

Group Study of Simulated Driving fMRI Data by Multiset Canonical Correlation Analysis

Yi-Ou Li · Tom Eichele · Vince D. Calhoun ·
Tulay Adali

Received: 18 August 2010 / Revised: 15 November 2010 / Accepted: 21 November 2010
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract In this work, we apply a novel statistical method, multiset canonical correlation analysis (M-CCA), to study a group of functional magnetic resonance imaging (fMRI) datasets acquired during simulated driving task. The M-CCA method jointly decomposes fMRI datasets from different subjects/sessions into brain activation maps and their associated time courses, such that the correlation in each group of estimated activation maps across datasets is maximized. Therefore, the functional activations across all datasets are extracted in the order of consistency across different dataset. On the other hand, M-CCA preserves the uniqueness

of the functional maps estimated from each dataset by avoiding concatenation of different datasets in the analysis. Hence, the cross-dataset variation of the functional activations can be used to test the hypothesis of functional-behavioral association. In this work, we study 120 simulated driving fMRI datasets and identify parietal-occipital regions and frontal lobe as the most consistently engaged areas across all the subjects and sessions during simulated driving. The functional-behavioral association study indicates that all the estimated brain activations are significantly correlated with the steering operation during the driving task. M-CCA thus provides a new approach to investigate the complex relationship between the brain functions and multiple behavioral variables, especially in naturalistic tasks as demonstrated by the simulated driving study.

Keywords Blind source separation · Canonical correlation analysis · fMRI · Simulated driving · Functional behavioral association

Y.-O. Li (✉) · T. Adali
Department of Computer Science and Electrical
Engineering, University of Maryland Baltimore County,
1000 Hilltop Circle, Baltimore, MD 21250, USA
e-mail: liyou1@umbc.edu, Yiou.Li@ucsf.edu

T. Adali
e-mail: adali@umbc.edu

T. Eichele
Department of Biological and Medical Psychology,
University of Bergen, Jonas Lies Vei 91,
5011 Bergen, Norway

V. D. Calhoun
The Mind Research Network, 1101 Yale Boulevard N.E.,
Albuquerque, NM 87106, USA
e-mail: vcalhoun@unm.edu

V. D. Calhoun
Department of Electrical and Computer Engineering,
University of New Mexico, Albuquerque, NM, USA

V. D. Calhoun
Department of Psychiatry, Yale University School
of Medicine, New Haven, CT, USA

1 Introduction

Driving is a naturalistic behavior in which multiple brain networks are involved. Brain function during driving has been studied through fMRI [1–4] and PET [5], using the conventional hypothesis-driven analysis based on the general linear model (GLM). On the other hand, data-driven methods such as independent component analysis (ICA) and canonical correlation analysis (CCA) have been successfully applied to the estimation of brain activation in fMRI data [6, 7]. Compared with conventional fMRI analysis based on GLM,

data-driven methods do not require pre-specified response model and hence can identify multiple active brain regions due to intrinsic connectivity as well as task activation in more flexible ways. Therefore, data-driven methods are especially attractive for analyzing brain function during naturalistic and spontaneous states, such as simulated driving, movie watching, and functional connectivity during the resting state.

Data-driven approach has been applied to study brain networks during driving and identified multiple neural correlates as well as their modulation with driving speed [8, 9]. The estimated time courses of the identified brain regions provide information about the temporal modulation of the corresponding brain networks, which is not particularly predictable and thus can not be obtained through a hypothesis-driven method such as GLM. Such a study typically involves processing of a group of subjects/sessions and the inference of group level brain activations. When multiple datasets are analyzed by data-driven methods independently, the results do not necessarily have a intrinsic correspondence, even if the datasets are collected using the same experimental paradigm. Hence, keeping the coherency among the estimated brain activations across different datasets is the key issue in a data-driven analysis of group fMRI data.

Toward this end, ICA has been extended to handle group data in two different ways. In one approach, spatial or temporal concatenation is first applied to group datasets and ICA is used to estimate independent sources within the aggregated sub-space containing the common features. Methods in this category include: Group ICA [10], tensorial ICA [11], and a unified multi-dataset maximum likelihood ICA model [12]. In an alternative approach, a corresponding source model is constructed among the latent sources across different datasets, and the independence between different groups of corresponding sources are maximized by linear demixing vectors specific to each dataset. Methods in this category include independent vector analysis (IVA) and M-CCA. IVA achieves individual decomposition without dataset concatenation and controls cross-subject correspondence by incorporating a joint density on sources in each independent vector with higher-order dependence [13]. However, the corresponding sources in IVA density model assume no second-order dependence, which could be impractical because in group study, second-order dependence is typically significant in characterizing source correspondence.

To achieve individual decomposition while incorporating second-order dependence across the corresponding sources, we propose a joint source separation

scheme on multiple datasets by M-CCA [14] such that sources are extracted from the datasets by maximizing the between-set source correlations. The method uses multiple stages, where in each stage, a linear projection is found for each multi-dimensional dataset such that correlation among the group of projected data i.e., the extracted sources, is maximized, and the sources extracted at different stages are constrained to be uncorrelated within each dataset.

Figure 1 illustrates the generative models of different data-driven group analysis methods, where (a) group ICA and (b) tensorial ICA are representative group analysis models based on data concatenation, (c) IVA and (d) M-CCA are models incorporating source correspondence. IVA and M-CCA model the statistical dependence between the corresponding sources in a complementary manner. That is, IVA models source correspondence by higher order statistical dependence while M-CCA models source correspondence by second-order correlation.

One important property of the M-CCA method is that it is robust to the increasing number of datasets incorporated into the joint analysis [15]. Since M-CCA models the latent sources in each dataset independently, it preserves the distinction of the source estimates for each dataset. Therefore, M-CCA is a promising candidate method for group inference from large number of datasets and study of brain function variations across different subjects.

In this work, a group of fMRI datasets from simulated driving are jointly separated by the M-CCA method, resulting in multiple brain networks. To draw inference from the rich information resulting from the group analysis, first, group level activation maps are obtained from the set of jointly extracted sources, i.e., the group of corresponding sources. This is done by a rank-one approximation using principal component analysis. Secondly, the cross-dataset variability, i.e., the eigenvector corresponding to the most significant principal component, is used to perform second level analysis to study the association with the behavioral factors recorded during the driving task using a multiple linear regression. M-CCA identifies four functional brain networks in order of cross-dataset consistency and their corresponding time courses. A regression analysis of each cross subject variation profile with the behavioral factors reveals significant association between functional activations to the steering behavior during simulated driving task.

This paper is organized as follows. In Section 2, we briefly introduce the modeling assumption of group fMRI data for M-CCA to achieve joint source separation, the two-sample intra-class structure to model

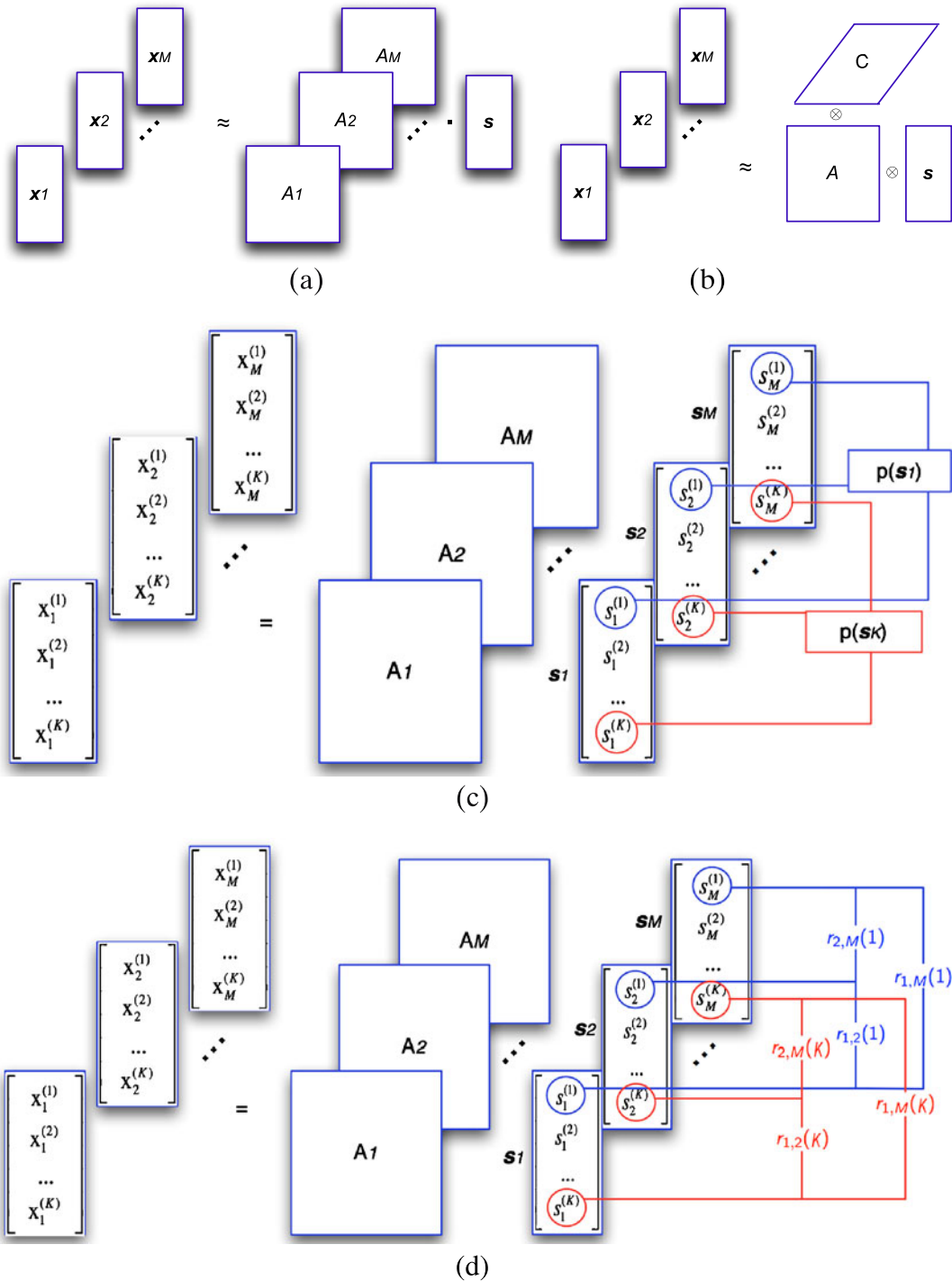


Fig. 1 Comparison of group analysis models: **a** group ICA and **b** tensorial ICA are representative group analysis models based on data concatenation, **c** IVA and **d** M-CCA are models incorporating source correspondence. IVA and M-CCA model

the statistical dependence between the corresponding sources in a complementary manner. That is, IVA models source correspondence by higher order statistical dependence while M-CCA models source correspondence by second order correlation.

the heterogeneous correlation structure within a group of corresponding sources, and the calculation of cross-dataset variation for behavior association study. In Section 3, we simulate group datasets to test the perfor-

mance of the proposed method and introduce the simulated driving experiment and scheme for group fMRI data analysis. The experimental results on simulated datasets and driving fMRI datasets are presented in

Section 4. In Section 5, we conclude our work with a discussion on the M-CCA method for group analysis.

2 Theory

2.1 Joint Estimation of Brain Sources by M-CCA Using Eigenvalue Maximization

We assume the following generative model for joint BSS on a group of mixtures:

- (i) For a group of M datasets, each dataset, $\mathbf{x}_m = [x_m^{(1)}, x_m^{(2)}, \dots, x_m^{(K)}]^T$, $m = 1, 2, \dots, M$ contains linear mixtures of K sources given in the source vector $\mathbf{s}_m = [s_m^{(1)}, s_m^{(2)}, \dots, s_m^{(K)}]^T$, mixed by a non-singular matrix \mathbf{A}_m , i.e.,

$$\mathbf{x}_m = \mathbf{A}_m \mathbf{s}_m, \tag{1}$$

where $\mathbf{x}_m, \mathbf{s}_m$ are K -dimensional random vectors, whose samples form the mixture dataset and source dataset respectively, \mathbf{A}_m is a $K \times K$ non-singular square matrix;

In group fMRI data analysis, the dataset from subject m , \mathbf{x}_m , contains the scanned brain volume at time points $1, 2, \dots, K$ where the samples of variable $x_m^{(k)}$ are drawn from all the voxels in the scanned volume image at time k . On the other hand, each variable $s_m^{(k)}$ in \mathbf{s}_m contains one estimated spatial independent source for subject m and the corresponding column in \mathbf{A}_m contains the time course of the spatial source.

- (ii) The correlation matrix of the concatenated source vector $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_M^T]^T$ are composed of diagonal blocks, i.e.,

$$\mathbf{R} \triangleq E\{\mathbf{s}\mathbf{s}^T\} = \begin{bmatrix} \mathbf{I} & \mathbf{\Lambda}_{1,2} & \dots & \mathbf{\Lambda}_{1,M} \\ \mathbf{\Lambda}_{2,1} & \mathbf{I} & \dots & \mathbf{\Lambda}_{2,M} \\ \dots & \dots & \dots & \dots \\ \mathbf{\Lambda}_{M,1} & \mathbf{\Lambda}_{M,2} & \dots & \mathbf{I} \end{bmatrix}$$

where $\mathbf{\Lambda}_{m,n} = \text{diag}([r_{m,n}^{(1)}, r_{m,n}^{(2)}, \dots, r_{m,n}^{(K)}])$ for $m, n \in \{1, 2, \dots, M\}$ is a diagonal matrix with the correlation values of the corresponding sources in \mathbf{s}_m and \mathbf{s}_n on its diagonal. In other words, the latent sources across different datasets are correlated only at the same index. Those correlated sources form the group of *corresponding sources*.

Based on the above model, the groups of corresponding sources can be jointly separated in multiple steps by M-CCA [15]. In step k , the overall correlations

among the extracted sources is maximized with respect to a group of demixing vectors:

$$\{\mathbf{w}_1^{(k)}, \mathbf{w}_2^{(k)}, \dots, \mathbf{w}_M^{(k)}\} = \arg \max_{\mathbf{w}} \left\{ \lambda_{\max}(\hat{\mathbf{R}}_k) \right\},$$

where λ_{\max} represents the greatest eigenvalue of its matrix argument and $\hat{\mathbf{R}}_k = E\{\hat{\mathbf{s}}^{(k)}\hat{\mathbf{s}}^{(k)T}\} =$

$$\begin{bmatrix} \mathbf{w}_1^{(k)T} E\{\mathbf{x}_1 \mathbf{x}_1^T\} \mathbf{w}_1^{(k)} & \mathbf{w}_1^{(k)T} E\{\mathbf{x}_1 \mathbf{x}_2^T\} \mathbf{w}_2^{(k)} & \dots & \mathbf{w}_1^{(k)T} E\{\mathbf{x}_1 \mathbf{x}_M^T\} \mathbf{w}_M^{(k)} \\ \mathbf{w}_2^{(k)T} E\{\mathbf{x}_2 \mathbf{x}_1^T\} \mathbf{w}_1^{(k)} & \mathbf{w}_2^{(k)T} E\{\mathbf{x}_2 \mathbf{x}_2^T\} \mathbf{w}_2^{(k)} & \dots & \mathbf{w}_2^{(k)T} E\{\mathbf{x}_2 \mathbf{x}_M^T\} \mathbf{w}_M^{(k)} \\ \dots & \dots & \dots & \dots \\ \mathbf{w}_M^{(k)T} E\{\mathbf{x}_M \mathbf{x}_1^T\} \mathbf{w}_1^{(k)} & \mathbf{w}_M^{(k)T} E\{\mathbf{x}_M \mathbf{x}_2^T\} \mathbf{w}_2^{(k)} & \dots & \mathbf{w}_M^{(k)T} E\{\mathbf{x}_M \mathbf{x}_M^T\} \mathbf{w}_M^{(k)} \end{bmatrix}$$

is the correlation matrix of the k th group of corresponding sources estimated by the demixing vectors $\{\mathbf{w}_1^{(k)}, \mathbf{w}_2^{(k)}, \dots, \mathbf{w}_M^{(k)}\}$.

In addition, an orthogonality constraint is imposed in each step to guarantee that the extracted sources from previous steps are not re-extracted in current step:

$$\hat{\mathbf{s}}_m^{(k)} \perp \{\hat{\mathbf{s}}_m^{(1)}, \hat{\mathbf{s}}_m^{(2)}, \dots, \hat{\mathbf{s}}_m^{(k-1)}\}, m = 1, 2, \dots, M.$$

Therefore, we have

$$\hat{\mathbf{s}}_m = \begin{bmatrix} \mathbf{w}_m^{(1)T} \\ \mathbf{w}_m^{(2)T} \\ \dots \\ \mathbf{w}_m^{(K)T} \end{bmatrix} \mathbf{x}_m, m = 1, 2, \dots, M,$$

as the separated sources for the m th dataset. Figure 1d illustrates the M-CCA decomposition on group datasets.

The M-CCA procedure outlined above optimizes demixing vectors such that the largest eigenvalue of source correlation matrix $\hat{\mathbf{R}}_k$ is maximized. This procedure is similar to principal component analysis (PCA) on the concatenation of multiple datasets $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_M^T]^T$ since both of them find the linear transformations that maximize the eigenvalue of the correlation matrix. However, when estimating multiple sets of corresponding sources, the orthogonality among the principal components of \mathbf{x} does not necessarily translate to orthogonality of the estimated sources within each individual dataset. In other words, when the correlation structure within the group of corresponding sources, $\{s_1^{(k)}, s_2^{(k)}, \dots, s_M^{(k)}\}$, has heterogeneous correlation values, the orthogonality among the principal components of \mathbf{x} , i.e., the pooled common factors of the group datasets, is not sufficient to characterize the orthogonality among the source estimates within each dataset. This case is studied in the next section.

2.2 Modeling the Correlation Structure Within each Group of Corresponding Sources

When $\{s_1^{(k)}, s_2^{(k)}, \dots, s_M^{(k)}\}$ have homogeneous correlation values, we can approximately model their correlation matrix \mathbf{R}_k using an intra-class correlation model

[15]. In this case, M-CCA and PCA on concatenation of all the datasets both can achieve joint BSS of the latent sources according to the proposed generative model.

When $\{s_1^{(k)}, s_2^{(k)}, \dots, s_M^{(k)}\}$ have heterogeneous correlation values, the number of independent correlation coefficients is $M(M - 1)/2$ and it is not straightforward to calculate relation between these coefficients and the eigenvalues of the correlation matrix. Considering a simplified case, we model \mathbf{R}_k by a 2×2 BLOCK matrix with the intra-class correlation matrix structure on its diagonal blocks, i.e.,

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_1 & \rho_3 & \rho_3 & \dots & \rho_3 \\ \rho_1 & 1 & \dots & \rho_1 & \rho_3 & \rho_3 & \dots & \rho_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_1 & \rho_1 & \dots & 1 & \rho_3 & \rho_3 & \dots & \rho_3 \\ \rho_3 & \rho_3 & \dots & \rho_3 & 1 & \rho_2 & \dots & \rho_2 \\ \rho_3 & \rho_3 & \dots & \rho_3 & \rho_2 & 1 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_3 & \rho_3 & \dots & \rho_3 & \rho_2 & \rho_2 & \dots & 1 \end{bmatrix}$$

where \mathbf{Q}_{11} and \mathbf{Q}_{22} are, respectively, $M_1 \times M_1$ and $M_2 \times M_2$ correlation matrices of the two sub-groups with correlation coefficient ρ_1 and ρ_2 , $\mathbf{Q}_{12} = \mathbf{Q}_{21}$ are the cross-correlation matrices with the cross-correlation

coefficient ρ_3 between the two sub-groups. In this way, the number of independent correlation coefficients is reduced to three and analytical solutions can be derived for the eigenvalues based on the correlation coefficients.

Notice that:

- (i) $\mathbf{Q}_{11} = (1 - \rho_1)\mathbf{I} + \rho_1\mathbf{1}\mathbf{1}^T$, where $\mathbf{1}$ is a column vector of all ones, it can be shown that $\lambda_{\max}(\mathbf{Q}_{11}) = 1 + (M - 1)\rho_1$, and $\lambda_{\min}(\mathbf{Q}_{11}) = 1 - \rho_1$ with multiplicity $(M - 1)$;
- (ii) $\mathbf{Q}_{12} = \mathbf{Q}_{21} = \rho_3\mathbf{1}\mathbf{1}^T$; and
- (iii) $\det(\mathbf{R}_k - \lambda\mathbf{I}) = \det(\mathbf{Q}_{11}) \cdot \det(\mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12})$.

Hence, it can be show that the eigenvalues of \mathbf{R}_k are:

$$\lambda = \begin{cases} 1 - \rho_1, & (M_1 - 1) \\ 1 - \rho_2, & (M_2 - 1) \\ \frac{1}{2} \left[(a + b) \pm \sqrt{(a + b)^2 - 4(ab - c)} \right] \end{cases} \quad \text{where } a, b,$$

$$\text{and } c \text{ are defined as } \begin{cases} a \triangleq 1 - \rho_1 + M_1\rho_1 \\ b \triangleq 1 - \rho_2 + M_2\rho_2 \\ c \triangleq M_1M_2\rho_3^2 \end{cases}$$

To demonstrate the relationship between the correlation values, ρ_1 , ρ_2 , and ρ_3 in \mathbf{R}_k and its first and second largest eigenvalues, we simulate a two sample intra-class

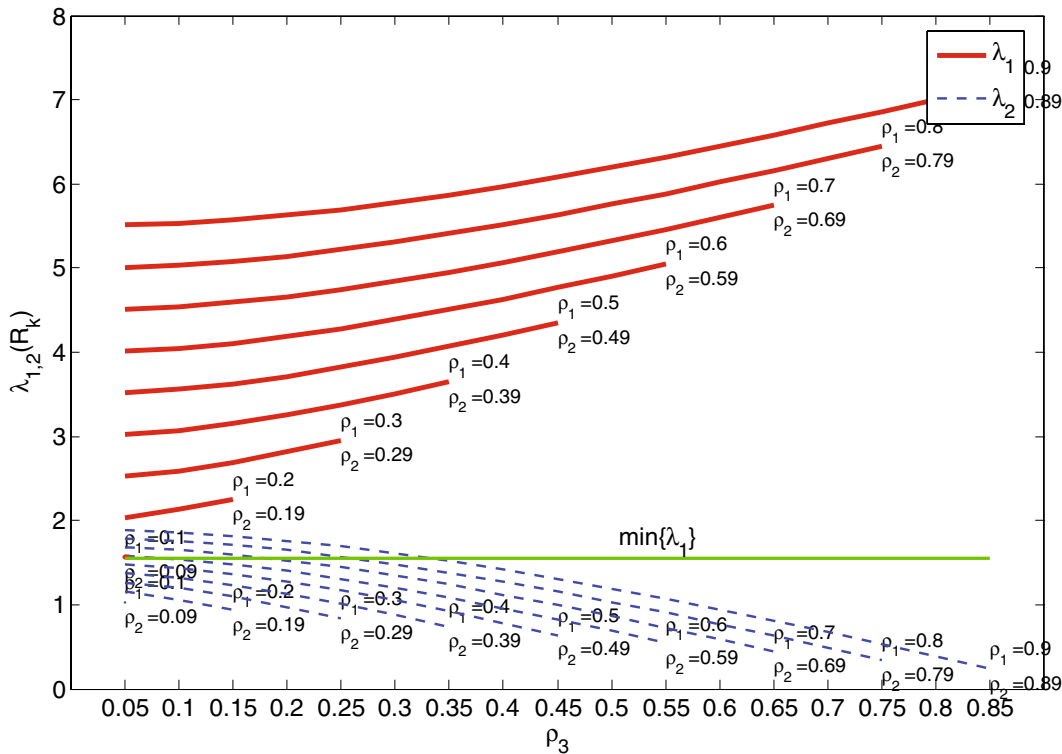


Fig. 2 First and second eigenvalue of the two sample intra-class correlation matrix with different configurations of correlation values. The difference between the first and second eigenvalues of \mathbf{R}_k reduces as the inter-class correlation ρ_3 decreases. The

second eigenvalue of \mathbf{R}_{k_1} with $\rho_1 = 0.9$, $\rho_2 = 0.89$, and $\rho_3 < 0.35$ becomes greater than the first eigenvalue of \mathbf{R}_{k_2} with $\rho_1 = 0.1$ and $\rho_2 = 0.09$, and $\rho_3 = 0.05$.

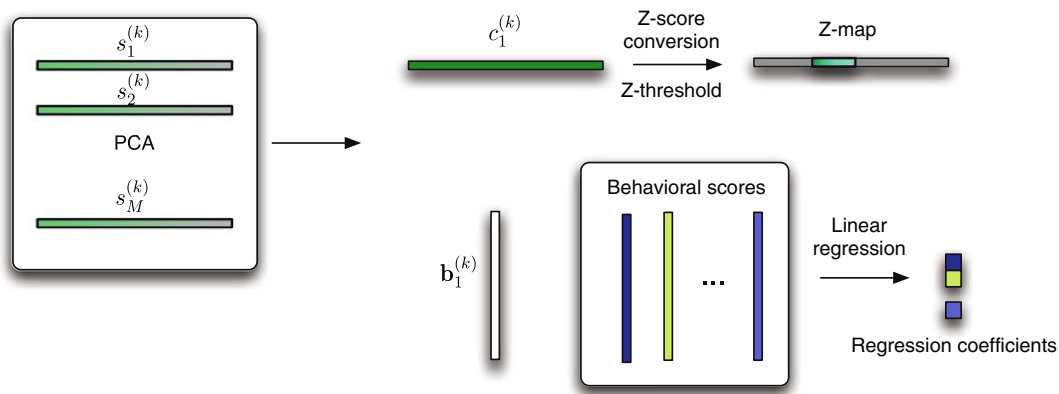


Fig. 3 Illustration of second level PCA to obtain group level source estimate and behavioral association study by multiple linear regression between the first principal eigenvector and the behavioral score vectors.

correlation matrix, where $M_1 = 13$ is the number of sources from the dominant sub-group and $M_2 = 3$ is the number of outliers forming the second sub-group. For $\rho_1 = 0.1, \dots, 0.9$ and $\rho_2 = \rho_1 - 0.01$, we calculate λ_1 and λ_2 of \mathbf{R}_k at different values of ρ_3 . Figure 2 shows the first eigenvalue (λ_1) and the second eigenvalue (λ_2) depending on different configurations of ρ_1 , ρ_2 , and ρ_3 . It is observed that the difference between the first and second eigenvalues of \mathbf{R}_k reduces as ρ_3 decreases. The second eigenvalue of \mathbf{R}_{k_1} with $\rho_1 = 0.9$, $\rho_2 = 0.89$, and $\rho_3 < 0.35$ becomes equal to or greater than the first eigenvalue of \mathbf{R}_{k_2} with $\rho_1 = 0.1$, $\rho_2 = 0.09$, and $\rho_3 = 0.05$. Therefore, when $\{s_1^{(k_1)}, s_2^{(k_1)}, \dots, s_M^{(k_1)}\}$ have heterogeneous correlation values, i.e., when ρ_3 is low, PCA on concatenated datasets places the secondary feature of a heterogeneous group (\mathbf{R}_{k_1}) over the first feature of the subsequent group (\mathbf{R}_{k_2}), resulting degraded separation of the sources in group k_2 . On the other hand, M-CCA provides reliable separation performance because of the orthogonality constraint imposed among the estimated sources within each dataset. Because the secondary or subsequent features in \mathbf{R}_{k_1} are prevented by the orthogonality constraint after the sources in group k_1 are estimated. We demonstrate the performance difference between M-CCA and group ICA using PCA on concatenated dataset in Section 4.1.

2.3 Second Level Analysis

In the second level analysis, PCA is applied to each group of corresponding sources estimated by M-CCA. The first principal component, i.e., the component with the largest variance is taken as the summary activation map for the group. The eigenvector corresponding to the first principal component is taken as the cross-

subject variation vector for this component. The procedure can be written as

$$\mathbf{s}^{(k)} = \mathbf{b}_1^{(k)} c_1^{(k)} + \sum_{m=2}^M \mathbf{b}_m^{(k)} c_m^{(k)} \tag{2}$$

where $\mathbf{s}^{(k)}$ is an $M \times 1$ vector that consists of $\{s_1^{(k)}, s_2^{(k)}, \dots, s_M^{(k)}\}$, $\mathbf{b}_1^{(k)}$ is the eigenvector corresponding to the largest eigenvalue of the correlation matrix $E\{\mathbf{s}^{(k)}(\mathbf{s}^{(k)})^T\}$. When the first pair of eigenvector and principal component $\mathbf{b}_1^{(k)} c_1^{(k)}$ is taken to represent $\mathbf{s}^{(k)}$, it corresponds to a rank-one approximation of the data matrix of $\mathbf{s}^{(k)}$. The percentage of variance explained by the approximation is the ratio between the largest eigenvalue and the sum of all the eigenvalues.

The summary activation map $c_1^{(k)}$ is converted to a Z-score map by dividing the voxel value by the standard deviation of the voxel values for the entire map, and thresholded at $Z = 1.5$. The supra-thresholded voxels are considered as the activated area. The eigenvector $\mathbf{b}_1^{(k)}$ represents the magnitude modulation of the summary activation map across the M subjects and a linear regression is performed on this modulation vector against the selected behavioral score vectors recorded during the driving task. The second level PCA procedure is illustrated in Fig. 3.

3 Methods

3.1 Simulations

We generate simulated sources as the latent sources for the group datasets. Each source set contains four 60×60 pixel images with a superimposed focus simulating typical activation patterns estimated from fMRI data. The activation regions on the simulated image

sources are originally located at the center of each quadrant of the image, as show in Fig. 4a. The activation region is then randomly dislocated around the original location, as illustrated in Fig. 4b, to form the group of corresponding sources. Different degrees of dislocation are arranged for each group of corresponding sources such that the average within-group source correlation is in decreasing order for the four groups.

To test the performance of the algorithms on jointly separating different number of sources, we also generate random sources sampled from a Laplacian distribution. Across different groups of random sources, the within-group correlation values are also arranged in decreasing order.

One source is selected from each group of corresponding simulated image sources and random sources, to form the set of K latent sources. The K sources are then mixed by a randomly generated $K \times K$ nonsingular mixing matrix to form a simulated dataset.

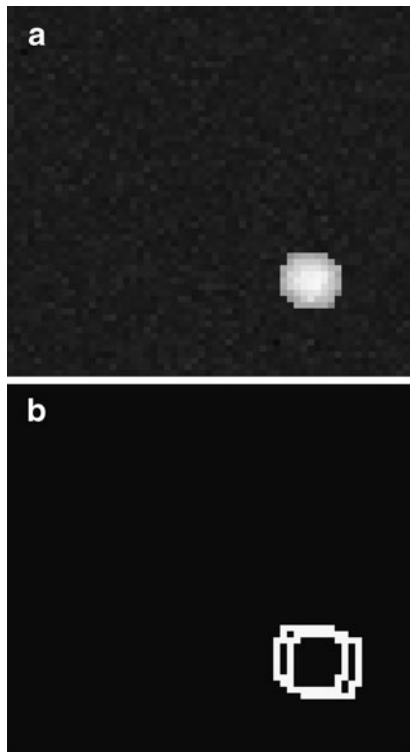


Fig. 4 Illustration of a simulated image source. **a** The activation region is originally located at the center of one quadrant of the image. **b** The activation region is then randomly dislocated around the original location to simulate the spatial variation of brain activation due to subject differences.

We test the joint blind source separation of MCCA and Group ICA [10] on three cases where the group contains $M = 10$, $M = 20$, $M = 40$, and $M = 80$ datasets. As an example the correlation matrix of the simulated sources from $M = 20$ datasets is shown in Fig. 5. In accordance with our generative model, the simulated latent sources in the group assume block diagonal correlation structure.

3.2 Subjects

The study included forty, healthy, right-handed men ($N = 20$) and women ($N = 20$), with a mean age of 24.75 ± 4.7 years. Potential participants were screened to eliminate those with any neurological disorder or DSM IV-TR Axis I psychiatric disorder including a history of substance abuse. At each visit, participants underwent a urine drug screen to test for drugs and pregnancy in women. All participants were nonsmokers, had good visual acuity without correction, valid driver licenses, good driving records assessed by self report. Participants were trained for 10 min. on the simulated driving paradigm.

3.3 Experiment Design

The driving simulator used custom built in-house software. The hardware setup was designed to mimic a realistic driving experience with a steering wheel, accelerator and brake pedals. All ferromagnetic components in this hardware were replaced with nonferromagnetic (plastic or copper-beryllium) parts. The controller itself was connected to a computer outside the scanner room through a waveguide in the wall. An LCD projector (SHARP XG-P25X) outside the scanner room and behind the scanner projected through a second waveguide to a translucent screen, which the subjects saw via a mirror attached to the head coil of the 3 Tesla MRI scanner (Allegra; Siemens, Erlangen, Germany). The screen subtended approximately a 25° field of view (FOV) that provided the subject with a straight line of sight.

The functional paradigm was blocked and consisted of three epochs, fixation (30 s), driving (90 s) and watching (60 s), each repeated three times and ending with a fixation epoch. During the fixation phase, the subject was instructed to focus on a + sign. During the driving block, the participant was asked to drive the car normally and safely, abide by all conventional traffic rules (posted speed limit [40 mph], stop signs, yielding for pedestrians, staying in the right lane except to pass). During the watching phase, the subject passively viewed a simulated driving scene. A screenshot

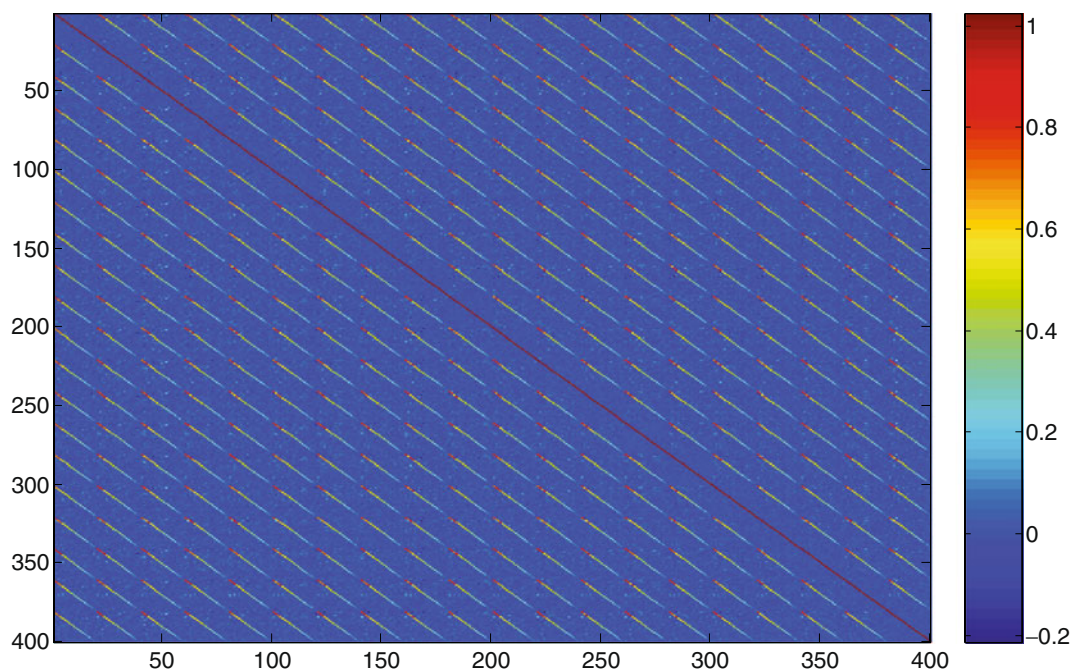


Fig. 5 Block diagonal correlation structure of the simulate sources for joint source separation: $M = 20$, $P = 20$. Blocks on the main diagonal are identity matrices, indicating that the generated sources within each dataset are uncorrelated and have unit variances. The block matrices at *upper* and *lower triangle*

are approximately *diagonal matrices*, indicating that the sources across different datasets have correlation only at *corresponding* indices. Each group of cross-correlated sources form a group of corresponding sources.

of the driving scene is shown in Fig. 6. The watching phase is the same for all subjects, i.e., a playback of a previously recorded driving session. Each run was repeated three times to increase the signal to noise ratio (SNR).

3.4 Data Acquisition

Driving behavior Continuous behavioral variables including, passenger side white line crossings, median/yellow line crossings, opposite side white line crossings, crashes, speed, and steering weave were recorded during the driving phase of the experiment.

Imaging Functional data were acquired on a Siemens Allegra 3T scanner at the Olin Neuropsychiatry Research Center at the Institute of Living using an echoplanar sequence using the following imaging parameters; repeat time (TR) = 1500 ms, echo time (TE) = 27 ms, field of view (FOV) = 22 cm, flip angle = 70° , acquisition matrix = 64×64 , voxel size = 3.44×3.44 , slice thickness = 5 mm, number of slices = 29, as-

pending acquisition. The scanner was equipped with 40 mT/m gradients and a standard quadrature head coil. To achieve longitudinal equilibrium, six dummy scans were performed at the beginning and discarded prior to analysis. Scanning was automatically triggered by the paradigm.

3.5 Data Processing

Behavior data Eight behavioral scores are constructed for each subject and running session: average speed (AS), differential of speed (DS), duration of over-speed-limit driving (OS), average steering weave (AR), differential of steering weave (DR), differential of pedal offset (DP), times of yellow line crossing (YLC) and times of white passenger-side line crossing (WPLC). Each behavioral score is normalized for all the datasets to make different behavioral scores commensurate. Other behavioral records such as the driver-side white line crossing, hit on curb, and collision are eliminated because of no occurrence or very rare occurrence.



Fig. 6 A typical screen shot taken from the driving simulation program.

fMRI data All images were preprocessed using the SPM2 Software (<http://www.fil.ion.ucl.ac.uk/spm/software/spm2/>). Motion correction was achieved using INRIAAlign [16, 17] to compensate for movement in the fMRI time series images. Each subjects' movement parameters were screened to ensure absence of excessive head motion. Motion corrected images were then spatially normalized to Montreal Neurological Institute (MNI) space by matching them to the standardized EPI template image in SPM. After spatial normalization, images were spatially smoothed by Gaussian kernel with $10 \times 10 \times 10$ mm full-width-half-maximum.

Before statistical analysis, the global signal in each dataset is removed. The global signals include the global temporal mean signal and the global spatial mean signal. Since localized brain activations account for relatively lower variance compared with the global fluctuations in fMRI, removing global signal improves

the sensitivity of subsequent multivariate analyses. Each dataset is dimension reduced by PCA and the dimension of signal subspace, i.e., the reduced number of principal components, is determined by information theoretic criteria with data sample dependence correction [18]. The information theoretic criteria such as Akaike information criterion (AIC) [19] and Minimum description length (MDL) [20] establish a measure in terms of the total information conveyed by the observations (in a likelihood sense) and balance the quantity by the number of free parameters used to describe the data through likelihood so that no overfitting will occur. In fMRI analysis, we use them to select the optimal number of principal components spanning the subspace of fMRI signals, embedded in multivariate Gaussian noise with uniform variances. To facilitate the group analysis, the maximum order selected by the MDL criterion is applied to all datasets, resulting in thirty-one

principal components retained for each dataset. One hundred and twenty datasets (forty subjects with three sessions for each subject) are analyzed by M-CCA algorithm to estimate thirty-one components for each dataset.

In the second level analysis, the first principal component in each group of corresponding source estimates is deemed to be the group level activation map. The group time course is calculated using the same weights as the group level activation maps and is linear regressed with the driving and watching paradigm to identify whether the identified source is driving or watching related. The first eigenvector is deemed to be the cross-subject variation modulation vector and is linear regressed with the eight selected behavioral factors.

4 Results

4.1 Simulated Datasets

Figure 7 shows a performance comparison of M-CCA and Group ICA for joint separation of sources with increasing number of datasets. When the number of datasets is small ($M = 10$) both M-CCA and group ICA achieves good performance on joint source separation. The separation performance of Group ICA degrades as the number of datasets increases ($M = 20, 40, 80$) while M-CCA is not affected.

As an example, Fig. 8 shows the estimated global matrix estimated by M-CCA and Group ICA. In source separation simulation, the global matrix is the product

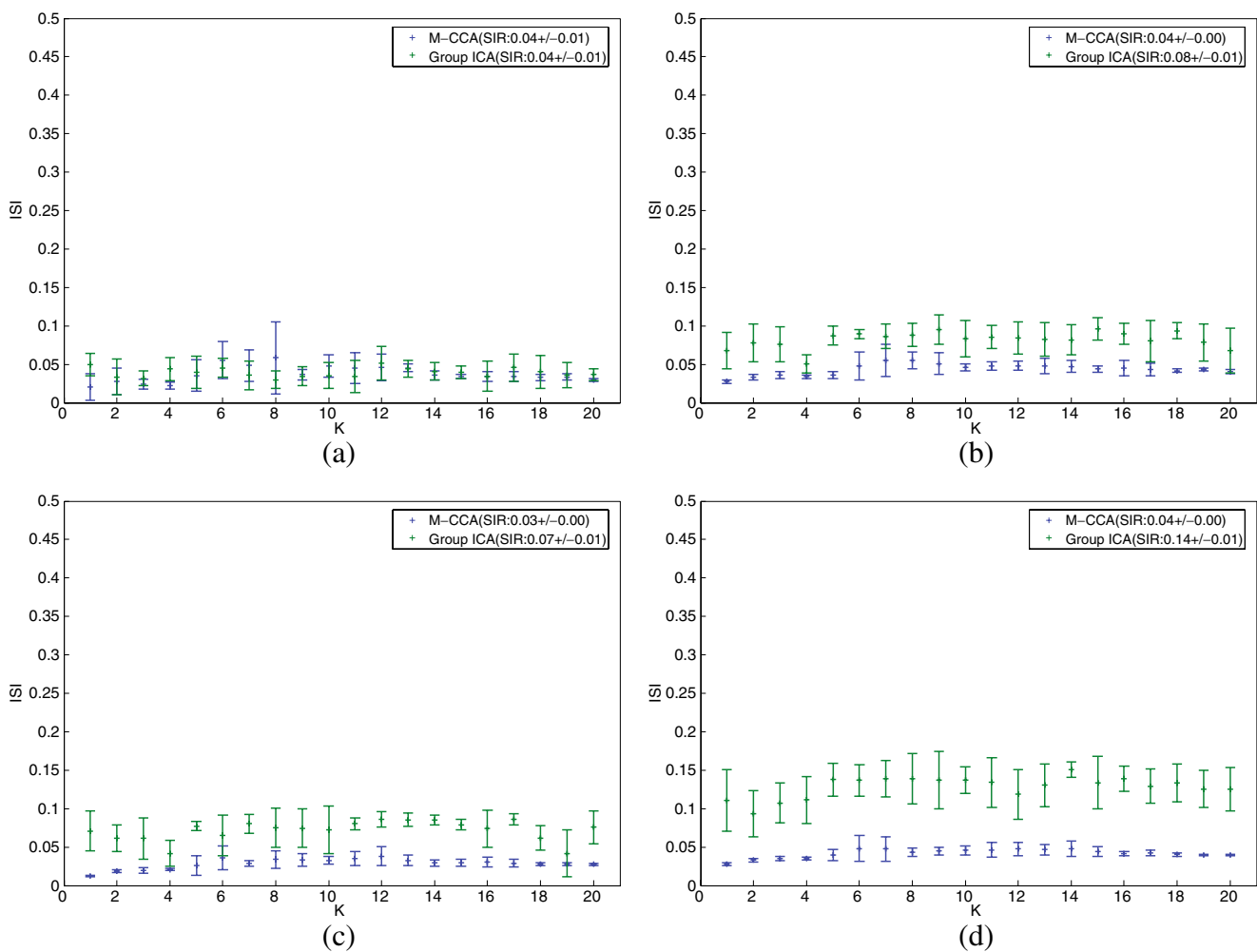


Fig. 7 Comparison of source separation performance between M-CCA and Group ICA. The separation performance is measured by inter-symbol interference (ISI) [26] of each jointly separated source. Four cases are studied with increasing number of datasets **a** $M = 10$, **b** $M = 20$, **c** $M = 40$, and **d** $M = 80$. When

the number of datasets is small ($M = 10$) both M-CCA and group ICA achieves good performance on joint source separation. The separation performance of Group ICA degrades as the number of datasets increases ($M = 20, 40, 80$) while M-CCA is not affected.

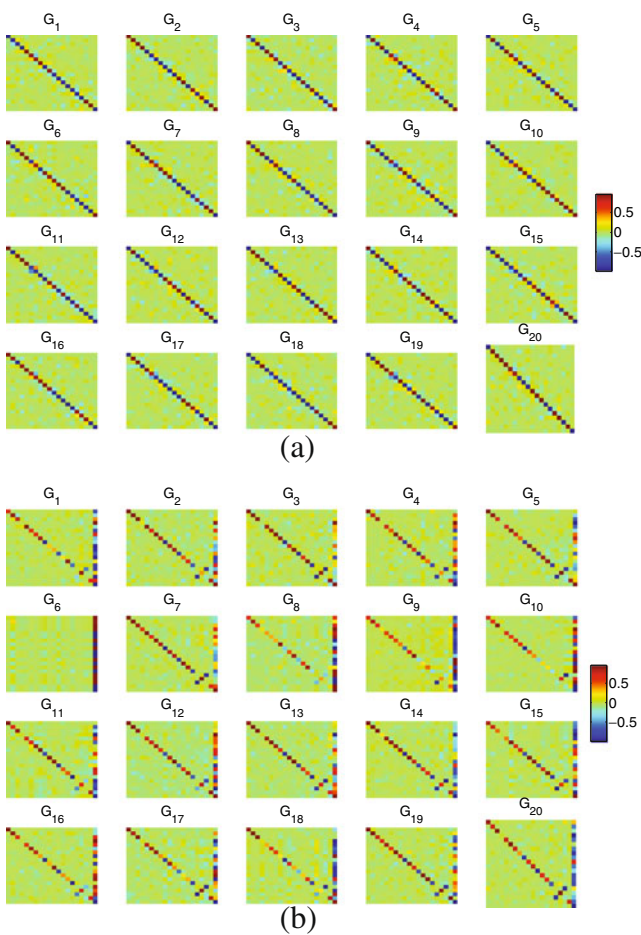


Fig. 8 Estimated global matrix \mathbf{G} of each dataset by **a** M-CCA and **b** Group ICA for the case of $M = 20$ and $P = 20$. Because Group ICA uses PCA on concatenated datasets for group dimension reduction, the last group of corresponding sources are not robustly separated, as shown in panel **b**. This is due to the interference between their first principal component and the secondary principal components of the previous source groups, as explain in Section 2.3. In contrast, M-CCA can robustly separate the last group of corresponding sources, as shown in panel **a** because of the orthogonality constraints are imposed on source estimates for each individual dataset.

of the estimated demixing matrix and the true mixing matrix. The global matrix is close to identity matrix in case of correct source separation. It is observed that, in this instance, Group ICA fails to separate sources in dataset 6, and for all the datasets, the last group of sources are not well separated from other sources. Because the last group of sources have relatively low cross-correlation, when PCA is applied to concatenation of datasets the eigenvalue of its correlation matrix is superseded by the secondary eigenvalue of the previous groups, as discussed in Section 2.2. There-

fore, for methods based on PCA on concatenation of datasets, *e.g.*, Group ICA, the last group of sources can not be properly estimated. On the other hand, M-CCA extracts sources from each individual dataset and imposes orthogonality constraints among the extracted sources. In this way, the extraction of sources with low cross-correlation is not interfered by the sources from previous groups.

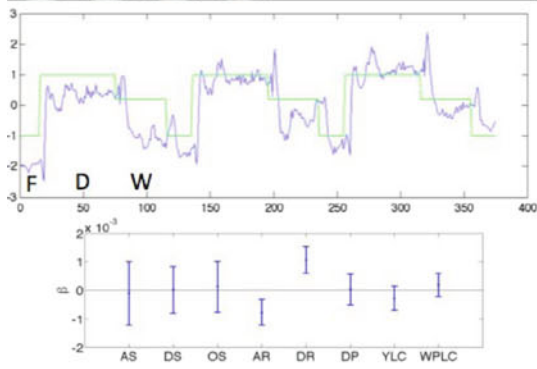
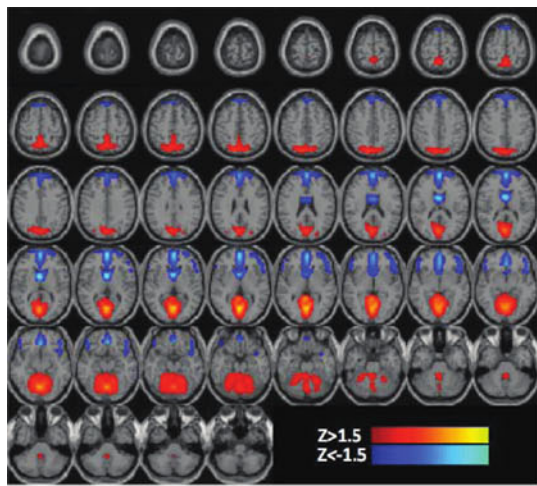
4.2 Estimated Sources of Interest (SOIs) from Simulated Driving fMRI Data

Because M-CCA jointly extract sources from each dataset such that the cross-dataset correlation among the extracted source is maximized, there is a natural order of the source estimates—an decreasing order on cross-dataset source correlation. Among the thirty-one group of jointly extracted sources, we present four SOIs based on the activation areas shown on their Z-maps. That is, we pick up those sources whose group maps show localized activations corresponding to functional brain regions.

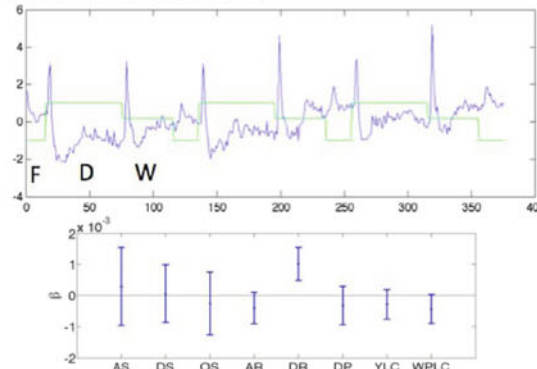
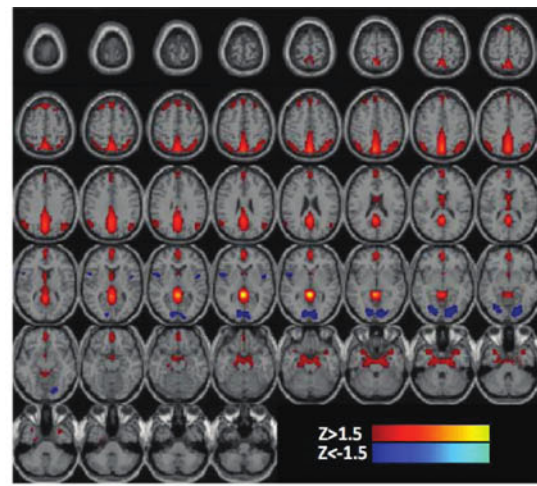
The first identified source of interest (SOI) is shown in Fig. 9a. The estimated spatial pattern has low inter-subject variability and shows correlation in parieto-occipital regions and anti-correlation in medial frontal regions. The time course of this component has high regression coefficients with the driving paradigm (0.48 ± 0.20)—indicating that the frontal and parieto-occipital brain regions are consistently active across all the subjects when performing the driving task. The observation that the frontal region is anti-correlated with the driving task agrees with a previous driving study in [8] (Blue component), *i.e.*, the frontal region is part of the “default mode network”, and is decreased as a driving task is performed.

The second SOI is shown in Fig. 9b. The active areas comprises the default model network (DMN) [21, 22]. Typically, the activation of DMN is suppressed when a cognitive task is performed. In a simulated driving experiment, driving is the dominant task, therefore, the activations in DMN is suppressed compared with watching and fixation. As a result, the time course of this component has a negative regression coefficient with the driving paradigm.

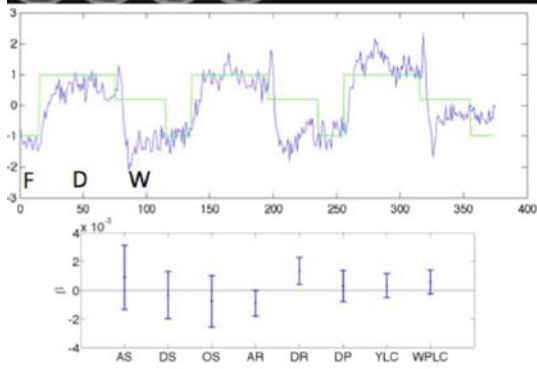
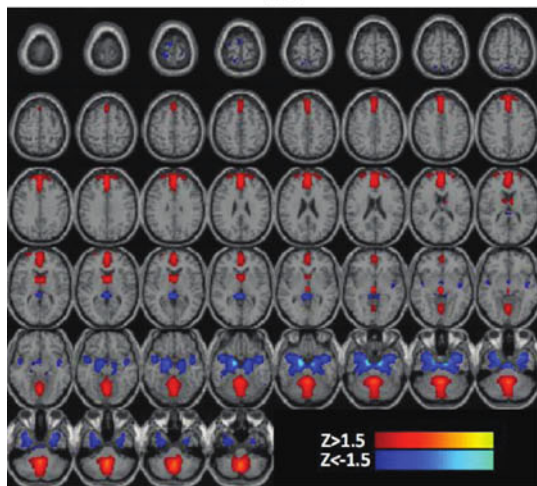
The third SOI is shown in Fig. 9c. Its spatial pattern shows correlation in medial frontal areas and cerebellum. The time course is significantly correlated with the driving paradigm. This is consistent with the involvement of the medial frontal regions in top down executive control, (thinking about driving) and with cerebellar regions being involved in the motor function during driving.



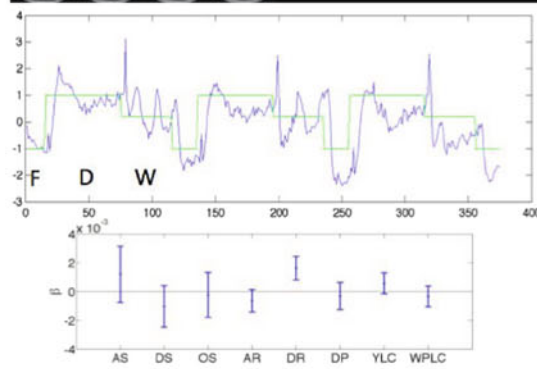
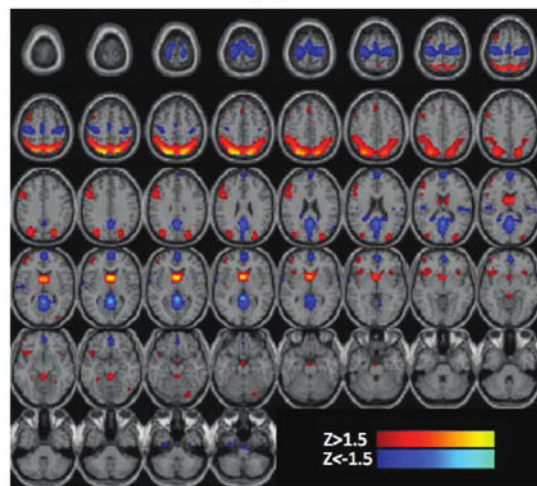
(a)



(b)



(c)



(d)

◀ **Fig. 9** For each SOI, the group activation map is shown on the *top*, mean time course overlaid with the paradigm for three repetitions of [F]ixation-[D]riving-[W]atching in the *middle*, and confidence interval (CI) of behavioral correlation at *bottom*. **a** SOI-1 shows correlation in parieto-occipital regions and anti-correlation in medial frontal regions. The time course has high regression coefficients associated with the driving paradigm (0.48 ± 0.20)—indicating that the frontal and parieto-occipital brain regions are highly consistent across all the subjects when performing the driving task. Among the eight behavioral factors defined in Section 3.5, this component has significant association with the average and differential of steering weave. **b** SOI-2 shows the default model network (DMN). Typically, DMN is suppressed when a cognitive task is performed. As a result, the time course of this component has a negative regression coefficient with the driving paradigm. **c** SOI-3 shows correlations in medial frontal areas and also in the cerebellum. The time course is significantly correlated with the driving paradigm. This is consistent with the involvement of the medial frontal regions in top down executive control, (thinking about driving) and with cerebellar regions being involved in the motor function during driving. **d** SOI-4 shows correlation at parietal lobe, anterior cingulate, and anti-correlation at motor cortex and posterior cingulate. The time course of this component is correlated with both driving and watching paradigm. The activation of SOI-4 contains multiple brain regions and the cross-dataset consistency of SOI-4 is the lowest among the four presented SOIs. The observation suggest that the cross-subject and cross-session variability becomes significant for brain regions involved with functions such as visual integration and attention (parietal lobe), high level motor control (pre- and post-central cortices).

The last SOI is shown in Fig. 9d. Its spatial pattern shows correlation at parietal lobe, anterior cingulate, and anti-correlation at motor cortex and posterior cin-

gulate. The time course is correlated with both driving and watching paradigm. The spatial map of this SOI contains multiple brain regions and the cross-dataset consistency of this SOI is the lowest among the four presented SOIs. The observation suggest that the cross-subject and cross-session variability becomes significant for brain regions involved with functions such as visual integration and attention (parietal lobe), high level motor control (pre- and post-central cortices). It's also interesting to observe a counter relationship between anterior and posterior cingulate.

4.3 Association of the Brain Activations with the Driving Behavior

Rank-one approximation is applied to each group of SOIs. The percentage of variance represented by the first principal component in each group is calculated to be: 0.89, 0.83, 0.76, and 0.73.

The cross-subject covariation of each group is regressed with eight behavioral factors as described in Section 3.5. The estimated 95% confidence interval of the regression coefficients of each behavioral factor is displayed in the bottom panel in Fig. 9a–d.

Study with the behavioral factors indicates that the variations of all four SOIs are significantly related to the differential of the steering weave, i.e., the frequency of steering manipulation—a very active behavioral variable involved in driving task, primarily involved with

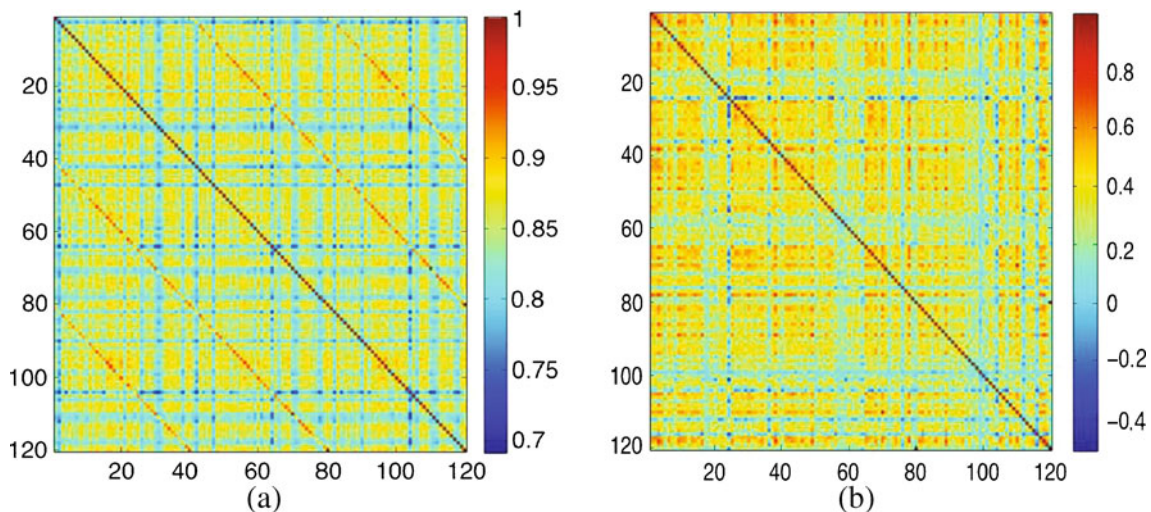


Fig. 10 **a** Spatial correlation matrix and **b** Temporal correlation matrix of the group of corresponding sources of SOI-1. It is observed that the spatial maps estimated from the three sessions of the same subject assume higher correlation than the correlation with other subjects, i.e., higher correlation values appear on the *diagonals* in each 40×40 matrix blocks. Since M-CCA treats

fMRI data of each subject at each session as an independent dataset, this correlation pattern reveals that the brain activation pattern is more consistent within the subject (across different running sessions) than across the different subjects. The estimated time course has lower degrees of correlation compared with the spatial maps.

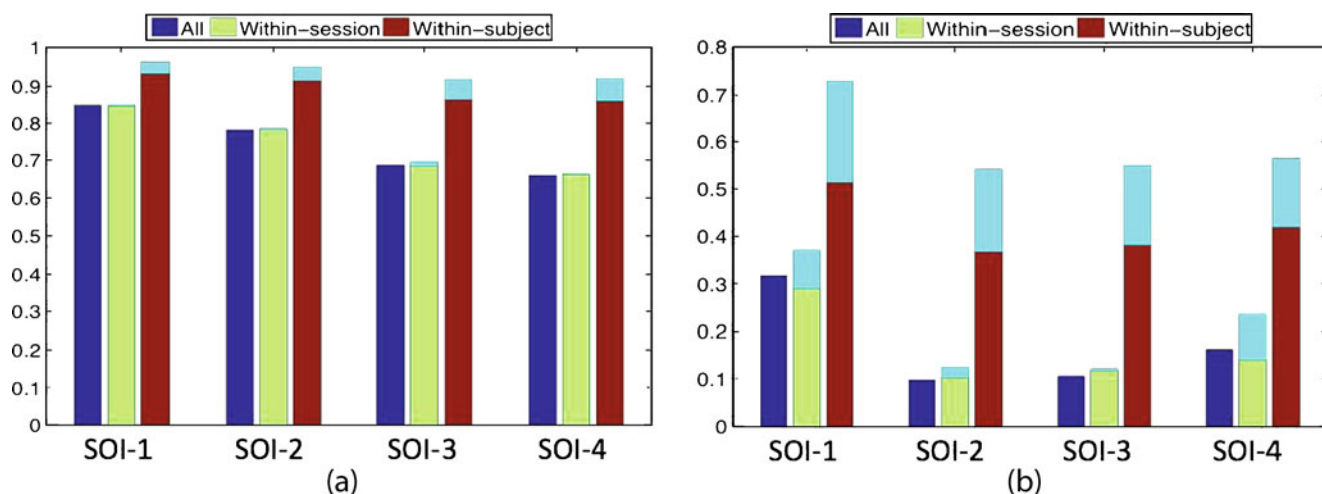


Fig. 11 Comparison of **a** spatial map correlation and **b** time course correlation of the four SOIs among all datasets, within-sessions, and within-subjects, the standard deviation is stacked onto the mean value of each case. Although the time courses

show lower level of cross-correlation compared with spatial maps, the trend of high within-subject correlation is consistently observed.

cognitive control such as online monitoring and adjustment during driving.

SOIs 1 and 3 also show significant association with the average steering weave, which is a measure on the overall extend that the car is changing its direction—this could be due to lane changing or turning direction. This manipulation involves high level decision making, attention, and possibly visual and motor functions.

4.4 The Source Correlation Structure Estimated by M-CCA

Different from the existing group analysis methods, M-CCA estimates the group of sources in the order of cross-correlation among the corresponding sources, i.e., the sum of all the pairwise source correlation values. Therefore, a natural ranking in terms of the cross-dataset consistency is imposed during the group analysis.

M-CCA reveals the correlation structure of each group of corresponding sources. This correlation structure has been used, through a rank-one decomposition, in the second level behavioral association study.

Furthermore, as an example, we plot the cross-correlation pattern of the spatial maps and time courses for SOI-1 in Fig. 10. It is observed that the spatial maps extracted from three sessions of the same

subject assume higher spatial correlation, i.e., higher correlation values appear on the diagonals in each 40×40 matrix blocks in Fig. 10a. Since M-CCA treats fMRI data of each subject at each session as an independent dataset, this correlation pattern reveals that the functional correlation is more consistent within a subject (across different running sessions) than across different subjects. In the temporal domain, the cross-correlation of time courses are significantly lower than that of the spatial maps as shown in Fig. 10b. However, the pattern of higher within-subject correlation than cross-subject correlation is consistent in spatial and temporal domain, as presented in Fig. 11 for all four SOIs.

5 Discussion

5.1 Contributions of each Brain Network to the Driving Task

Since M-CCA is based on the principle of correlation maximization, the components obtained are not as sparse as the independent components in ICA. Therefore, multiple activation regions could co-exist in the same component. However, these regions together compose the spatial map that are arranged in order of decreasing cross-dataset correlation.

Table 1 List of activation regions in each M-CCA component, the time course regression coefficients with driving and watching paradigm, and their confidence intervals (CI).

Component	SOI-1	SOI-2	SOI-3	SOI-4
β_D	0.85 CI: [0.75,0.94]	-0.63 CI: [-0.75, -0.52]	0.86 CI: [0.77,0.94]	0.64 CI: [0.53,0.75]
β_W	0.42 CI: [0.33,0.52]	-0.39 CI: [-0.51, -0.28]	0.15 CI: [0.07,0.23]	0.61 CI: [0.50,0.72]
Parieto-occipital sulcus	+			
Parietal lobe				+
Medial frontal			+	
Inferior frontal	-			
Motor cortex				-
Cerebellum			+	
Superior parietal				
Inferior parietal		+		
Anterior cingulate				+
Posterior cingulate		+		-

Table 1 summarizes the activation regions for each SOI, with the Brodmann labeling shown in Table 2. As a reference, time course regression coefficients of driving (β_D) and watching (β_W) are also presented in the table for each component. The four selected SOIs all have significant correlation with driving, indicating that these functionally correlated regions are driving related, hence, their variation across subjects can be used for the association study with the driving behavior scores.

It is observed that the first source has the highest regression coefficient with the driving paradigm and demonstrates highly consistent spatial activations. This observation suggests that M-CCA, as a second-order based method, reveals the variance in the fMRI datasets imposed by the experimental design, which is common across all the datasets [23]. The co-existence of positive and negative correlations as a brain network

is similarly observed in the results of model-driven analysis [1]. The new observation in this work suggests that the inferior frontal lobe, which involves high level executive functions tends to be suppressed during low level online visuomotor guidance, i.e., when the parieto-occipital cortex are active. Similar observation is reported in [24].

SOI-2 includes major activation regions at inferior parietal and posterior cingulate cortices. The time course of this source has negative regression coefficients with the driving task, which agrees with the fact that the default mode network is decreased when performing a task [21, 22]. However, it is worth noting that the different parts of this brain default network are also functionally specialized [25], which explains the association of the posterior cingulate cortex with the parietal, motor, and anterior cingulate cortices in forming a separate system shown in Fig. 9d. The overlapping

Table 2 List of activation regions in each M-CCA component and their anatomical label, Brodmann area label, volume, and peak values, and their Talairach coordinates.

Component	Anatomical region	Brodmann area	Volume (L/R)	Peak Z-score and coordinates (L/R)
SOI-1	Cuneus	17, 18	9.9/8.4	2.5(-9, -99, 8)/2.1(3, -96, 5)
	Lingual gyrus	17, 18	7.8/6.7	2.3(-3, -79, 1)/2.1(3, -79, 1)
SOI-2	Precuneus	7	13.6/11.0	4.6(0, -71, 45)/3.2(6, -70, 50)
	Cingulate gyrus	31	5.6/1.9	4.1(0, -51, 27)/3.0(6, -54, 28)
SOI-3	Medial frontal gyrus	9	6.7/6.0	2.8(-6, 42, 20)/2.7(6, 42, 20)
	Culmen	-	6.3/6.0	3.2(-6, -50, -15)/3.5(0, -53, -18)
SOI-4	Precuneus	7	10.4/9.7	4.6(-21, -70, 50)/3.9(18, -70, 48)
	Inferior parietal lobule	40	8.4/5.0	3.2(-39, -50, 52)/4.3(18, -67, 53)
	Precentral gyrus	6, 4	3.9/3.7	-2.9(-33, -17, 67)/-2.8(30, -17, 67)
	Posterior cingulate	29, 30, 31	3.7/5.6	-3.6(0, -46, 22)/-3.4(3, -52, 11)

of the activation regions indicates that a single region is part of more than one network which is also observed with ICA based methods [8].

Since BSS methods assumes a linear mixture model which is more flexible than the conventional models such as the general linear model, it is better fitted into the scenario where exploratory interpretation of the data is desired. In this study, M-CCA achieves source separation based on cross-dataset consistency and hence efficiently utilizes the statistical power in group dataset.

5.2 Different Ways to Associate Neural Correlates with Behavioral Factors

In a cognitive experiment, it is always intriguing to reveal the link between the neural correlates with behavioral factors. There are two typical ways of exploring this association.

In one way, behavioral factors recorded during the cognitive experiment can be used to design regressors to infer the associated neural correlates by model-driven analysis such as GLM [3]. In this way, the resulting neural correlates are specific to each dataset. A group inference can be drawn from all the neural correlates identified from each dataset. Although, this method requires precise record of the behavior and assumes that the behavior records can be used to explain the brain activation through an event-related design with the modulation of hemodynamic response.

As an alternative approach, the association can be identified by the co-variation of the activation levels with the behavioral scores [4, 5, 8]. In this approach, the underlying hypothesis is that the engagement of functional networks is concordant to behavioral factors and the hypothesis is tested by linear regression of the functional activation level with the summary behavioral factors such as number of event occurrence, driving speed, and vehicle operation parameters as studied in Section 4.3. This approach is suitable for data-driven analysis since no response model needs to be assumed.

5.3 Conclusion and Future Work

In this work, we use a novel statistical method, M-CCA, to study brain functions and functional-behavioral associations during simulated driving task. The M-CCA method jointly estimates brain activations from

large number of fMRI datasets based on the cross-correlation structures of the underlying sources. In addition, the association of functional and behavioral factors is revealed by regressing cross-dataset activation variations with the recorded behavioral scores. The proposed framework is generally applicable to group analysis of multi-dimensional datasets for study of multiple features and their variations. The extension of intra-class correlation model in Section 2.2 provides a model of the heterogeneity within a group of corresponding sources, hence, is potentially useful in modeling and comparing contrasting sub-group structures such as healthy controls and patients.

Acknowledgements This research is supported in part by the NIH grant R01 EB 000840.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Walter, H., Vetter, S. C., Grothe, J., Wunderlich, A. P., Hahn, S., & Spitzer, M. (2001). The neural correlates of driving. *NeuroReport*, *12*, 1763–1767.
2. Graydon, F. X., Young, R., Benton, M. D., Genik II, R. J., Posse, S., Hsieh, L., et al. (2004). Visual event detection during simulated driving: Identifying the neural correlates with functional neuroimaging. *Transportation Research*, *F7*, 271–286.
3. Spiers, H. J., & Magurie, E. A. (2007). Neural substrates of driving behaviour. *NeuroImage*, *36*, 245–255.
4. Uchiyama, Y., Ebe, K., Kozato, A., Okada, T., & Sadato, N. (2003). The neural substrates of driving at a safe distance: A functional MRI study. *Neuroscience Letters*, *352*, 199–202.
5. Jeong, M., Tashiro, M., Singh, L. N., Yamaguchi, K., Horikawa, E., Miyake, M., et al. (2006). Functional brain mapping of actual car-driving using [18f]FDG-PET. *Annals of Nuclear Medicine*, *20*, 623–28.
6. McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J., et al. (1998). Analysis of fMRI data by blind separation into independent components. *Human Brain Mapping*, *6*, 160–188.
7. Friman, O., Borga, M., Lundberg, P., & Knutsson, H. (2002). Exploratory fMRI analysis by autocorrelation maximization. *NeuroImage*, *16*, 454–64.
8. Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., & Pearlson, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Human Brain Mapping*, *16*, 158–167.
9. Meda, S. A., Calhoun, V. D., Astur, R. S., Turner, B. M., Ruopp, K., & Pearlson, G. D. (2009). Alcohol dose effects

- on brain circuits during simulated driving: An fmri study. *Human Brain Mapping*, 30, 1257.
10. Calhoun, V. D., Adahi, T., Pekar, J. J., & Pearlson, G. D. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14, 140–151.
 11. Beckmann, C. F., & Smith, S. M. (2005). Tensorial extensions of independent component analysis for group fMRI data analysis. *NeuroImage*, 25, 294–311.
 12. Guo, Y., & Pagnoni, G. (2008). A unified framework for group independent component analysis for multi-subject fMRI data. *NeuroImage*, 42, 1078–1093.
 13. Kim, T., Attias, H. T., Lee, S. Y., & Lee, T. W. (2007). Blind source separation exploiting higher-order frequency dependencies. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 15, 70–79.
 14. Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika*, 58, 433–51.
 15. Li, Y.-O., Wang, W., Adahi, T., Calhoun, V. D. (2009). Joint blind source separation by multi-set canonical correlation analysis. *IEEE Transactions on Signal Processing*, 57, 3918–3929.
 16. Freire, L., & Mangin, J. F. (2001). Motion correction algorithms may create spurious activations in the absence of subject motion. *NeuroImage*, 14, 709–722.
 17. Freire, L., Roche, A., & Mangin, J. F. (2002). What is the best similarity measure for motion correction in fMRI time series? *IEEE Transactions on Medical Imaging*, 21, 470–484.
 18. Li, Y.-O., Adahi, T., & Calhoun, V. D. (2007). Estimating the number of independent components in fMRI data. *Human Brain Mapping*, 28, 1251–1266.
 19. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proc. 2nd international symposium on information theory*. Tsahkadsor, Armenian SSR; Hungary.
 20. Rissanen, J. (1978). Modeling by the shortest data description. *Automatica*, 14, 465–471.
 21. Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98, 676–682.
 22. Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, 16, 1484–1492.
 23. Viviani, R., Grön, G., & Spitzer, M. (2005). Functional principal component analysis of fMRI data. *Human Brain Mapping*, 24, 109–129.
 24. Galletti, C., Kutz, D. F., Gamberini, M., Breveglieri, R., & Fattori, P. (2003). Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Experimental Brain Research*, 153, 158–170.
 25. Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., & Fox, P. T. (2009). Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *Journal of Neuroscience*, 29, 14496–14505.
 26. Amari, S., Cichocki, A., & Yang, H. H. (1996). A new learning algorithm for blind signal separation. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in neural information processing systems* (Vol. 8, pp. 757–763). Cambridge, MA: MIT Press.



Yi-Ou Li received his B.S. degree in electrical engineering from Beijing University of Posts and Telecommunications. He graduated from the Machine Learning for Signal Processing Laboratory in University of Maryland Baltimore County with a Ph.D. degree in electrical engineering. He joined the Neural Connectivity Lab in the department of Radiology and Biomedical Imaging at University of California San Francisco in July 2009. He currently works on machine learning for medical imaging data analysis to study traumatic brain injury. His research interests include statistical signal processing, multivariate analysis, quantitative evaluation of data-driven methods, and data analytic applications.



Tom Eichele went to medical school in Greifswald, Germany and became a licensed physician in 2004. He received a Ph.D. in biological and medical psychology at the University of Bergen, Norway in 2007. He is currently working as a post-doctoral researcher at the Bergen fMRI group (<http://fmri.uib.no>) and teaches psychology undergraduates in cognitive electrophysiology and functional neuroanatomy. He also is a visiting fellow at the Mind Research Network in Albuquerque, NM. Dr. Eichele is interested in multimodal analysis of neuroimaging data (fMRI and EEG) on trial-by-trial level to answer questions in cognitive neuroscience, and has so far contributed to more than 30 journal articles and book chapters, and has organized workshops and sessions at major conferences. He is a member of the Organization for Human Brain Mapping (OHBM), the society for psychophysiological research (SPR) and the cognitive neuroscience society (CNS). He is a reviewer for many international journals and funding agencies.



Vince D. Calhoun received a bachelor's degree in Electrical Engineering from the University of Kansas, Lawrence, Kansas, in 1991, master's degrees in Biomedical Engineering and Information Systems from Johns Hopkins University, Baltimore, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland Baltimore County, Baltimore, in 2002. He worked as a senior research engineer at the psychiatric neuroimaging laboratory at Johns Hopkins from 1993 until 2002. He then served as the director of medical image analysis at the Olin Neuropsychiatry Research Center and as an associate professor at Yale University.

Dr. Calhoun is currently Chief Technology Officer and Director of Image Analysis and MR Research at the Mind Research Network and is a Professor in the Departments of Electrical and Computer Engineering (primary), Neurosciences, Psychiatry and Computer Science at the University of New Mexico. He is the author of more than 160 full journal articles and over 300 technical reports, abstracts and conference proceedings. Much of his career has been spent on the development of data driven approaches for the analysis of brain imaging data. He has won over \$18 million in NSF and NIH grants on the incorporation of prior information into independent component analysis (ICA) for functional magnetic resonance imaging, data fusion of multimodal imaging and genetics data, and the identification of biomarkers for disease.

Dr. Calhoun is a senior member of the IEEE, the Organization for Human Brain Mapping, the International Society for Magnetic Resonance in Medicine, and the American College of Neuropsychopharmacology. He is a chartered grant reviewer for NIH. He has organized workshops and special sessions at multiple conferences. He is currently serving on the IEEE Machine Learning for Signal Processing (MLSP) technical committee and previously served as the general chair of the 2005 meeting. He is a reviewer for many journals and is on the editorial board of the Human Brain Mapping and Neuroimage journals.



Tulay Adali received the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, in 1992 and joined the faculty at the University of Maryland Baltimore County (UMBC), Baltimore, the same year. She is currently a Professor in the Department of Computer Science and Electrical Engineering at UMBC.

Prof. Adali assisted in the organization of a number of international conferences and workshops including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), the IEEE International Workshop on Neural Networks for Signal Processing (NNSP), and the IEEE International Workshop on Machine Learning for Signal Processing (MLSP). She was the General Co-Chair, NNSP (2001–2003); Technical Chair, MLSP (2004–2008); Program Co-Chair, MLSP (2008 and 2009), 2009 International Conference on Independent Component Analysis and Source Separation; Publicity Chair, ICASSP (2000 and 2005); and Publications Co-Chair, ICASSP 2008.

Prof. Adali chaired the IEEE SPS Machine Learning for Signal Processing Technical Committee (2003–2005); Member, SPS Conference Board (1998–2006); Member, Bio Imaging and Signal Processing Technical Committee (2004–2007); and Associate Editor, IEEE Transactions on Signal Processing (2003–2006), Elsevier Signal Processing Journal (2007–2010). She is currently Chair of Technical Committee 14: Signal Analysis for Machine Intelligence of the International Association for Pattern Recognition; Member, Machine Learning for Signal Processing and Signal Processing Theory and Methods technical committees; Associate Editor, IEEE Transactions on Biomedical Engineering and Journal of Signal Processing Systems for Signal, Image, and Video Technology, and Senior Editorial Board member, IEEE Journal of Selected Areas in Signal Processing.

Prof. Adali is a Fellow of the IEEE and the AIMBE and the past recipient of an NSF CAREER Award. Her research interests are in the areas of statistical signal processing, machine learning for signal processing, and biomedical data analysis.