# Principal Component Regression Approach for Functional Connectivity of Neuronal Activations Measured by Functional MRI

*Eini I. Niskanen[1]†, Mika P. Tarvainen[1], Mervi Könönen[2], Hilkka Soininen[3], and Pasi A. Karjalainen[1]*

[1]University of Kuopio
Dept. of Applied Physics
P.O.Box 1627, FIN-70211 Kuopio
FINLAND
† Tel. +358-17-162367, Fax: +358-17-162373
† E-mail: Eini.Niskanen@uku.fi

[2]Kuopio University Hospital
Dept. of Clinical Neurophysiology
P.O.Box 1777, FIN-70211 Kuopio
FINLAND

[3]University of Kuopio
Dept. of Neuroscience and Neurology
P.O.Box 1627, FIN-70211
FINLAND

## ABSTRACT

*A principal component regression (PCR) based approach for studying the functional connectivity of blood oxygenation level dependent (BOLD) responses of functional magnetic resonance imaging (fMRI) is proposed. The temporal dependency of BOLD responses from different spatial areas is determined from the eigenvectors of the correlation matrix.*

## 1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a noninvasive method for studying human brain function and constructing whole brain activation maps for sensory and mental functions with relatively high spatial resolution. The most frequently used fMRI technique, the Blood Oxygenation Level Dependent (BOLD) technique, is based on different magnetic properties of oxygenated and deoxygenated hemoglobin in brain capillaries. Neuronal activity causes local increases in cerebral blood flow and cerebral blood volume. Although the neuronal activity consumes oxygen thus increasing the concentration of deoxygenated hemoglobin, the oxygen consumption is exceeded by the import of oxygenated blood. This causes a surplus of oxygenated hemoglobin which in $T_2^*$-weighted image is seen as an intensity increase and is called the BOLD response.

The relationship between stimulus, neural activation, and BOLD response has been studied since fMRI was introduced in 1992 [1], [2], but it is still not yet thoroughly understood. It has been found that the shape of the BOLD response varies across subjects and also within subject depending on the type of the stimulus and active cortical area [3]. The BOLD response is relatively slow, the peak of the response arises about 4-5 seconds after the stimulus and it may take even 10 seconds or more for the response to return to the baseline after that. For short interstimulus intervals this is problematic because the summation of the consecutive responses is highly nonlinear. However, linear deconvolution has been found to be effective, if the stimuli are separated by at least 4 seconds [4].

Normally the purpose of fMRI studies is to locate the cortical areas involved in processing the given stimulus. The analysis is done on a voxel-by-voxel basis not concentrating on the interactions or dependencies of different cortical areas. However, brain regions do not act in isolation, but connected and they communicate with each other. The term functional connectivity is defined as the temporal correlations among neurophysiological events between spatially remote cortical areas [5]. Functional connectivity studies have, however, been mainly concentrated on block-designs [6], [7], whereas temporal dependency of event-related responses has gained less attention.

In this paper, we propose a principal component regression (PCR) [8] based approach for studying the functional connectivity of BOLD responses. In the method, BOLD responses from different cortical areas are joined together to form an augmented PCR system. The temporal correlations of BOLD responses from different cortical areas can then be determined from the eigenvectors of the augmented data correlation matrix. The method is tested by using simulated fMRI data.

## 2. METHODS

### 2.1 fMRI measurements and simulations

For the simulations a young healthy right-handed volunteer was scanned in the Department of Clinical Radiology in the Kuopio University Hospital with a Siemens Magnetom Vision 1.5 T MRI scanner in order to obtain real noise and signal fluctuations caused by head motion, cardiac and respiratory cycles, and hardware-related signal drifts. Approximately 700 $T_2^*$-weighted gradient-echo echo-planar (EP) images and a $T_1$-weighted anatomical image were acquired. During the scanning of the EP images the volunteer was instructed to do nothing in order to acquire data as

clean as possible with no intentional actual activations. All EP images were acquired using interscan interval of 2.5 seconds. The EP images comprised of 16 slices with in-plane resolution of $4 \times 4$ mm. The anatomical $T_1$-image comprised of 180 sagittal slices with in-plane resolution of $1 \times 1$ mm covering the whole head.

The data was preprocessed using SPM2 (Statistical Parametric Mapping) developed by the Wellcome institute in London, UK (`http://www.fil.ion.ucl.ac.uk/spm/`). The preprocessing involved correction of movements, correction of differences in acquisition times between slices, coregistration of the EP images to the anatomical image, and normalization of the images to the SPM2 template.

Data from two different cortical areas were selected for analysis. These were the primary visual area V1 (area 1) and the primary motor area, Brodmann area 4, (area 2). A voxel from both of these areas was chosen and 70 artificial event-like BOLD responses were added to both time series.

The artificial BOLD responses were generated using a balloon model. The balloon model is a simple biomechanical model for the hemodynamic changes during brain activation [9], [10]. The central idea of the model is that the venous compartment is modeled as an expandable balloon. The neuronal activity causes an increase in cerebral blood flow (CBF) that increases the cerebral blood volume (CBV) i.e. inflates a venous "balloon". The two dynamical variables in defining the actual BOLD effect are the total deoxyhemoglobin content $q$ and the volume of the balloon $v$. The time dependent changes in $v$ and $q$ are driven by changes in CBF, CBV, and cerebral metabolic rate of oxygen (CMRO2). The balloon model is a nonlinear model for the BOLD signal change $\frac{\Delta S}{S}$

$$\frac{\Delta S}{S} = V_0 \left[ k_1 (1-q) + k_2 \left(1 - \frac{q}{v}\right) + k_3 (1-v) \right] \quad (1)$$

where $V_0$ represents the baseline CBV and $k_1$, $k_2$ and $k_3$ are dimensionless parameters that depend on the imaging parameters among other things.

There is always some delay between the onset of the BOLD response and the neuronal stimulus. This neuronal delay in area 1, $\delta_{n1}$, was assumed to be $\chi^2$ distributed with 0.5 degrees of freedom. The neuronal delay in area 2, $\delta_{n2}$, was assumed to be $\chi^2$ distributed with 0.6 degrees of freedom. In addition, a constant delay $\delta_c = 300$ms between the responses in area 1 and area 2 was assumed. We created two data sets. The mean of the neuronal delays in area 1 was 490ms in both data sets. In the first data set the response in area 2 was independent on the neuronal delay in area 1 i.e. $\delta_{tot2} = \delta_c + \delta_{n2}$, where $\delta_{tot2}$ is the total delay between the stimulus and the onset of the BOLD response in area 2. The mean of total delays in area 2 was 720ms. In the second data set the response in area 2 was dependent on the neuronal delay in area 1 i.e. $\delta_{tot2} = \delta_{n1} + \delta_c + \delta_{n2}$. The mean of total delays in the dependent case in area 2 was 1210ms.

### 2.2 Principal component regression

The voxel time series from both areas were divided into adequate BOLD responses. Let us denote the $m$'th such response from area 1 with $z_m^1(t) = \left[z_m^1(1), \cdots, z_m^1(N)\right]^T$ and that from area 2 with $z_m^2(t)$ correspondingly. An augmented $2N \times M$ data matrix is then formed as

$$Z = \begin{bmatrix} z_1^1(t) & \cdots & z_M^1(t) \\ \\ z_1^2(t) & \cdots & z_M^2(t) \end{bmatrix} \quad (2)$$

where $M$ is the number of responses and $N$ is the length of responses.

In the principal component regression, the vector containing the measured data is presented as a weighted sum of orthogonal basis vectors. The basis vectors are selected to be the eigenvectors of either covariance or correlation matrix of the data. In this study, the correlation matrix is used since it preserves the mean of the data.

The correlation matrix of the augmented BOLD responses can be estimated as

$$R_Z = \frac{1}{M} Z Z^T. \quad (3)$$

The eigenvectors $u$ and the corresponding eigenvalues $\lambda$ can be solved from the eigendecomposition. Quantitatively, the first eigenvector $u_1$ is the best mean square fit of a single waveform to the set of augmented BOLD responses. Thus, the first eigenvector is often similar to the mean. In case the BOLD responses are otherwise similar but there is some time variation in response onsets, the second eigenvector covers this latency jitter and is often similar to the first derivative of the response [11]. If the time variation of response onsets in areas 1 and 2 are independent, then it is expected that both the second and third eigenvactors are required to cover the variation. The significance of each eigenvector is described by the corresponding eigenvalue. Information about the trial-to-trial responses can be obtained from the principal components

$$\hat{\theta}_{PC} = U^T Z \quad (4)$$

where $U = [u_1, \cdots, u_M]$.

## 3. Results

The 70 simulated BOLD responses for both areas for the independent and dependent cases are illustrated in

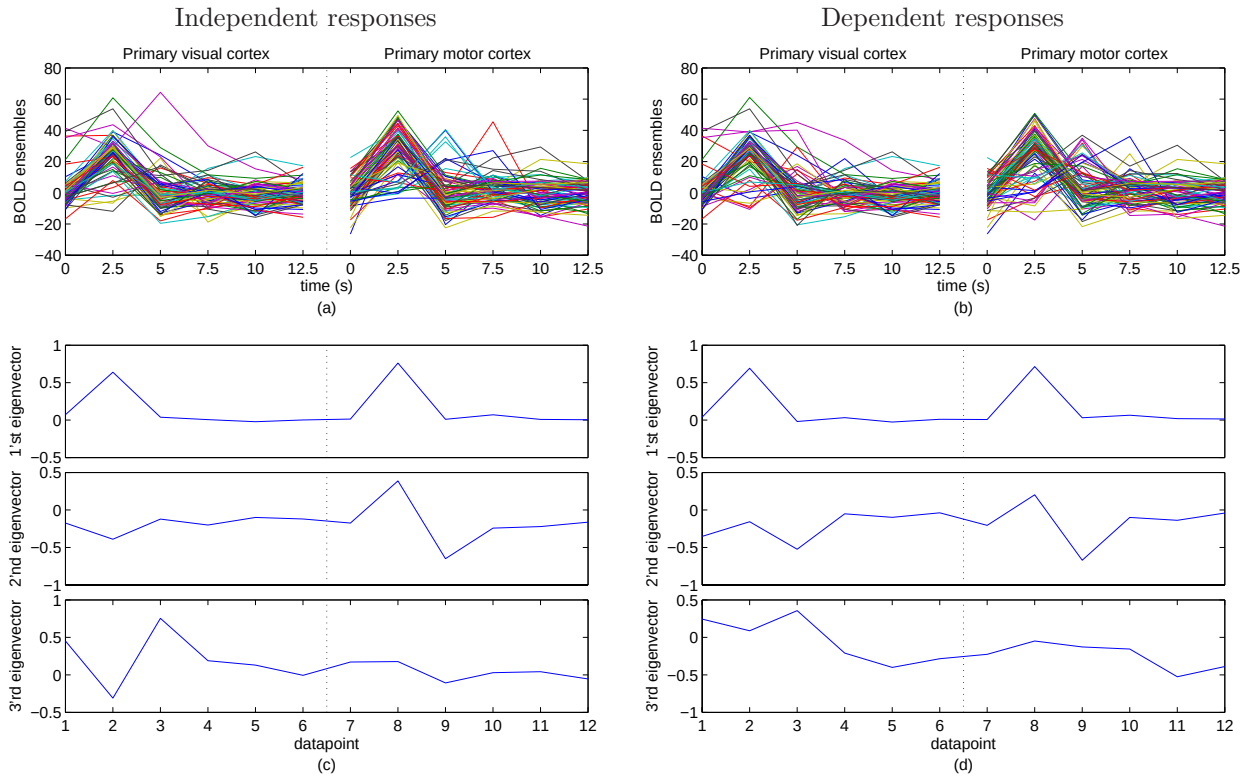


**Fig. 1.** The comparison of the independent and dependent data sets. (a) The 70 simulated BOLD responses for independent data set and (b) for the dependent data set. (c) The first (top), the second (middle) and the third (bottom) eigenvector for the independent data set. (d) The first (top), the second (middle) and the third (bottom) eigenvector for the dependent data set.

Fig. 1 (a) and (b), respectively. The length of each response was 12.5 seconds and each of them consisted of six data points. As can be seen, there is more variance in the onsets of the BOLD responses in area 2 in the dependent case [Fig. 1.(b)] than in the independent case [Fig. 1.(a)]. The BOLD responses for areas 1 and 2 were then concatenated according to (2) for independent and dependent cases, thus, creating two data sets. The eigenvectors and the corresponding eigenvalues were calculated for both data sets. The first three eigenvectors of the correlation matrix $R_Z$ are shown in Fig.1 (c) for the independent case and in Fig.1 (d) for the dependent case in points. The first eigenvectors are quite similar in both data sets. The interesting difference between the data sets is found in the second and third eigenvectors. In the independent case the second eigenvector covers the latency jitter of the BOLD responses in area 2 and the third eigenvector the latency jitter of the BOLD responses in area 1. In the dependent case, on the other hand, the second eigenvector covers most of the latency jitter of BOLD responses in both areas. The first three eigenvalues for the independent data set were $\lambda_1 = 0.5968$, $\lambda_2 = 0.1220$, and $\lambda_3 = 0.0850$ and for the dependent data set $\lambda_1 = 0.6055$, $\lambda_2 = 0.1390$, and $\lambda_3 = 0.0711$.

## 4. Discussion

We have proposed an approach for studying the functional connectivity of distinct cortical areas. We have shown that using a principal component regression based approach the dependency of the cortical areas can be determined from the second and third eigenvectors. In the independent case the second and third eigenvectors cover the independent time variations of the BOLD responses in different areas. Whereas, in the dependent case, the time variation can be mainly covered by one eigenvector. This can also be seen from the eigenvalues, which reflect the significance of the eigenvectors. In the independent case the second and third eigenvalues are somewhat closer to each other than in the dependent case.

Trial-to-trial information of the BOLD responses can be further estimated from the principal components and will be studied in future. The method will also be further tested with real fMRI data and its connection to the balloon model will be studied.

## REFERENCES

[1] K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskof, B. P. Poncelet, D. N. Kennedy, B. E. Hoppel, M. S. Cohen, R. Turner, H.-M. Cheng, T. J. Brady, and B. R. Rosen, "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation," *Proc. Natl. Acad. Sci. USA*, vol. 89, pp. 5675–5679, June 1992.

[2] A. M. Blamire, S. Ogawa, K. Ugurbil, D. Rothman, G. McCarthy, J. M. Ellerman, F. Hyder, Z. Rattner, and R. G. Shulman, "Dynamic mapping of the human visual cortex by high-speed magnetic resonance imaging," *Proc. Natl. Acad Sci. USA*, vol. 89, pp. 11069–11073, November 1992.

[3] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human, BOLD hemodynamic responses," *NeuroImage*, vol. 8, pp. 360–369, November 1998.

[4] A. L. Vazquez and D. C. Noll, "Nonlinear aspects of the BOLD response in functional MRI," *NeuroImage*, vol. 7, no. 2, pp. 108–118, 1998.

[5] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak, "Functional connectivity: The principal-component analysis of large (PET) data sets.," *Journal of Cerebral Blood Flow and Metabolism*, vol. 13, pp. 5–14, 1993.

[6] M. J. Lowe, M. Dzemidzic, J. T. Lurito, V. P. Mathews, and M. D. Phillips, "Correlations in Low-Frequency BOLD Fluctuations Reflect Cortico-Cortical Connections," *NeuroImage*, vol. 12, no. 5, pp. 582–587, 2000.

[7] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore, "Detection of functional connectivity using temporal correlations in MR images," *Human Brain Mapping*, vol. 15, no. 4, pp. 247–262, 2002.

[8] I. T. Jolliffe, *Principal Component Analysis*. Springer-Verlag New York, Inc., second ed., 2002.

[9] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of Blood Flow and Oxygenation Changes During Brain Activation: The Balloon Model," *Magnetic Resonance in Medicine*, vol. 39, pp. 855–864, 1998.

[10] T. Obata, T. T. Liu, K. L. Miller, W.-M. Luh, E. C. Wong, L. R. Frank, and R. B. Buxton, "Discrepancies between BOLD and flow dynamics in primary and sipplementary motor areas: application of the balloon model to the interpretation of BOLD transients," *NeuroImage*, vol. 21, pp. 144–153, 2004.

[11] J. Möcks, "The Influence of Latency Jitter in Principal Component Analysis of Event-Related Potentials," *Psychophysiology*, vol. 23, no. 4, pp. 480–484, 1986.