{Y}$ :

$$\mathbf{X}'\mathbf{Y} \approx \mathbf{U}'\Lambda\mathbf{V}.$$

However the analysis is not invariant to linear transformations of predictors or voxels. This is partly overcome by making  $\mathbf{X}$  orthogonal; this makes it invariant to rotations of the predictors, but not invariant to arbitrary linear transformations. What it means is that the analysis will be different depending on how the predictors are scaled. For example, doubling a predictor, or choosing different predictors within the same space, will give different answers. Inference about the latent variables is performed non-parametrically using a permutation test. This is valid for PET data, but not fMRI data, since the observations have a temporal correlation structure.

## 2.3 Orthonormalized PLS

We could overcome the lack of invariance noted above by simply orthonormalizing the predictors first, as follows:

$$(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{Y} \approx \mathbf{U}'\Lambda\mathbf{V}.$$

This is now invariant to linear transformations of predictors, but not voxel blood flows, so it combines some of the advantages of PLS and Canonical Variates Analysis (see below). However the problems of temporal correlation and sequential latent root testing still remain.

## 2.4 Standard CVA (Canonical Variates Analysis)

$$(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1/2} \approx \mathbf{U}'\mathbf{\Lambda}\mathbf{V}.$$

Here we orthonormalize both the predictors and blood flow responses. This is invariant to linear transformations of predictors or voxel blood flows, but because there are more voxels than scans, it is not possible to implement, since  $(\mathbf{Y}'\mathbf{Y})^{-1/2}$  cannot be found.

## 2.5 SVD-CVA (Preliminary SVD, then CVA)

Friston *et al.* (1995b, 1996) overcome the problem of more voxels than scans by reducing the number of voxels in a preliminary SVD step:

$$\mathbf{Y} \approx \mathbf{U}^{*'}\mathbf{\Lambda}^*\mathbf{V}^*,$$

replacing  $\mathbf{Y}$  by the reduced set of variables

$$\mathbf{Y}^* = \mathbf{Y}\mathbf{V}^{*'},$$

then finding canonical variates between  $\mathbf{X}$  and  $\mathbf{Y}^*$ :

$$(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{Y}^*(\mathbf{Y}^{*'}\mathbf{Y}^*)^{-1/2} \approx \mathbf{U}'\mathbf{\Lambda}\mathbf{V}.$$

Although the distribution of the data is altered by the preliminary SVD step, it can nevertheless be shown that the usual parametric statistical inference for this procedure is still valid for independent scans such as PET (see Appendix A.1). The method has been successfully employed by Fletcher *et al.* (1996). However for temporally correlated data such as fMRI (Friston *et al.*, 1995b), the usual parametric statistical inference is invalid (see the note added in proof to Worsley & Friston (1995)). Even for PET data, there is some evidence that the proposed MLM method given next is more sensitive (see Appendix A.2).

## 2.6 Proposed MLM (Multivariate Linear Models)

To allow for temporal correlation, denoted by the scans  $\times$  scans matrix  $\mathbf{\Sigma}$ , we shall normalize the predictors not by their sum of squares matrix  $\mathbf{X}'\mathbf{X}$ , but by the matrix  $\mathbf{X}'\mathbf{\Sigma}\mathbf{X}$ , which is the variance of a column of  $\mathbf{X}'\mathbf{Y}$ . This leads to

$$(\mathbf{X}'\mathbf{\Sigma}\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{Y} \approx \mathbf{U}'\mathbf{\Lambda}\mathbf{V}. \quad (2.1)$$

Another way of motivating this is via linear models. If we fit a linear model to the blood flows of the form  $\mathbf{Y} \approx \mathbf{X}\mathbf{B}$ , where  $\mathbf{B}$  is a predictors  $\times$  voxels matrix of least-squares regression coefficients, then it can be shown that

$$(\mathbf{X}'\mathbf{\Sigma}\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{Y} = \text{Var}(\mathbf{B})^{-1/2}\mathbf{B}, \quad (2.2)$$

where  $\text{Var}(\mathbf{B})$  is the variance matrix of any column of  $\mathbf{B}$ , corresponding to a single voxel. Thus our proposed method attempts to approximate the normalised matrix of linear model coefficients by a simpler SVD structure  $\mathbf{U}'\mathbf{\Lambda}\mathbf{V}$ . It is still invariant to linear transformations of predictors, and in addition, inference for this approach is also invariant to the temporal correlation structure. It can therefore be applied to PET or fMRI data.

### 3 A test at a single voxel

The spatial correlation structure of PET and fMRI data is usually well approximated by the convolution of a Gaussian point spread function with itself (Friston *et al.*, 1991; Worsley *et al.*, 1992). The temporal correlation structure of fMRI data can be approximated by the convolution of the hemodynamic response with itself (Friston *et al.*, 1995a) or by more sophisticated methods (Zarahn *et al.*, 1997). An obvious possibility is to deconvolute or whiten the spatial and temporal data to remove these correlations, which leads to the maximum likelihood estimators of the parameters, but this is very sensitive to the correct specification of the correlation structure of the noise. Instead, Friston *et al.* (1995a) suggest using least squares methods to derive the test statistic, then correcting its distribution for the (known) spatial and temporal correlations (Worsley & Friston, 1995). We shall use the same principle here.

Before doing this, one slight generalisation is necessary. As is usual in linear models, we partition the predictors into two subsets: a set of background or nuisance predictors of no interest, such as a constant term, a covariate for global blood flow, or a linear time trend; and a set of predictors of interest, such as stimulus intensity or measured performances, or a subset of these. We usually wish to characterize the effect of the predictors of interest, allowing for the predictors of no interest.

To test for temporal activation at a given voxel  $i$ , let  $\mathbf{b}_i$  be the vector of the least squares estimators of the effects of the subset of the predictors of interest at voxel  $i$ ,  $i = 1, \dots, N$ . Let

$$F_i = \mathbf{b}_i' \text{Var}(\mathbf{b}_i)^{-1} \mathbf{b}_i / h \quad (3.1)$$

be the mean sum of squares of these effects, inversely weighted by their variances, where  $h$  is the number of predictors in the subset (see Appendix A.3). The null distribution of  $F_i$  is well approximated by an  $F$  distribution with  $h$  and  $\nu$  degrees of freedom, where  $\nu$  is the effective temporal degrees of freedom (see Appendix A.4).

The advantage of using the least squares estimator of the effects, as opposed to the maximum likelihood estimator, is that its variance does not involve whitening the  $n$  time points, but only whitening the  $h$  components of  $\mathbf{b}_i$ . This is a far more stable procedure and much less sensitive to miss-specification of the temporal correlation of the observations. The disadvantage is that it may not be as sensitive as using the maximum likelihood estimator, although in some cases (see Worsley & Friston, 1995) there is no loss of sensitivity at all, and only a small loss of sensitivity for a square-wave temporal response generated by a periodic on-off stimulus.

The voxels can be searched for local maxima of  $F_i$  to find the regions most significantly related to the predictors of interest (Büchel *et al.*, 1996). The significance of these local maxima can be assessed using results in Worsley *et al.* (1996) for the maxima of  $F$  fields.

### 4 A global test

A global test statistic for temporal and spatial activation is the average of  $F_i$  across voxels:

$$S = \sum_{i=1}^N F_i / N. \quad (4.1)$$

(This statistic is the likelihood ratio test statistic if there is no spatial or temporal correlation.) If the spatial point spread function is Gaussian, Worsley *et al.* (1995) define the effective spatial degrees of freedom as

$$d = \text{RESELS}(4 \log_e 2/\pi)^{D/2}, \quad (4.2)$$

where  $D$  is the number of dimensions and RESELS is the volume of the search region divided by the product of the effective FWHMs of the experimental point spread function of the observed random noise in each of the  $D$  dimensions (see Worsley *et al.*, 1992, for methods of estimating this). Provided  $\nu$  is large ( $> 10$ ) then the null distribution of  $S$  is well approximated by a multiple of an  $F$  distribution, as follows. Let

$$\nu_1 = dh, \quad \nu_2 = d\nu - \frac{(d-1)(4h+2\nu)}{h+2}, \quad F = \frac{\nu-2}{\nu} \frac{\nu_2}{\nu_2-2} S. \quad (4.3)$$

Then the null distribution of  $F$  is well approximated by an  $F$ -distribution with  $\nu_1$  and  $\nu_2$  degrees of freedom. The details of this derivation are given in Appendix A.5, and it is validated in Appendix A.2.

If there is just one predictor ( $h = 1$ ), such as a difference between two tasks, or a difference between early and late responses to the same task (Friston *et al.*, 1995c), then it can be shown that  $F_i$  is identical to the square of the  $t$ -statistic for testing for this effect. Moreover  $S$  is then identical to the mean sum of squares of the  $t$ -statistic map, a test statistic already proposed by Worsley *et al.* (1995) to test for the global effect of a linear combination of predictors. In this way the methods proposed here can be seen as straightforward generalizations to the case of more than one predictor.

## 5 Characterizing the response

As described in Section 2, the aim of CVA is to find what (global) linear combination of the predictors best describes the response across the voxels. One way of interpreting this is to find which linear combination of the effects of the predictors has the greatest variability. This suggests a Principal Components Analysis of the normalized effects, defined as

$$\mathbf{Z}_i = \text{Var}(\mathbf{b}_i)^{-1/2} \mathbf{b}_i. \quad (5.1)$$

Note that  $F_i = \mathbf{Z}_i' \mathbf{Z}_i / h$ . If we ignore the unequal voxel variances then  $\mathbf{Z}_i$  is the same as the  $i$ th column of  $\text{Var}(\mathbf{B})^{-1/2} \mathbf{B}$  from (2.2). The mean sums of squares and cross-products matrix is then

$$\mathbf{S} = \sum_{i=1}^N \mathbf{Z}_i \mathbf{Z}_i' / N. \quad (5.2)$$

The eigenvalues of  $\mathbf{S}$  are  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_h$ , and the corresponding eigenvectors (principal components) define the linear combinations of the normalized effects that capture the greatest variability; the variability captured by the  $j$ th principal component is  $\lambda_j$ . From these we can find the corresponding temporal and spatial effects (see Appendix A.3).

For pure noise data, the components of  $\mathbf{Z}_i$  are all uncorrelated standard Gaussian random variables, provided the effective temporal degrees of freedom is large. If a temporal signal is

present at position  $i$  and can be represented as a linear combination of the predictors, then the mean of  $\mathbf{Z}_i$  will be non-zero. If the same temporal signal is spatially distributed across voxels then it is consistently estimated by the first principal component. If there are  $q$  such signals then they are estimated by the spatial and temporal responses associated with the first  $q$  components.

We may ask the question: Do these first  $q$  principal components capture *all* the signal, or is there still some signal left in the remaining  $h - q$  principal components? There are several ways of testing such a hypothesis. The standard method suggested in the literature is to use the ratio of the arithmetic to the geometric mean of  $\lambda_{q+1}, \dots, \lambda_h$ , also known as the sphericity test (see Anderson, 1984, page 475). If the last  $h - q$  population eigenvalues are equal, indicating no outstanding principal components after the first  $q$ , this ratio should be close to one. However this test is designed for data with unknown voxel variance  $\sigma_i^2$ , and it is designed to detect *unequal* eigenvalues. In our case the variance of the data is normalized to a standard deviation of one at each voxel by dividing by an estimator  $\hat{\sigma}_i^2$ , so the null population eigenvalues are one. Moreover we are more interested in detecting eigenvalues greater than one, rather than unequal eigenvalues. For these reasons we suggest a slightly different test.

One possibility is to use the average of the last  $h - q$  principal components, a test similar to the Lawley-Hotelling trace (Anderson, 1984, page 323):

$$S_q = \sum_{j=q+1}^h \lambda_j / (h - q). \quad (5.3)$$

(Again it can be shown that, provided there is no spatial or temporal correlation, this is the likelihood ratio test statistic for a simultaneous test of the last  $h - q$  population eigenvalues equal to a constant greater than one.) There is a simple approximate form for its null distribution: provided  $\nu$  is large ( $> 10$ ) and the signal in the first  $q$  principal components is large, then  $S_q$  has the same distribution as  $S$  but with  $h$  replaced by  $h - q$  (the proof of this follows the arguments on page 476 of Anderson (1984)). Let

$$\nu_{1q} = d(h - q), \quad \nu_{2q} = d\nu - \frac{(d - 1)(4(h - q) + 2\nu)}{h - q + 2}, \quad F_q = \frac{\nu - 2}{\nu} \frac{\nu_{2q}}{\nu_{2q} - 2} S_q. \quad (5.4)$$

Then the null distribution of  $F_q$  is well approximated by an  $F$ -distribution with  $\nu_{1q}$  and  $\nu_{2q}$  degrees of freedom. The case of  $q = 0$  (some signal in all the principal components) is of course identical to the global test based on  $S = S_0 = \sum_j \lambda_j / h = \text{trace}(\mathbf{S}) / h$ .

A second possibility, similar in spirit to Roy's maximum root (Anderson, 1984, page 328), is based on just the largest of the last  $h - q$  eigenvalues,  $\lambda_{q+1}$ . This is the likelihood ratio test statistic for a single non-unity population eigenvalue in the last  $h - q$  population eigenvalues. It may be more sensitive to a single principal component than  $S_q$ , but it has the drawback that there is no simple formula for its null distribution, and tables are available only for  $h - q \leq 6$ . For these reasons we prefer  $S_q$ .

## 6 Summary of procedure

The following summary of the procedure has been implemented in SPM:

**Step 1: Voxel tests.** Calculate the voxel  $F$ -statistic  $F_i$  for testing for an effect of the predictors on blood flow at voxel  $i$ .

**Step 2: Global test.** Test for a global effect using  $S$ , the average of the voxel  $F$ -statistics  $F_i$ . If this is not significant, then stop and conclude that the predictors have no effect on the response. If it is significant, then move to Step 3.

**Step 3: Characterizing the response.** The response can be characterized by a principal components analysis of the normalized effects of the predictors, as follows:

1. Test for more than 1 component using  $S_1$ ; if this is significant:
2. Test for more than 2 components using  $S_2$ ; if this is significant:
- .
- .
- $q$ . Test for more than  $q$  components using  $S_q$ ; if this is not significant, then stop.

We can then use the first  $q$  principal components as an adequate description of the temporal and spatial responses in the data.

As an optional procedure in the Step 1, not necessary for subsequent steps,  $F_i$  can be searched over voxels for local maxima that can then be tested for their significance (Büchel *et al.*, 1996) using the unified  $p$ -value for  $F$ -fields from Worsley *et al.* (1996). Note that localized focal signals may be detected by local maxima but not by  $S$ , whereas distributed non-focal signals may be detected by  $S$  but not by local maxima; no test is best at detecting all types of signal.

## 7 Assumptions

The above theory is based on several parametric assumptions, which are listed below. Note that if any of these assumptions are in doubt, then PET data (but not fMRI data) can always be analyzed by PLS, which is non-parametric. The assumptions are:

1. The linear model is correctly specified. Strong coherent signals that are not linear combinations of the predictors may go undetected. If there is some doubt about the model, extra predictors could be added to ensure that the signal is correctly modeled.
2. The temporal correlation is correctly specified. Friston *et al.* (1995a) suggest estimating this by convolving the hemodynamic response function with itself. Note that the lag need not be specified, only the shape. More sophisticated methods are given by Zarahn *et al.* (1997), who model the power spectrum of fMRI data directly. As discussed in Section 3, the type of analysis has been chosen to make it robust against miss-specification of the temporal correlation structure.



3. The noise component of the images can be modeled as a stationary Gaussian random field. Departures from this assumption are not too serious because the Central Limit Theorem guarantees that estimated effects are approximately Gaussian even if the individual scans are not Gaussian (see Worsley *et al.* , 1992, for a discussion of these issues).
4. The point spread function is Gaussian (only needed for the calculation of the effective spatial degrees of freedom,  $d$ , from (4.2)). This is not too problematic, since PET and fMRI data is often smoothed with a Gaussian filter prior to analysis to increase the signal-to-noise ratio.
5. Large number of scans relative to predictors, so that the effective temporal degrees of freedom  $\nu$  is large ( $\nu > 10$ ). This is needed to ensure that  $F$  has the appropriate  $F$ -distribution, so that the false positive rate for the global test is correctly specified. This is usually not a problem since  $\nu > 30$  is common for most PET and fMRI data sets. This is validated in Appendix A.2, along with a comparison of its sensitivity relative to the SVD-CVA method, but assuming that there is no spatial or temporal correlation.
6. If present, the signal must be large enough to ensure that the sequential tests based on  $F_q$  used to characterize the response are valid, so that the method does not pick up more signal components than are actually present. This will be validated in Section 8 for temporally and spatially correlated data, but assuming that  $\nu = \infty$ .

Note that sequential testing of latent roots is a standard statistical procedure, but it must be used with caution. Suppose a  $p$ -value of  $\alpha$  is used for all sequential tests. If no signal is present, it always protects against false positives with probability  $\alpha$ , since the procedure stops before the first stage with a probability of  $1 - \alpha$  irrespective of what might happen after that. If signal is present in a finite number of components, and it is strong enough, then the same argument applies and no extra false components are detected (with probability  $1 - \alpha$ ). If the signal is weak, it may not be detected at earlier stages but this still protects against finding more components than are truly present.

## 8 Validation

First of all, we checked that if no signal was present, then no signal components were detected (with the pre-specified false positive rate). Second, we checked that if one signal component was added then it was reliably detected, and no further components were detected (with the pre-specified false positive rate). Since the requirement that  $\nu > 10$  is validated at a single voxel in Section A.2, then we carried out the validation of the sequential tests with  $\nu = \infty$ , which is equivalent to the case of known  $\sigma_i^2$ . This reduced the amount of simulation by a factor of about 50 by doing all the simulations in Fourier space (see Appendix A.6).

As far as possible, we tried to duplicate the experiment of Friston *et al.* (1995b). We generated 100 null data sets by smoothing 3D Gaussian white noise with a 10mm FWHM Gaussian spatial filter and a 6.65 second FWHM Gaussian hemodynamic response function.

The data was sampled in a  $30 \times 35 \times 10$  grid of  $3 \times 3 \times 6$ mm voxels ( $N = 10500$ , RESELS=567) and  $n = 120$  time intervals of 3 seconds duration, divided into 4 repeats of 3 different tasks, each of 10 scans duration. A model with  $h = 12$  predictors of interest, identical to those used by Friston *et al.* (1995b), was used for  $\mathbf{X}$ . These were intended to capture a differential response (4 sine terms) between the 3 different tasks (3 indicator variables), giving a total of 12 predictors (the first 32 rows are shown as the matrix  $\mathbf{X}$  in Figure A.6). Note that the predictors cover 12 scans to allow for a 2 scan overlap with the next task. A set of 7 nuisance variables corresponding to a constant term and 6 low-frequency Fourier components, was used for the predictors of no interest. A temporal signal with a 20% RMS signal to noise ratio was created by duplicating the biphasic temporal response observed in Friston *et al.* (1995b). The amplitude was varied spatially throughout the brain by multiplying the temporal signal by a white noise image smoothed with the same spatial filter as above. This was added to each null data set to produce a simulated signal data set with one principal component.

Both null and simulated signal data sets were analyzed as above. Figure A.6 shows a histogram of  $S$ , calculated as outlined in Appendix A.3. Superimposed on this is the theoretical scaled  $F$ -distribution (4.3) for  $S$ . The effective spatial degrees of freedom is  $d = 470$  and the effective temporal degrees of freedom is  $\nu = 35.6$ . However the simulations were carried out for known  $\sigma_i^2$ , so that  $\nu = \infty$  and  $\nu_2 = \infty$  were used to calculate the null distribution of  $S$ . The fit to the null data appears to be reasonable. The  $P = 0.05$  threshold of the null distribution of  $S$  is 1.036, and the proportion of simulated  $S$  values exceeding this is  $0.04 \pm 0.07$  for the null data and all of them for the simulated signal data.

Figure A.6ab shows the average eigenvalues together with the true eigenvalues (1 for the null data, and 2.9 for the first eigenvalue of the simulated signal data). Figure A.6ab shows the  $p$ -values of  $S_q$  found using the scaled  $F$ -distribution (5.4). For the null data, we conclude that there are no significant components in 96% of the cases, and one (false) component in 4% of the cases. For the simulated signal data, we conclude that there is just one significant signal component in 98% of the cases, and a second (false) component in 2% of the cases. Figure A.6ab shows the predicted temporal response of the first principal component for 6 randomly chosen data sets. For the simulated signal data the predicted temporal response slightly overestimates the true response that was added to the data. Note that the sign of the predicted temporal response is arbitrary, so in each case the sign was chosen to best match the true signal. For the null data the predicted response is a random combination of the basis functions used to describe the differential responses.

These simulations were close to the expected, which tells us that the methods do appear to work as claimed even when a single moderate strength signal component is present. Such a signal is reliably detected, and the false positive rate of finding more signal components is controlled at the specified level.

## 9 Application

The above methods were applied to the same data set analyzed by Friston *et al.* (1995b,c). In this fMRI activation experiment, there were three conditions: a rest condition (“rest”), a motor sequencing condition where the subject moved his or her right and left hand in a fixed alternating order in response to a visual cue (“fixed”), and a motor sequencing condition in

which the subject moved either the right or the left hand as instructed visually in a random sequence (“random”). In the random condition the subject could not anticipate or prepare the exact movement before seeing the cue.

The global test statistic is  $S = 1.81$ , which is highly significant ( $P < 0.0001$ ). Inspection of the latent roots (Figure A.6c) shows that just one component is significant (Figure A.6c), as in the simulated signal data. Note that the spread of the remaining 11 latent roots in Figure A.6c is greater than that for the simulated data. This suggests that there is some weak signal still present in the higher components that has gone undetected. The spatial response (Figure A.6a) shows several foci in the right anterior cortex. The predicted temporal response, shown in Figures A.6c and A.6b, suggests a delayed response for the random condition but only weak responses for the other two conditions.

This is quite different from that reported by (Friston *et al.*, 1995b, Figure 5) but very similar to that reported in (Friston *et al.*, 1995c, Figures 4 and 5) (reproduced here in Figure A.6d). In the latter paper, the authors used a simple univariate analysis with predictors tailored to detect the differential response suggested by the multivariate analysis. A significant differential response was detected in the anterior cingulate gyrus (A.6d). (Note that the reported  $\text{SPM}\{Z\}$  values are incorrect; see the note added in proof to Worsley & Friston (1995).) The resemblance between this temporal response and that from our analysis is quite striking, particularly for the random condition. Moreover the spatial responses of our method match quite closely those of the univariate analysis; both are compatible with activation in the anterior cingulate. In contrast, the spatial response reported in Friston *et al.* (1995b) shows a diffuse response with no clear foci.

Formally, these localized interpretations are difficult to test; our procedure only indicates that some signal is present in the temporal and spatial responses, without specifying where. However, these spatial patterns should be relevant to the set of subjects’ cerebral regions involved in the experimental paradigm in the sense that they best represent the variability of the covariance between the acquired data and the predictors.

## 10 Conclusion

We have proposed a parametric method that decomposes the correlation between the data and a set of predictors in spatio-temporal patterns (spatial and temporal responses). These components best represent the correlation structure in a parsimonious way. This method extends Partial Least Squares by accounting for the temporal structure of the data and allows inference on the successive components. It also extends omnibus tests for one predictor to a set of predictors under the assumptions described above.

An important issue is functional versus effective connectivity. Although canonical variate analysis and related methods do not directly inform on the effective connectivity (the influence of one region on another), their output can be used to generate hypotheses on the effect or modulation of the activity of one part of the brain depending on the response observed in other locations. Of particular interest might be the comparison of first and second or subsequent components bearing in mind that these are constructed such that they are uncorrelated to each other. In this respect, unpredicted spatio-temporal patterns with meaningful neurophysiological (or possibly artefactual) interpretations can be observed in

second, third or further components. The Multivariate Linear Model (MLM) technique permits the formal testing of the signal left in these components while removing the variation explained by the previous patterns.

We expect the proposed multivariate analyses to have more and more importance in the future for the following reason: in many occasions, brain responses will be modeled by a series of functions (regressors or predictors) for which no specific contrasts can be formed, and for which the parameter of interest is the overall variance explained by this set of functions. While the  $F$ -statistic map provides position specific information on the “significance” and the shape of the observed signal, the MLM method extracts a set of unique and global functions which inform on the commonalities of the brain temporal and spatial hemodynamic responses to a stimulus. Obviously, observed patterns will be dependent on the space spanned by the set of functions and it is likely that variations found in the results for different spaces could inform on the complexity or the nature of the brain functions. This remain to be investigated and relates to model selection techniques for which MLM could be used to give an overall assessment of the significance of a set of regressors.

## A Appendix

### A.1 Validity of SVD-CVA for PET data

The preliminary SVD step of Friston *et al.* (1995b, 1996) induces correlations between the scans (in particular the sum of squares of scan measurements are fixed at 1) so that the assumptions for validity of the usual multivariate statistical methods do not hold. Nevertheless, we shall show that results still remain valid, provided the scans are independent, so that the method can be used for PET data, as for example in Fletcher *et al.* (1996).

The argument is as follows. It is not hard to show that canonical correlations for any pair  $\mathbf{X}, \mathbf{Y}^*$  are invariant under any rotation of the scans, that is, pre-multiplication by an orthonormal matrix  $\mathbf{R}$ . Moreover the canonical correlations (and hence Wilk’s  $\Lambda$ ) depend on  $\mathbf{Y}^*$  only through the standardized observations,  $\mathbf{W} = \mathbf{Y}^*(\mathbf{Y}^{*'}\mathbf{Y}^*)^{-1/2}$ . In fact the null distribution of  $\mathbf{W}$  is characterized by the fact that  $\mathbf{W}'\mathbf{W} = \mathbf{I}$ , the identity matrix, and that the distribution of  $\mathbf{W}$  is invariant under rotations. These distributional assumptions are met by the temporal SVD components  $\mathbf{Y}^* = \mathbf{Y}\mathbf{V}'$  of  $\mathbf{Y}$ , provided the scans are independent and normally distributed, for the following reason. The temporal SVD components of  $\mathbf{R}\mathbf{Y}$  are  $\mathbf{R}\mathbf{Y}^*$ , giving standardized observations  $\mathbf{R}\mathbf{W}$ . The multivariate normal distribution of  $\mathbf{R}\mathbf{Y}$  is the same as that of  $\mathbf{Y}$ , hence the distribution of  $\mathbf{R}\mathbf{W}$  is the same as that of  $\mathbf{W}$ . The distribution of  $\mathbf{W}$  is thus rotationally invariant and  $(\mathbf{R}\mathbf{W})'\mathbf{R}\mathbf{W} = \mathbf{I}$ , guaranteeing that the canonical correlations of  $\mathbf{X}, \mathbf{Y}^*$  have the same distribution as if the rows of  $\mathbf{Y}^*$  were independent normal. Note that this argument breaks down when the scans are correlated as in fMRI data, because then the multivariate normal distribution of  $\mathbf{R}\mathbf{Y}$  is not the same as that of  $\mathbf{Y}$ .

## A.2 Sensitivity of SVD-CVA for PET data

The idea of reducing voxels by a singular value decomposition, as proposed by Friston *et al.* (1995b, 1996), appears at first to be very sensitive, but in this section we shall provide some evidence that our proposed method based on  $F$ -statistics is more sensitive at detecting both a distributed signal and a focal signal.

To simplify the discussion, suppose there is a single temporal predictor, and a single underlying temporal response to it that is expressed by different strengths in different brain regions. The first temporal SVD component should pick out these responses from the background noise, without any input from the predictor.

Friston *et al.* (1995b, 1996) then propose to test for the temporal response by regressing the first temporal SVD component  $\mathbf{Y}^*$  on the temporal predictor and using a standard multivariate test statistic, Wilk's  $\Lambda$ . In this case, Wilk's  $\Lambda$  becomes a simple  $t$ -test for correlation between the SVD component and the predictor (we shall refer to this as the SVD  $t$ -test). Even though the distribution of the SVD component is not Gaussian, it can be shown that the  $t$ -statistic does have the usual null  $t$ -distribution as if the observations were Gaussian, provided there is no temporal correlation (see Appendix A.1). This remarkable result allows us to set the specificity of the test accurately, and at the same time to enhance the signal beforehand, without seemingly sacrificing anything.

However our simulations provide some evidence that a straightforward  $F$ -test is more sensitive. The  $F$ -test is based on the sum of squares of the separate  $t$ -statistics, one for each voxel, obtained by regressing the temporal predictor on the raw temporal voxel data, without first extracting the temporal SVD component. We simulated Gaussian white noise at 100 independent voxels (roughly equal to the number of RESELS in 20mm smoothed data) and 10 independent time points (roughly equal to the number of temporal RESELS), and added two types of signal: a distributed signal generated from 100 Gaussian random variables, and a focal signal covering 5% of the voxels. In both cases the signal to noise ratio was 0%, 10% and 20%. Figure A.6 shows the sensitivity from 1000 simulations as a function of false positive rate.

The specificity of both tests is well controlled: at 0% signal the SVD  $t$ -statistic has a  $t$ -distribution with 9 degrees of freedom; the  $F$ -statistic has an approximate scaled  $F$ -distribution with 100 and 174 degrees of freedom (see (4.3)). However in every case the  $F$ -statistic is more sensitive by 10-50%, especially for the focal signal where we might have thought the SVD  $t$ -statistic would do better.

It should be mentioned that the SVD-CVA method has the advantage that it does not depend on the assumed spatial correlation structure, whereas the effective degrees of freedom of the  $F$ -statistic (4.2) does depend on an assumed Gaussian-shaped point response function. Since fMRI data is often smoothed prior to analysis with a Gaussian-shaped filter, this assumption seems reasonable.

To conclude: the SVD-CVA approach is only valid for independent scans, but it makes no assumption about the spatial correlation structure. Even when conditions are favorable for the SVD-CVA approach (no temporal correlation), a simple  $F$ -test appears to do better.

### A.3 Matrix details

$F_i$  can be written in matrix notation as follows; these results can be found in Worsley & Friston (1995), but they are repeated here for clarity. Let  $\mathbf{Y}_i$  be the column vector of scan measurements at voxel  $i$ ,  $i = 1, \dots, N$ . The linear model at voxel  $i$  is

$$\mathbf{Y}_i = \mathbf{X}\beta_i + \mathbf{G}\gamma_i + \epsilon_i, \quad (\text{A.1})$$

where  $\mathbf{X}$  is the design matrix of the subset of predictors of interest, whose rows are the scans and whose columns are the predictors of interest, possibly filtered by the hemodynamic response function;  $\mathbf{G}$  is a similar design matrix for the predictors of no interest such as a constant term and a linear temporal drift. The error vector  $\epsilon_i$  has a Gaussian distribution with mean zero and variance-covariance matrix  $\Sigma\sigma_i^2$ , where  $\sigma_i^2$  is an unknown scalar and the scans $\times$ scans matrix  $\Sigma$  can be modeled as the convolution of the hemodynamic response function with itself (Friston *et al.*, 1995a) or by more sophisticated models (Zarahn *et al.*, 1997). The complete design matrix  $\mathbf{D} = [\mathbf{X}\mathbf{G}]$  is assumed to have full rank. Then the least squares estimator of the effects  $\beta_i$  of the predictors of interest, allowing for the effects  $\gamma_i$  of the predictors of no interest, is

$$\mathbf{b}_i = (\mathbf{X}'_G \mathbf{X}_G)^{-1} \mathbf{X}'_G \mathbf{Y}_{Gi},$$

where

$$\begin{aligned} \mathbf{X}_G &= \mathbf{X} - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{X}, \\ \mathbf{Y}_{Gi} &= \mathbf{Y}_i - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{Y}_i. \end{aligned}$$

Let  $\mathbf{R} = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$ , where  $\mathbf{I}$  is the scans $\times$ scans identity matrix. Then the residuals are  $\mathbf{r}_i = \mathbf{R}\mathbf{Y}_i$  and an unbiased estimator of  $\sigma_i^2$  is

$$\hat{\sigma}_i^2 = \mathbf{r}'_i \mathbf{r}_i / \text{trace}(\mathbf{R}\Sigma).$$

Hence an unbiased estimator of the variance of  $\mathbf{b}_i$  is

$$\text{Var}(\mathbf{b}_i) = (\mathbf{X}'_G \mathbf{X}_G)^{-1} \mathbf{X}'_G \Sigma \mathbf{X}_G (\mathbf{X}'_G \mathbf{X}_G)^{-1} \hat{\sigma}_i^2.$$

If we define the mean regression sum of squares of the normalized predictors of interest as

$$H_i = \mathbf{Y}'_{Gi} \mathbf{X}_G (\mathbf{X}'_G \Sigma \mathbf{X}_G)^{-1} \mathbf{X}_G' \mathbf{Y}_{Gi} / h,$$

then

$$F_i = \mathbf{b}'_i \text{Var}(\mathbf{b}_i)^{-1} \mathbf{b}_i = H_i / \hat{\sigma}_i^2.$$

The normalized effects can be calculated as follows. Let

$$\mathbf{M} = \mathbf{X}'_G \Sigma \mathbf{X}_G.$$

Then  $\mathbf{M}^{-1/2}$  can be chosen to be the transpose of the inverse of the Cholesky factor of  $\mathbf{M}$ , so that  $\mathbf{M}^{-1/2} \mathbf{M} (\mathbf{M}^{-1/2})'$  is the identity matrix. This then defines

$$\mathbf{Z}_i = \mathbf{M}^{-1/2} \mathbf{X}'_G \mathbf{Y}_{Gi} / \hat{\sigma}_i.$$

The mean sums of squares and cross-products matrix is then

$$\mathbf{S} = \sum_{i=1}^N \mathbf{Z}_i \mathbf{Z}_i' / N.$$

The eigenvalues of  $\mathbf{S}$  are  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_h$  which form the diagonal elements of  $\mathbf{\Lambda}$ , and the corresponding eigenvectors (principal components) are  $\mathbf{u}_1 \dots \mathbf{u}_h$ , which are analogous to the rows of  $\mathbf{U}$  in Figure A.6. Note that the signs of these principal components are arbitrary. The spatial response corresponding to the  $j$ th principal component at voxel  $i$ , analogous to  $\mathbf{V}$  in Figure A.6, is

$$v_{ij} = \mathbf{Z}_i' \mathbf{u}_j / \sqrt{\lambda_j}.$$

The corresponding observed temporal response removing the effect of  $\mathbf{G}$ , analogous to  $\mathbf{YV}'$  of Figure A.6, is

$$\mathbf{y}_j = \sum_{i=1}^N v_{ij} \mathbf{Y}_{\mathbf{G}i} / (\hat{\sigma}_i N).$$

Finally, the corresponding predicted temporal response removing the effect of  $\mathbf{G}$ , analogous to  $\mathbf{XU}'$  in Figure A.6, is the predicted values from a regression of  $\mathbf{y}_j$  on  $\mathbf{X}_{\mathbf{G}}$ :

$$\hat{\mathbf{y}}_j = \mathbf{X}_{\mathbf{G}} (\mathbf{X}_{\mathbf{G}}' \mathbf{X}_{\mathbf{G}})^{-1} \mathbf{X}_{\mathbf{G}}' \mathbf{y}_j = \sqrt{\lambda_j} \mathbf{X}_{\mathbf{G}} (\mathbf{X}_{\mathbf{G}}' \mathbf{X}_{\mathbf{G}})^{-1} \mathbf{M}^{1/2} \mathbf{u}_j.$$

The scaling of the observed and predicted temporal responses are arbitrary, but they were chosen in the following way. Suppose  $\beta_i / \sigma_i = \tilde{\beta} \tilde{v}_i$ , where  $\sum_i \tilde{v}_i^2 / N = 1$ , that is the effect of interest divided by the standard deviation of the observations is proportional to a single common effect  $\tilde{\beta}$ , with the voxel RMS of the constant of proportionality equal to 1. Suppose also that the error  $\epsilon_i$  in the linear model (A.1) is zero. Then it can be shown that there is only one response ( $q = 1$ ), the spatial response is  $v_{i1} = \tilde{v}_i$ , and the observed and predicted temporal responses are  $\mathbf{y}_1 = \hat{\mathbf{y}}_1 = \mathbf{X}_{\mathbf{G}} \tilde{\beta}$ . Thus 1 on the scale of the observed and predicted temporal responses corresponds to  $\sigma_i$  on the scale of the observations.

## A.4 Temporal degrees of freedom

The exact degrees of freedom of the mean regression sum of squares  $H_i$  is  $h$ , and the effective (temporal) degrees of freedom of  $\hat{\sigma}_i^2$  is

$$\nu = \text{trace}(\mathbf{R}\mathbf{\Sigma})^2 / \text{trace}(\mathbf{R}\mathbf{\Sigma}\mathbf{R}\mathbf{\Sigma}).$$

However  $H_i$  is not independent of  $\hat{\sigma}_i^2$ , even assuming that  $\nu \hat{\sigma}_i^2 / \sigma_i^2$  has a  $\chi^2$  distribution with  $\nu$  degrees of freedom, so that the distribution of  $F_i$  is not quite  $F_{h,\nu}$ . It can be shown that the correlation between  $H_i$  and  $\hat{\sigma}_i^2$  is

$$\rho = \frac{\text{trace}(\mathbf{X}_{\mathbf{G}} (\mathbf{X}_{\mathbf{G}}' \mathbf{\Sigma} \mathbf{X}_{\mathbf{G}})^{-1} \mathbf{X}_{\mathbf{G}}' \mathbf{\Sigma} \mathbf{R} \mathbf{\Sigma})}{\sqrt{h} \text{trace}(\mathbf{R}\mathbf{\Sigma}\mathbf{R}\mathbf{\Sigma})}. \quad (\text{A.2})$$

It is also possible to show that this correlation is small but always positive. The effect of positive correlation between the numerator and denominator of the  $F$  ratio is to slightly increase

the effective numerator and denominator degrees of freedom (by factors of  $1/(1 - \rho\sqrt{h/\nu})$  and  $1/(1 - \rho\sqrt{\nu/h})$  respectively, see (A.5)), resulting in slightly larger critical thresholds. Thus using  $h$  and  $\nu$  as degrees of freedom will result in a slightly conservative test. For the example used in this article,  $\rho = 0.018$  and the effective numerator and denominator degrees of freedom are increased by factors of 1.005 and 1.015, respectively; clearly these are negligible.

The above factors were obtained by the following argument. Suppose  $V$  and  $W$  are two positive random variables with unit expectation, and our goal is to choose the degrees of freedom of an  $F$  distribution to match the first two moments of  $V/W$ ; in our case  $V = H_i$  and  $W = \hat{\sigma}_i^2$ . Using a Taylor series expansion about  $V = 1$  and  $W = 1$ , and taking expectations, it can be shown that

$$\begin{aligned} E(V/W) &\approx 1 + \text{Var}(W) - \text{Cov}(V, W), \\ \text{Var}(V/W) &\approx \text{Var}(V) - 2\text{Cov}(V, W) + \text{Var}(W), \end{aligned} \quad (\text{A.3})$$

plus terms in third order moments of  $(V, W)$  and higher. If  $F$  has an  $F$  distribution with  $v$  and  $w$  degrees of freedom then

$$\begin{aligned} E(F) &\approx 1 + 2/w, \\ \text{Var}(F) &\approx 2/v + 2/w, \end{aligned} \quad (\text{A.4})$$

plus quadratic terms in  $(1/v, 1/w)$  and higher. Equating the right hand sides of (A.3) and (A.4) implies that

$$\begin{aligned} v &= 2/(\text{Var}(V) - \text{Cov}(V, W)) = v^*/(1 - \rho\sqrt{v^*/w^*}), \\ w &= 2/(\text{Var}(W) - \text{Cov}(V, W)) = w^*/(1 - \rho\sqrt{w^*/v^*}), \end{aligned} \quad (\text{A.5})$$

where  $v^* = 2/\text{Var}(V) = h$ ,  $w^* = 2/\text{Var}(W) = \nu$  and  $\rho = \text{Cor}(V, W)$ , given by (A.2).

## A.5 Spatial degrees of freedom

We now turn to the spatial correlation structure of  $F_i$ . It is straightforward to show that any linear combination of stationary Gaussian random fields with identical spatial correlation function  $\rho(\mathbf{x})$  has the same spatial correlation function  $\rho(\mathbf{x})$ . From this it can be shown that the spatial correlation function of any quadratic form in these Gaussian random fields, such as  $H_i$  or  $\hat{\sigma}_i^2$ , is equal to  $\rho(\mathbf{x})^2$ . We now consider the ratio of two quadratic forms. In general, if  $\text{Cor}(V_1, V_2) = \text{Cor}(W_1, W_2) = \rho_{12}$ , and  $\text{Cor}(V_1, W_1) = \text{Cor}(V_2, W_2) = 0$  then Taylor series methods give  $\text{Cor}(V_1/W_1, V_2/W_2) \approx \rho_{12}$ . This implies that the spatial correlation function of  $F_i$  is approximately equal to  $\rho(\mathbf{x})^2$ .

Worsley *et al.* (1995) show that the effect of averaging the square of a Gaussian random field is to divide the variance by the effective spatial degrees of freedom  $d$  (4.2). From this it follows that the effect of averaging  $F_i$  over voxels is to decrease the variance of the  $F$  distribution by the same factor  $d$ , which implies that the mean and variance of  $S$  are

$$\frac{\nu}{\nu - 2} \quad \text{and} \quad \frac{2\nu^2(h + \nu - 2)}{dh(\nu - 2)^2(\nu - 4)},$$



respectively. Following the argument in Worsley *et al.* (1995), it seems appropriate to approximate the distribution of a multiple of  $S$  by an  $F$ -distribution with degrees of freedom  $\nu_1$  and  $\nu_2$  chosen to match the mean and variance of  $S$ . To do this, we must equate the coefficients of variation, which leads to

$$\frac{h + \nu - 2}{dh(\nu - 4)} = \frac{\nu_1 + \nu_2 - 2}{\nu_1(\nu_2 - 4)}. \quad (\text{A.6})$$

This does not uniquely specify  $\nu_1$  and  $\nu_2$ ; one possibility is to match higher moments. Instead, we have found from simulations that a satisfactory choice is to fix  $\nu_1 = dh$ , which satisfies (A.6) for infinite  $\nu$  and  $\nu_2$ , then solve (A.6) for  $\nu_2$ , which leads to (4.3). Note that this gives the correct distribution for  $S$  when there is just one voxel and  $d = 1$ .

## A.6 Simulations in Fourier space

Considerable savings in computations were achieved by carrying out all simulations in Fourier space. The principle behind this is that all the methods advocated in this paper, the regression and the principal components, are invariant under orthonormal transformations in space or time. A Fourier transform of the data does not therefore affect the calculations, but it reduces the smoothing step to multiplication by the spectrum of the filter. Moreover many of the spectrum values are negligibly small, so that only 2292 of the  $N = 10500$  voxels were actually needed for the spatial data. The only exception to this is the operation of division by  $\hat{\sigma}_i^2$ , so in all our simulations we replaced it by its known variance  $\sigma_i^2$ , so that  $\nu = \infty$ . A further reduction is possible in the temporal domain since the only random variables needed for the simulations are  $\mathbf{Z}_i$ , whose  $h = 12$  components have independent Gaussian distributions, and not the entire temporal sequence of  $n = 120$  observations.

The columns of a matrix of  $2292 \times 12$  independent standard Gaussian random variables was multiplied by the products of the spatial spectrum to create the Fourier space equivalent of the spatially smoothed null data,  $\mathbf{Z}_i$ . To create the simulated signal data, a single instance of the spatially smoothed data, transformed to Fourier space as above, was used to represent the spatial distribution of a single population principal component. This was multiplied by a signal-to-noise ratio of 0.2 and added to the first component of  $\mathbf{Z}_i$ . Note that this spatial distribution is created once only, and added to each null data set, so that the population principal component is identical in all simulations.

## Acknowledgement

The authors would like to thank Fred Bookstein for correcting an earlier version of this paper, and the referees for very helpful comments.

## References

- Anderson, T. W. 1984. *An Introduction to Multivariate Statistical Analysis*, 2nd edition. New York: Wiley.

- Büchel, C., Wise, R. J. S., Mummary, C. J., Poline, J. B., & Friston, K. J. 1996. Non-linear regression in parametric activation studies. *NeuroImage*, **4**, 60–66.
- Bullmore, E.T., Rabe-Hesketh, S., Morris, R.G., Williams, S.C.R., Gregory, L., Gray, J.A., & Brammer, M.J. 1996. Functional magnetic resonance image analysis of a large scale neurocognitive network. *NeuroImage*, **4**, 16–33.
- Fletcher, P. C., Dolan, R. J., Shallice, T., Frith, C. D., Frackowiak, R. S. J., & Friston, K. J. 1996. Is multivariate analysis of PET data more revealing than the univariate approach? Evidence from a study of episodic memory retrieval. *NeuroImage*, **3**, 209–215.
- Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. 1991. Comparing functional (PET) images: The assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism*, **11**, 690–699.
- Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. 1993. Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism*, **13**, 5–14.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, B. J., Williams, C. R., Frith, C. D., Frackowiak, R. S. J., & Turner, R. 1995a. Analysis of fMRI time-series revisited. *NeuroImage*, **2**, 45–53.
- Friston, K. J., Frith, C. D., Frackowiak, R. S. J., & Turner, R. 1995b. Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage*, **2**, 166–172.
- Friston, K. J., Frith, C. D., Turner, R., & Frackowiak, R. S. J. 1995c. Characterizing evoked hemodynamics with fMRI. *NeuroImage*, **2**, 157–165.
- Friston, K. J., Poline, J.-B., Holmes, A. P., Frith, C. D., & Frackowiak, R. S. J. 1996. A multivariate analysis of PET activation studies. *Human Brain Mapping*, **4**, 140–151.
- McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. 1996. Spatial pattern analysis of functional brain images using Partial Least Squares. *NeuroImage*, **3**, 143–157.
- Worsley, K. J., & Friston, K. J. 1995. Analysis of fMRI time-series revisited - again. *NeuroImage*, **2**, 173–181.
- Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, **12**, 900–918.
- Worsley, K. J., Poline, J. B., Vandal, A. C., & Friston, K. J. 1995. Tests for distributed, non-focal brain activations. *NeuroImage*, **2**, 183–194.
- 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. *Human Brain Mapping*, **4**(1), 58–73.

Zarahn, E., Aguirre, G.K., & D'Esposito, M. 1997. Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null hypothesis conditions. *NeuroImage*, **5**, 179–197.

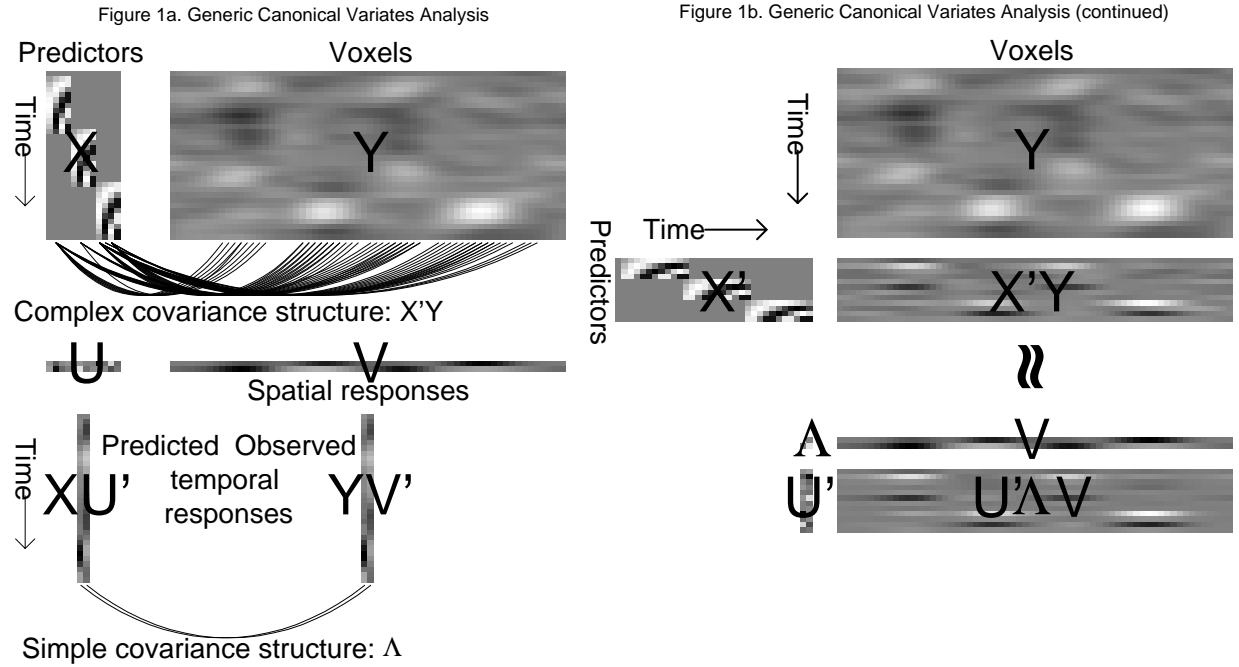


Figure 1: Generic Canonical Variates Analysis (CVA) for a simulated fMRI data set. The data consists of set of  $h = 12$  predictor variables  $\mathbf{X}$  and cerebral blood flow measurements  $\mathbf{Y}$  at  $N = 64$  voxels, each sampled at  $n = 32$  time points.  $\mathbf{U}$  and  $\mathbf{V}$  contain  $q = 2$  (orthonormal) latent variables for predictors and voxels, respectively, which act as weights. The weighted predictors  $\mathbf{XU}'$  and weighted voxels  $\mathbf{YV}'$  are the predicted and observed temporal responses, respectively. The aim is to capture the complex covariance structure  $\mathbf{X}'\mathbf{Y}$  of the original data by a simpler covariance structure  $\mathbf{\Lambda}$  of latent roots. This is done by approximating  $\mathbf{X}'\mathbf{Y}$  by  $\mathbf{U}'\mathbf{\Lambda}\mathbf{V}$ .

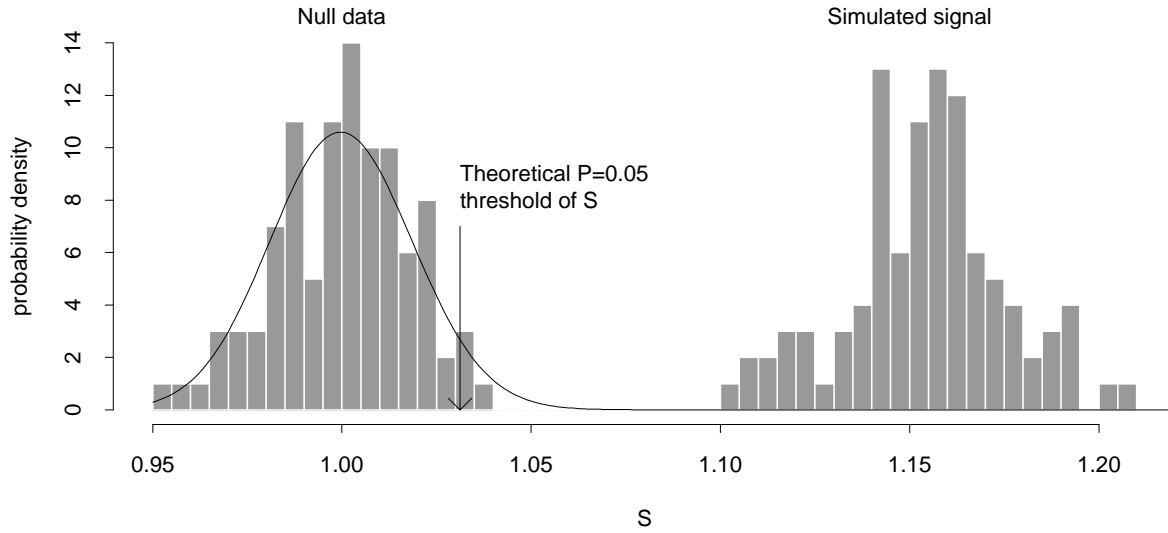
Fig 2: Histogram of  $S$ 

Figure 2: Histograms of the global test statistic  $S$  for (a) null, and (b) simulated signal data. Superimposed is the theoretical null  $F$  distribution and its  $P = 0.05$  threshold. The simulated signal data is highly significant.

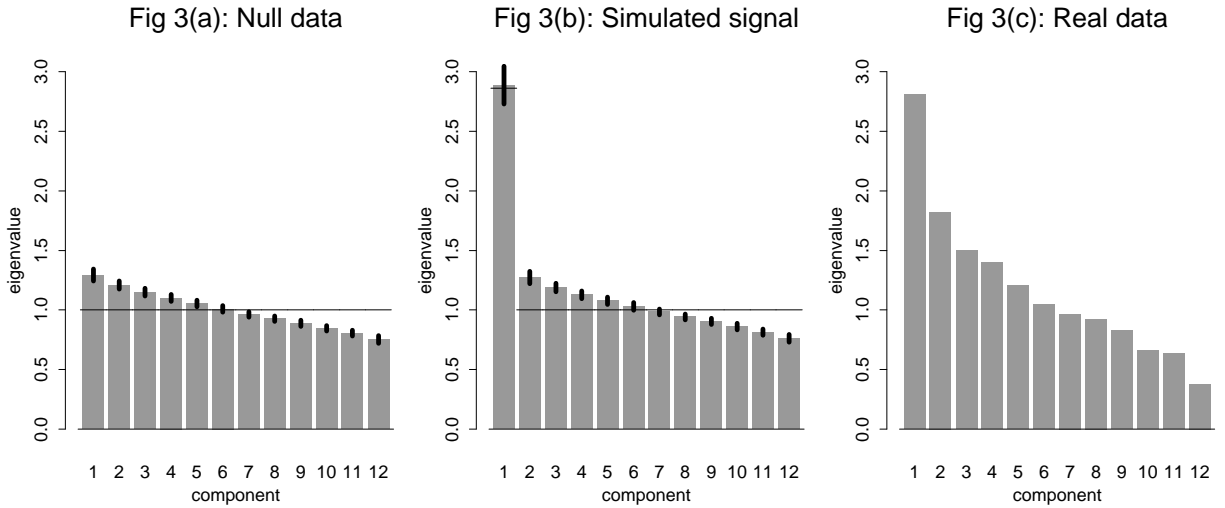


Figure 3: The average eigenvalues  $\lambda_j$  plotted against  $j$  for (a) null, (b) simulated signal, and (c) real data, together with one standard deviation error bars. Superimposed are the expected eigenvalues (lines): 1 for the non-null data, and 2.9 for the first component of the simulated data. The real data appears to show one strong component.

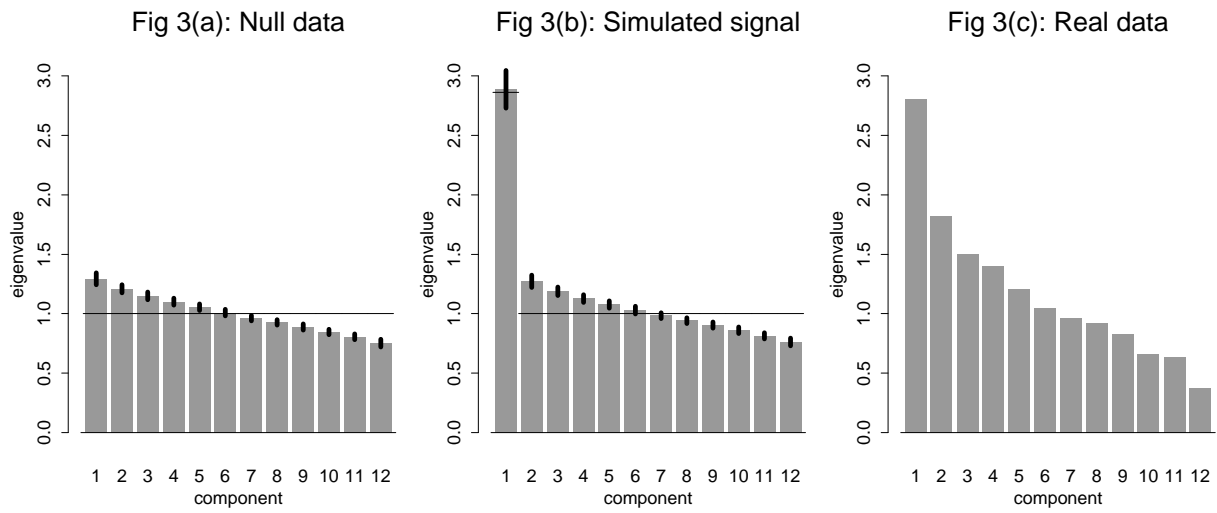


Figure 4: The  $p$ -values of the test statistics  $S_q$  for (a) null, (b) simulated signal, and (c) real data. The horizontal dashed line is the  $P = 0.05$  significance threshold. For the null data there are no significant components in 96% of the cases; for the simulated data, there is just one significant principal component in 98% of the cases; for the real data, there is just one significant principal component.

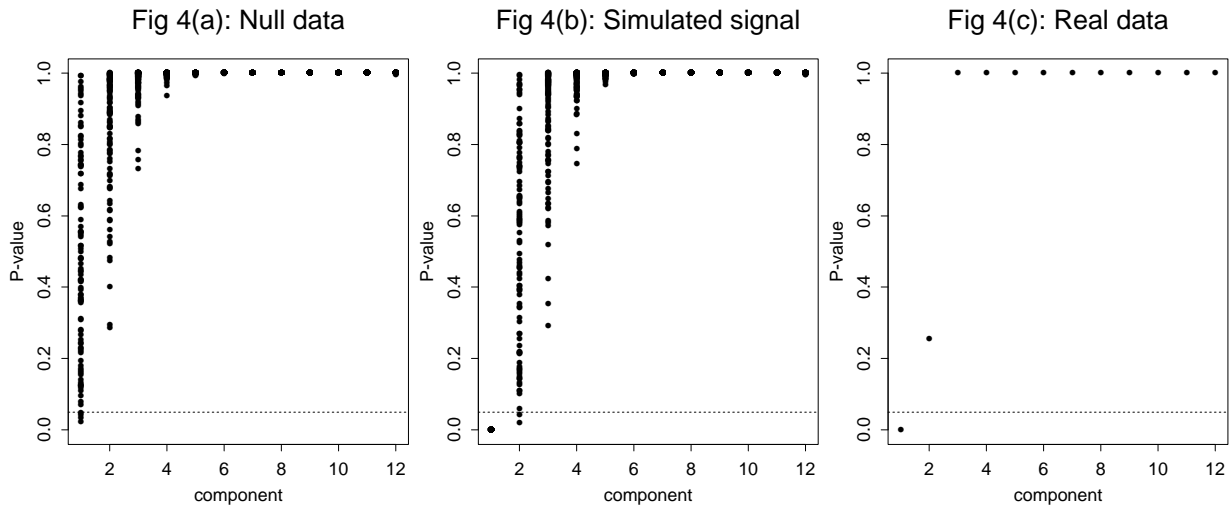


Figure 5: The predicted temporal response of the first principal component  $\hat{y}_1$ . The experiment consisted of 3 different tasks, each repeated 4 times, giving 12 epochs in all. Data from each epoch has been superimposed. Although the epochs were repeated after 10 scans, the predictor variables covered 12 scans to allow for a 2 scan overlap with the following epoch. Accordingly, 12 scans are shown on the temporal axis. The principal components were selected from (a) null, and (b) simulated signal data sets (3 second scans). Superimposed is the true temporal response that was added to the data (heavy lines): zero (no signal) for the null data, and a differential response for the three different tasks of the simulated data, taken from that reported by Friston *et al.* (1995b). The dotted lines represent the rest epochs, the broken lines represent the fixed epochs, and the solid lines the random epochs. The real data (c) shows a strong delayed response for the random condition, which matches quite closely the temporal response reported by (Friston *et al.* , 1995c, Figure 5) in the anterior cingulate, reproduced here as (d). The full predicted and observed temporal responses are shown in Figure A.6(b).

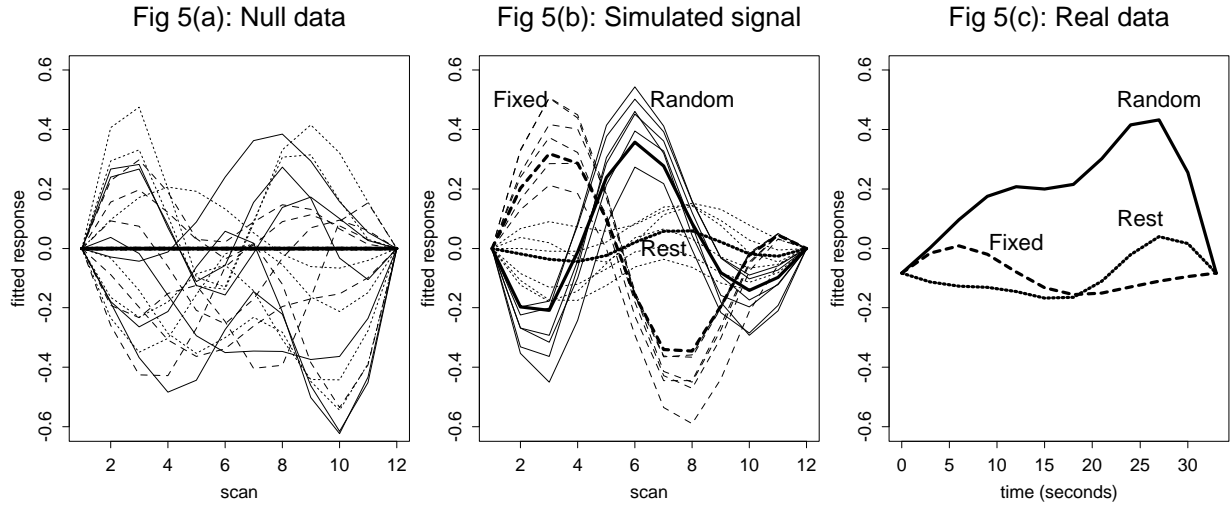


Figure 5(d): Anterior cingulate response from Friston *et al.* (1995c, Figure 5)

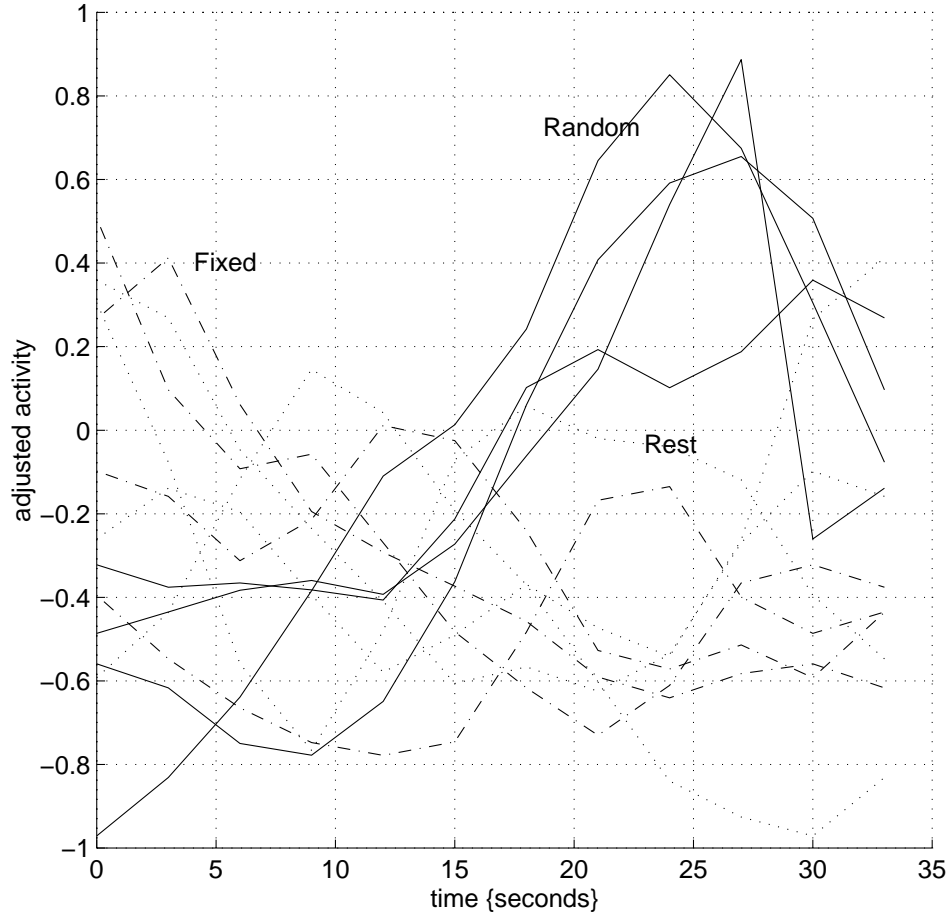


Figure 6: (a) Spatial response and (b) predicted (solid) and observed (dashed) temporal response for the first component of the real data. 1 on the scale of the observed and predicted temporal responses corresponds to one standard deviation of the observations. The predicted temporal response from each epoch is shown in Figure A.6(c). The spatial response is consistent with a differential response detected in the anterior cingulate in Friston *et al.* (1995c).



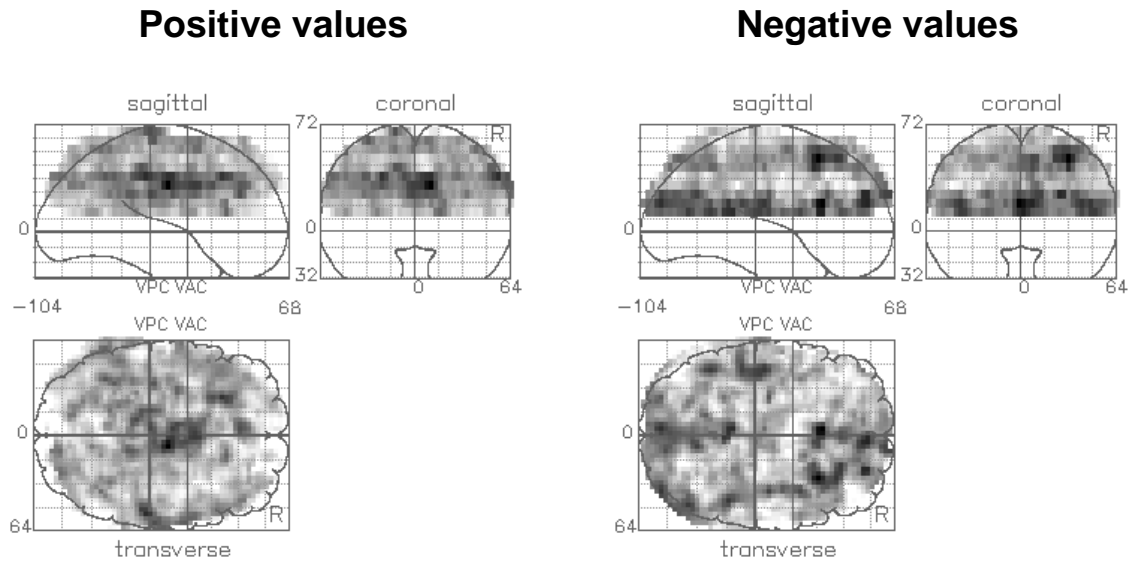
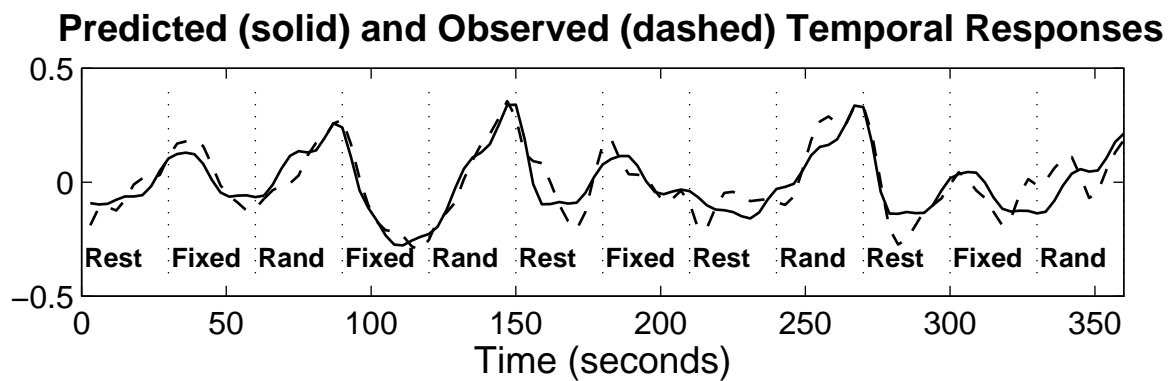
**Fig 6a. Spatial response****Fig 6b. Temporal responses**

Figure 7: Sensitivity of the  $F$ -test and SVD  $t$ -test to (a) a distributed signal and (b) a focal signal, at 0%, 10% and 20% signal to noise ratio. Also shown is the theoretical sensitivity at 0% signal, equal to the false positive rate (straight line). In all cases the  $F$ -test is more sensitive than the SVD  $t$ -test.