Statistica Sinica **18**(2008), 1233-1252

A SUPERVISED SINGULAR VALUE DECOMPOSITION FOR INDEPENDENT COMPONENT ANALYSIS OF fMRI

Ping Bai[1], Haipeng Shen[2], Xuemei Huang[3] and Young Truong[2]

[1]*PayPal Inc.,* [2]*University of North Carolina at Chapel Hill and* [3]*Pennsylvania State University*

Abstract: Functional Magnetic Resonance Imaging (fMRI) is a non-invasive technique for studying the brain activity. The data acquisition process results a temporal sequence of 3D brain images. Due to the high sensitivity of MR scanners, spikes are commonly observed in the data. Along with the temporal and spatial features of fMRI data, this artifact raises a challenging problem in the statistical analysis. In this paper, we introduce a supervised singular value decomposition technique as a data reduction step of independent component analysis (ICA), which is an effective tool for exploring spatio-temporal features in fMRI data. Two major advantages are discussed: first, the proposed method improves the robustness of ICA against spikes; second, the method uses the fMRI experimental designs to guide the fully data-driven ICA, yielding a more computationally efficient procedure and highly interpretable results. The advantages are demonstrated using spatio-temporal simulation studies as well as a data analysis.

Key words and phrases: Basis expansion, functional Magnetic Resonance Imaging, robustness, singular value decomposition, spatio-temporal data.

1. Introduction

Functional Magnetic Resonance Imaging (fMRI) is a noninvasive technique that is used to study brain functions in vivo by neuroscientists (Jezzard, Matthews and Smith (2001) and Huettel, Song and McCarthey (2004)). Since the early 1990s, it has gained growing popularity in both clinical and basic neuroscience researches, and has influenced our understanding of the neurobiology of human behavior.

In a typical fMRI experiment, functional images of the brain are recorded every few seconds while the subject is performing a task sequence or receiving a certain stimulus. The images are taken using a Magnetic Resonance (MR) sequence which is sensitive to changes in local blood oxygenation levels; hence, during a certain activation, the parts of the brain that are activated would demonstrate blood oxygenation level dependent (BOLD) signal changes similar to the time course of the activation paradigm.

A single MR image usually consists of a certain number of slices and each slice is made up of individual cuboid elements called voxels. Hence an fMRI data set can be viewed as a three dimensional matrix of voxels that is repeatedly sampled over time. Statistically, the data set is four-dimensional, and is usually represented as a spatio-temporal matrix $\mathbf{X}$ of dimension $M \times N$: each column of $\mathbf{X}$ corresponds to an fMRI image with M voxels, and each row of $\mathbf{X}$ is a time series of N time points for one voxel. In most fMRI experiments, the number of time points is far less than the number of voxels ($N \ll M$). For example, the fMRI data in Section 5 have $M = 153,594$ voxels and $N = 200$ time points.

Various statistical methods have been proposed for the analysis of fMRI data. The methods can be categorized into two complementary groups, model-driven and data-driven. Model-driven approaches include statistical parametric mapping (SPM) (Friston, Jezzard and Turner (1994) and Friston, Holmes and Worsley (1995)), which is by far the most popular fMRI analyzer, and a variety of Bayesian techniques (Genovese (2000), Gössl, Fahrmeir and Auer (2001) and Gössl, Auer and Fahrmeir (2001)). Data-driven approaches consist of cluster analysis (Goutte et al. (1999)), principal component analysis (Kherif et al. (2002)), and independent component analysis (ICA) (McKeown et al. (1998b), Petersen et al. (2000) and Calhoun et al. (2003)). More recently, SPM-ICA was proposed by Hu et al. (2005) to unify SPM, a model-driven approach, with ICA, a data-driven approach. Lange (1999) compared nine analytic methods currently used in BOLD fMRI analysis. Although some procedures outperform others in certain contexts, there is no globally optimal procedure for the analysis of a particular fMRI study.

The current paper focuses on ICA. As an exploratory data analysis tool, ICA is capable of extracting, from the recorded fMRI signals, the individual signals that correspond to such different sources as experiment stimuli, cardiac and respiratory effects, and subject/machine movements. However, most ICA algorithms are sensitive to outliers (McKeown et al. (1998b)). One of our contributions is to propose a technique that makes ICA more robust against outliers.

Outliers usually appear as spikes in fMRI data (Kao and MacFall (2000)). The spikes refer to data points with relatively high signal magnitudes, and they are inevitable in fMRI data due to radio-frequency problems in MR scanners, static discharge caused by synthetic fibers, or even abrupt subject movements. For illustration purposes, Figure 1 plots the recorded time series corresponding to three voxels in the fMRI data analyzed later in Section 5. The big spikes around time points 110, 100 and 120 are examples of many other spikes that are commonly detected in fMRI data. The spatial locations of the voxels are indicated by the line crossings on the images in the left column. What is more interesting is that the highlighted voxel in the first row even lies outside of the

brain tissue; hence, the corresponding spike clearly does not represent any brain function. To the best of our knowledge, there has not been much research explicitly addressing statistical issues related to the spikes. Luo and Nichols (2003) presented an exploratory diagnostic tool to identify outliers in fMRI data.

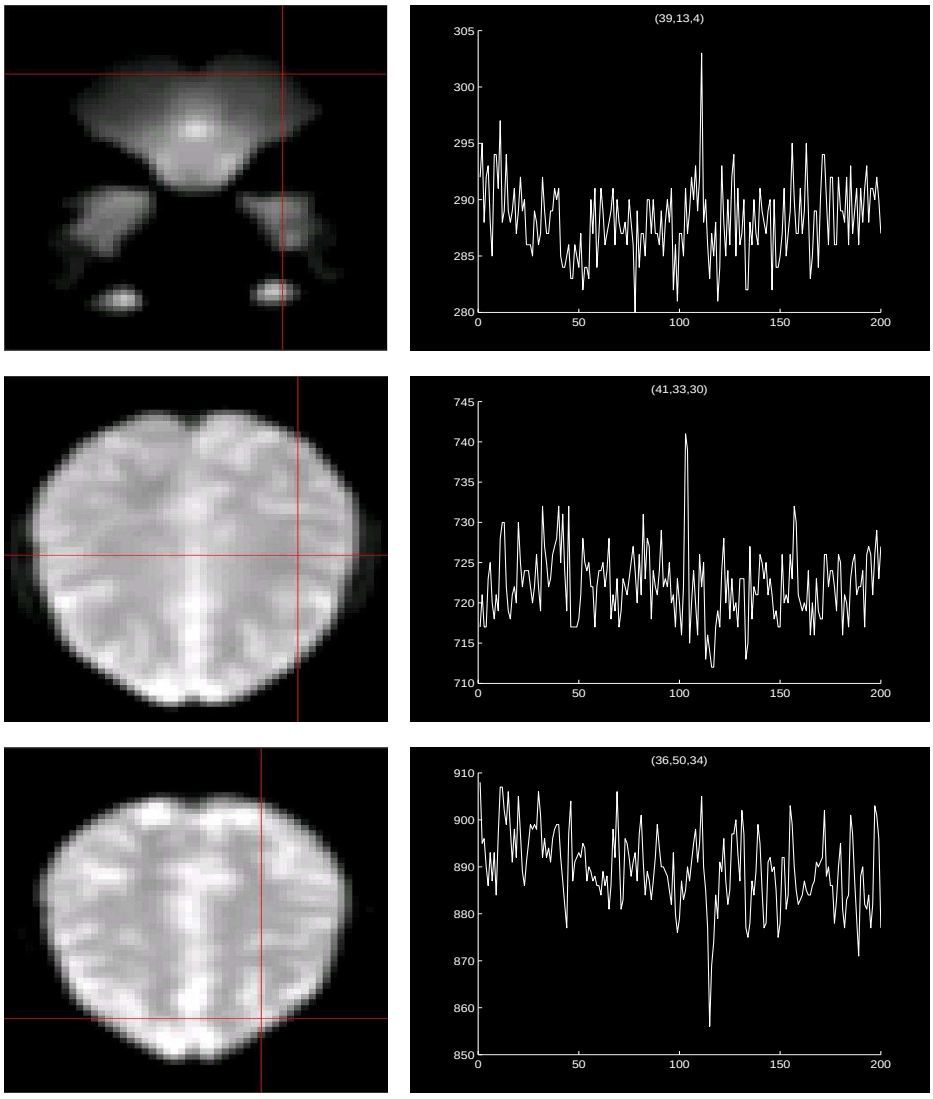

Figure 1. Examples of spikes in our fMRI data set. The left column contains the images for three slices of the brain. The right column plots the time series corresponding to the voxels highlighted by the line crossings on the slices. The big spikes around time points 110, 100 and 120 are examples of many other spikes in the data.

ICA of fMRI usually involves various preprocessing stages (Calhoun et al. (2003)). As mentioned above, fMRI data are usually of high dimension with $M \ll N$, for example, $153,594 \times 200$ in our data. Computationally, it is too expensive to apply ICA algorithms directly. To make ICA more efficient, data reduction is usually performed before ICA through dimension reduction techniques such as singular value decomposition (SVD) (Petersen et al. (2000)). Such techniques aim at reducing the high-dimensional full space to a much smaller feature subspace as retained by SVD, then the ICA decomposition can be focused on the subspace. Section 2.2 gives more detail about this step.

To make ICA more robust, one possibility is to remove the spikes from the subspace on which ICA is performed. However, as a least squares-based method, the conventional SVD is known to be highly susceptible to outliers. In this paper, we propose a supervised SVD (SSVD) that is less sensitive to the spikes. Our proposal is motivated by one interpretation of SVD as a low rank matrix approximation technique (Harville (1997)). The SVD components can then be obtained from solving a sequence of minimization problems. The conventional SVD is unsupervised and purely data-driven. To achieve supervised low rank approximations, and hence improve robustness, we introduce some regularization through basis expansion in the corresponding optimization problems. The basis expansion is constructed using the information of the fMRI experimental designs, particularly the frequencies of the stimulus sequences. Such supervision focuses the SVD on the experimentally interesting directions, which makes the decomposition more efficient and meanwhile less sensitive to spikes, as illustrated in Sections 4 and 5. Our approach is a hybrid between data-driven and model-driven methods, which aligns with the research effort of Hu et al. (2005).

The current research is motivated by one of the few fMRI studies focusing on elucidating the neurocircuitry involved in Parkinson's disease (PD) and its motor dysfunction. PD patients have problems performing simple motor activities, such as finger tapping. One research goal is to identify the brain regions, or Regions of Interest (ROIs), that are associated with finger tapping in these subjects. To detect all the ROIs associated with performing motor activities robustly, this study employs a block design involving alternating right/left-hand finger tapping (Figure 2). Our SSVD-ICA method makes use of this design information.

The rest of the paper is organized as follows. Section 2 provides an overview of ICA, as well as the conventional data reduction step before ICA. In Section 3, we introduce the supervised SVD procedure and its application in ICA for fMRI. Section 4 illustrates the performance of our proposed method using simulation. In Section 5, we compare our method with the conventional ICA using fMRI data collected in the aforementioned study. We close the paper with concluding remarks in Section 6, and technical details in the Appendix.

2. Independent Component Analysis (ICA)

2.1. Overview of ICA for fMRI

ICA has been recently used in fMRI studies to extract independent source signals from the recorded fMRI signals (McKeown et al. (1998b), Petersen et al. (2000) and Calhoun et al. (2003)). The basic idea of ICA can be illustrated using the classic "cocktail party" problem (Hyvärinen, Karhunen and Oja (2001) and Stone (2004)). Suppose many people talk simultaneously at a party, and several microphones are recording in different locations. The recorded signals are then mixtures of different voices. Using only these recorded mixtures as inputs, ICA aims at identifying the individual voices of different people.

Mathematically, let $\mathbf{y}$ and $\mathbf{s}$ denote respectively r-dimensional random vectors with

$$\mathbf{y} = \mathbf{M}\mathbf{s}, \tag{2.1}$$

where $\mathbf{M}$ is an $r \times r$ non-random *mixing* matrix. The signal mixture $\mathbf{y}$ is observable, while the source $\mathbf{s}$ is a vector of latent variables which are assumed to be independent, and hence they are called independent components (IC). The objective of independent component analysis (ICA) is to recover the source by finding a matrix $\mathbf{W}$ so that $\mathbf{s} = \mathbf{W}\mathbf{y}$. The source should not have a Gaussian distribution, otherwise the mixing matrix $\mathbf{M}$ is not identifiable (Hyvärinen, Karhunen and Oja (2001)). There are many algorithms for ICA, for example, that operate by maximization of the non-Gaussianity of $\mathbf{W}\mathbf{y}$, maximum likelihood, or minimization of mutual information (Hyvärinen, Karhunen and Oja (2001) and Stone (2004)).

Consider an fMRI data set represented by a spatio-temporal matrix $\mathbf{X}$ of dimension $M \times N$, with $\operatorname{rank}(\mathbf{X}) = r$. Using (2.1) and the singular value decomposition (see Section 2.2), $\mathbf{X}$ can be decomposed as

$$\mathbf{X} = \mathbf{A}\mathbf{S}, \tag{2.2}$$

where $\mathbf{A}$ and $\mathbf{S}$ are $M \times r$ and $r \times N$ matrices respectively. It might be useful to recall that the right side of (2.2) can be viewed as the outer product of two matrices, defined as the sum of the ith column of $\mathbf{A}$ times the ith row of $\mathbf{S}$, so that the essence of ICA is to decompose $\mathbf{X}$ into the sum of products of spatial maps, columns of $\mathbf{A}$, and the associated source temporal signals, rows of $\mathbf{S}$, when the number of independent components is r.

There exist two distinct ICA options for fMRI signals, spatial ICA (sICA) and temporal ICA (tICA). The sICA aims to find independent image components (columns of $\mathbf{A}$), while tICA looks for independent time courses (rows of $\mathbf{S}$). In either case, a single ICA component can be interpreted as one spatially distributed

set of voxels (i.e., one column of $\mathbf{A}$) that is activated by the corresponding time course in one row of $\mathbf{S}$ (McKeown, Hansen and Sejnowski (2003)).

2.2. The data reduction step before ICA

fMRI data are usually of high dimension, especially in the spatial domain. In addition, ICA algorithms are often computationally intensive. Hence before applying ICA, it is common to first perform dimension reduction using SVD (Petersen et al. (2000)). ICA algorithms are then applied in the reduced subspace.

Suppose $\text{rank}(\mathbf{X}) = r \leq \min(M, N)$. The SVD decomposes $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \tag{2.3}$$

where $\mathbf{U}$ is the $M \times r$ matrix of *orthonormal* left singular vectors, $\mathbf{V}$ is the $N \times r$ matrix of *orthonormal* right singular vectors, and $\mathbf{D}$ is the $r \times r$ diagonal matrix of *positive* singular values. Here $\mathbf{U}$ and $\mathbf{V}$ can be viewed as the basis vectors that span the spatial patterns and temporal sequences, respectively.

As aforementioned, the sICA of $\mathbf{X}$ looks for independent image components. When performing sICA, we can focus on the subspace spanned by $\mathbf{Y} \equiv \mathbf{U}\mathbf{D}$, which is of dimension $M \times r$, much smaller than the original $M \times N$. Apply (2.1) to the columns of $\mathbf{Y}^T$ yielding $\mathbf{Y} = \mathbf{A}\mathbf{M}$, where the columns of $\mathbf{A}$ are r independent spatial maps of size M. The original time series at (2.2) can then be reconstructed as $\S = \mathbf{M}\mathbf{V}^T$. Similarly, tICA can be performed on the subspace retained by $\mathbf{Y}^* \equiv \mathbf{D}\mathbf{V}^T$ to identify independent time series.

In applications, the rank r may still be too large and a common approach is to set the number of ICs, or one simply uses the common default value of 20. In fact, one then carries out ICA by using the largest 20 singular values of $\mathbf{D}$, followed by a search for the temporal component that matches closely with the stimulus of interest. As long as the number of ICs is less than the rank r, the process continues until the desired match is found.

In summary, the idea of the data reduction step before ICA is to reduce the dimension of the matrices to be used as inputs for ICA algorithms. Considering the fact that most fMRI experiments have far less time points than voxels, this preprocessing step greatly improves computational efficiency.

3. Supervised Singular Value Decomposition

Both ICA and SVD are sensitive to the spikes frequently encountered in fMRI data. To overcome this problem, we propose to reduce the effect of the spikes during the data-reduction step. The idea is to replace the conventional SVD with a modified SVD technique that is supervised by the experimental design.

Such supervision makes the procedure more robust, as it reduces the dependence on data. Consequently, this makes the follow-up ICA more robust because ICA focuses on the subspace obtained by the data reduction.

3.1. Low rank approximation via SVD

To motivate our approach, we first note that SVD can be viewed as a matrix low rank approximation method. In the SVD of $\mathbf{X}$ at (2.3), let $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_r]$, $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_r]$, and $\mathbf{D} = \text{diag}\{d_1, \ldots, d_r\}$.

For an integer $l \leq r$, write $\mathbf{X}^{(l)} \equiv \sum_{k=1}^{l} d_k \mathbf{u}_k \mathbf{v}_k^T$. Then, $\mathbf{X}^{(l)}$ is the closest rank-l matrix approximation to $\mathbf{X}$ (Harville (1997)), in the sense that $\mathbf{X}^{(l)}$ minimizes the squared Frobenius norm between $\mathbf{X}$ and any rank-l matrix $\mathbf{X}^*$: $\|\mathbf{X} - \mathbf{X}^*\|_F^2 = \text{tr}\{(\mathbf{X} - \mathbf{X}^*)(\mathbf{X} - \mathbf{X}^*)^T\}$.

Suppose, for example, we seek the best rank-one matrix approximation of $\mathbf{X}$. Note that any $M \times N$ rank-one matrix can be written as $\mathbf{u}\mathbf{v}^T$, where $\mathbf{u}$ is a norm-1 M-vector and $\mathbf{v}$ is a N-vector. The problem can be formulated as

$$\min_{\mathbf{u},\mathbf{v}} \|\mathbf{X} - \mathbf{u}\mathbf{v}^T\|_F^2. \tag{3.1}$$

Then, SVD's low-rank approximation property implies the solution $\mathbf{u} = \mathbf{u}_1$, $\mathbf{v} = d_1\mathbf{v}_1$. The subsequent pairs $(\mathbf{u}_k, d_k, \mathbf{v}_k)$, $k > 1$, provide best rank-one approximations of the corresponding residual matrices. For example, $d_2\mathbf{u}_2\mathbf{v}_2^T$ is the best rank-one approximation of $\mathbf{X} - d_1\mathbf{u}_1\mathbf{v}_1^T$.

3.2. Supervised SVD

In block design fMRI studies, experimental tasks or stimuli are typically applied in alternating blocks. In the areas activated by these stimuli, we observe temporal data that are correlated with the experiment design. Figure 2 shows the time series (solid line) of one voxel that is activated by the experimental stimulus (dashed line) in the fMRI study reported on later in Section 5. We note that the (periodic) time components can be approximately modeled as sinusoidal curves plus some noise in such experiments. This observation motivates us to propose the following supervised SVD (SSVD) procedure.

Consider the time component of interest, $\mathbf{v} = (v(t_1), \ldots, v(t_N))^T$, where t_i is the ith scanning time. To fix ideas, suppose we model the component as

$$v(t_i) = a\sin(2\pi\omega t_i + \phi) = a\cos\phi\sin(2\pi\omega t_i) + a\sin\phi\cos(2\pi\omega t_i), \tag{3.2}$$

where a is the amplitude, ϕ is the phase, and ω is the frequency. Let $\boldsymbol{\psi} = (a\cos\phi, a\sin\phi)^T$ and $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2)$, where $\mathbf{b}_1 = (\sin(2\pi\omega t_1), \ldots, \sin(2\pi\omega t_N))^T$, and $\mathbf{b}_2 = (\cos(2\pi\omega t_1), \ldots, \cos(2\pi\omega t_N))^T$. Then, in matrix form, one can see that

$$\mathbf{v} = \mathbf{B}\boldsymbol{\psi}, \tag{3.3}$$

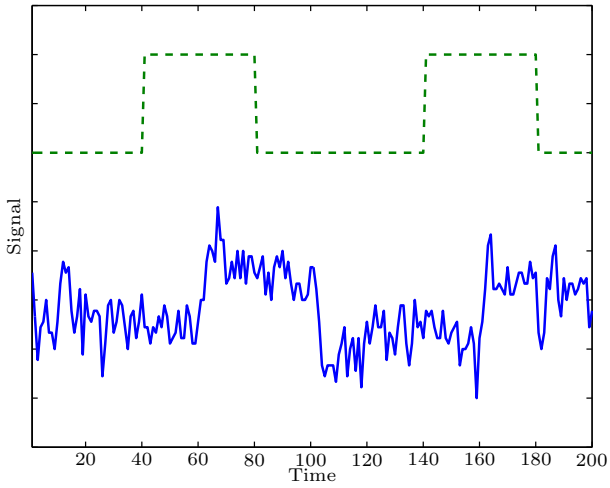

Figure 2. The recorded time course (solid line) of a voxel that is activated by the experimental stimulus sequence (dashed line). The lower and higher levels of the dashed line stand for rest and activation periods of the experiment, respectively.

suggesting that $\mathbf{v}$ lies in the linear space spanned by the bases $\mathbf{b}_1$ and $\mathbf{b}_2$. We remark that the choice of the sinusoidal function at frequency ω in (3.2) is an appoximation for the true periodic time component. See Section 6 for more discussion about the rationale for this choice.

To make use of the sinusoidal nature of $\mathbf{v}$, we propose to impose the constraint (3.3) on $\mathbf{v}$ in the optimization problem (3.1), and re-formulate it as

$$\min_{\mathbf{u},\boldsymbol{\psi}} \|\mathbf{X} - \mathbf{u}\mathbf{v}^T\|_F^2 \qquad \text{subject to } \mathbf{v} = \mathbf{B}\boldsymbol{\psi}.$$

We name this formulation *Supervised SVD*, because it *supervises* SVD by restricting it to find the best low rank approximation within a certain subspace.

For identifiability purpose, we can standardize $\mathbf{u}$ and $\mathbf{v}$ to unit length and introduce a slope parameter d. The problem is then rewritten as

$$\min_{d,\mathbf{u},\boldsymbol{\psi}} \|\mathbf{X} - d\mathbf{u}\mathbf{v}^T\|_F^2 \qquad \text{subject to } \mathbf{v} = \mathbf{B}\boldsymbol{\psi}, \mathbf{u}^T\mathbf{u} = 1, \mathbf{v}^T\mathbf{v} = 1. \tag{3.4}$$

For simplicity here, the SSVD is illustrated using only the fundamental Fourier bases in $\mathbf{B}$; one can easily extend $\mathbf{B}$ to accommodate higher order Fourier bases, which makes it more flexible for modeling periodic signals.

3.3. Solution and practical implementation

The solution of the Supervised SVD (3.4) can be obtained by solving two generalized eigen-problems as stated below.

Theorem 1. *The minimizer of* (3.4), $\{d, \mathbf{u}, \boldsymbol{\psi}\}$, *satisfies*

$$\begin{cases} \max_{\mathbf{u}} \ \mathbf{u}^T\mathbf{X}\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}^T\mathbf{u} & \text{subject to } \mathbf{u}^T\mathbf{u} = 1, \\ \max_{\boldsymbol{\psi}} \ \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{X}\mathbf{B}\boldsymbol{\psi} & \text{subject to } \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{B}\boldsymbol{\psi} = 1, \\ d = \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{u}. \end{cases}$$

The proof of the theorem is relegated to the Appendix. The generalized eigen-problems can be solved using standard methods as shown there. Below we summarize the computational algorithm to obtain the first supervised time component as well as its corresponding spatial component. See the Appendix for the technical justification. The same algorithm can be applied repeatedly on the residual matrices until the desired number of components is obtained.

Algorithm 1. Supervised SVD (SSVD)

(1) Choose the frequency of interest ω, and form the basis matrix $\mathbf{B}$.

(2) Apply Cholesky decomposition on $\mathbf{B}^T\mathbf{B}$ to get $\mathbf{B}^T\mathbf{B} = \mathbf{R}_{\mathbf{B}}^T\mathbf{R}_{\mathbf{B}}$.

(3) Apply SVD on $\mathbf{X}\mathbf{B}\mathbf{R}_{\mathbf{B}}^{-1}$ to derive its first left singular vector $\mathbf{u}$ and the first right singular vector $\tilde{\boldsymbol{\psi}}$.

(4) Set $\boldsymbol{\psi} \equiv \mathbf{R}_{\mathbf{B}}^{-1}\tilde{\boldsymbol{\psi}}$ and $d \equiv \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{u}$, which leads to $\mathbf{v} = \mathbf{B}\boldsymbol{\psi}$. Hence obtain the first SSVD triplet $\{d, \mathbf{u}, \mathbf{v}\}$.

The choice of the sinusoidal frequency ω in the above algorithm can be easily made in most of our fMRI experiments, as the frequency ω of interest is usually known a priori. For instance, components corresponding to the experimental stimuli and cardiopulmonary systems. To the extent that the experimenter is interested in some unknown signals, we propose to estimate ω through spectral analysis on the $\mathbf{v}$ components extracted from the conventional SVD. Although the SVD components are affected by the spikes, their dominating frequencies are less likely to be affected, unless the number of spikes is overwhelmingly large and they occur in a periodic manner. The robustness of the frequencies is supported by the numerical studies in Sections 4 and 5.

3.4. Application to ICA

Once we obtain the desired number of SSVD components, we propose to apply an ICA algorithm to them. As the SSVD components are insensitive to

the spikes, the results from the follow-up ICA procedure are robust as well, as illustrated in Section 4 via simulation. In addition, because our procedure (SSVD-ICA) uses the design information of the fMRI experiment, it is more powerful in detecting activated brain regions of interest, as shown in Section 5.

As for the number of SSVD components to extract, we propose the following. If we know all the frequencies of interest, we can extract the corresponding components; otherwise, we suggest extracting the first 20 (for example) SVD components, and perform spectral analysis on them to identify possibly interesting frequencies, before applying SSVD. This proposal is consistent with the common practice in ICA for fMRI data as described at the end of Section 2.

4. Simulation Studies

Below, we compare the SSVD-ICA procedure with the conventional ICA that uses SVD via two simulation studies. The first simulation is presented in detail in Sections 4.1 to 4.3, while the second one is briefly described in Section 4.4, to save space. The results confirm that SSVD-ICA is more robust than ICA. In addition, the second simulation shows that our approximation of modeling temporal components using sinusoidal curves works well.

4.1. Data description

According to the ICA model (2.2), to simulate an $M \times N$ fMRI data matrix $\mathbf{X}$, we should first simulate the $M \times r$ spatial component matrix $\mathbf{A}$ and the $r \times N$ time series matrix $\mathbf{S}$ separately. The data matrix $\mathbf{X}$ is then obtained as $\mathbf{X} = \mathbf{AS}$.

In this study, we set $r = 5$, $M = 30 \times 30 \times 10$ and $N = 240$. Basically, there are 5 underlying independent components; each column of $\mathbf{A}$ is a spatial map that consists of 10 slices and each slice contains 30×30 voxels; while each row of § is a time series of length 240 that corresponds to the relevant spatial component in $\mathbf{A}$. We simulated the data based on a simple rest-activation block design (the dotted line in the time plot within each panel of Figure 3). Each rest or activation period lasted 18 seconds.

The first four time components were simulated using simple sinusoidal functions plus Uniform noises generated between $[-0.05,\ 0.05]$. The noise distribution had to be non-Gaussian for ICA to be well defined (Section 2). The components are plotted in Figure 3 along with their corresponding spectral plots highlighting the frequencies. The amplitudes of the first four components were 0.5, 0.45, 0.35 and 0.45, respectively. The corresponding frequencies were 0.06Hz, 1Hz, 0.3Hz and 0.7Hz, carefully chosen so that they had physical interpretations. For example, one can imagine that the first one corresponds to the stimulus of some artificial experiment; the second and third frequencies are for the heart beat and

breath; the fourth one is for some artifact effect. The last time component was pure noise sampled uniformly between -0.05 and 0.05.

The voxels in $\mathbf{A}$ were assigned a 0 or 1. In each component, the voxels with value 1 corresponded to the regions activated by the corresponding time stimulus; they are plotted as dark red areas in Figure 3.

To simulate spikes in the data, we randomly selected 10% of the entries in the simulated $\mathbf{X}$, and replaced them with random noise generated from Uniform$[-8, -2]$ and Uniform$[2, 8]$. The noise distributions were chosen so that the simulated spikes were indeed outliers, judged by the 1.5-Inter-Quartile-Range rule of thumb.

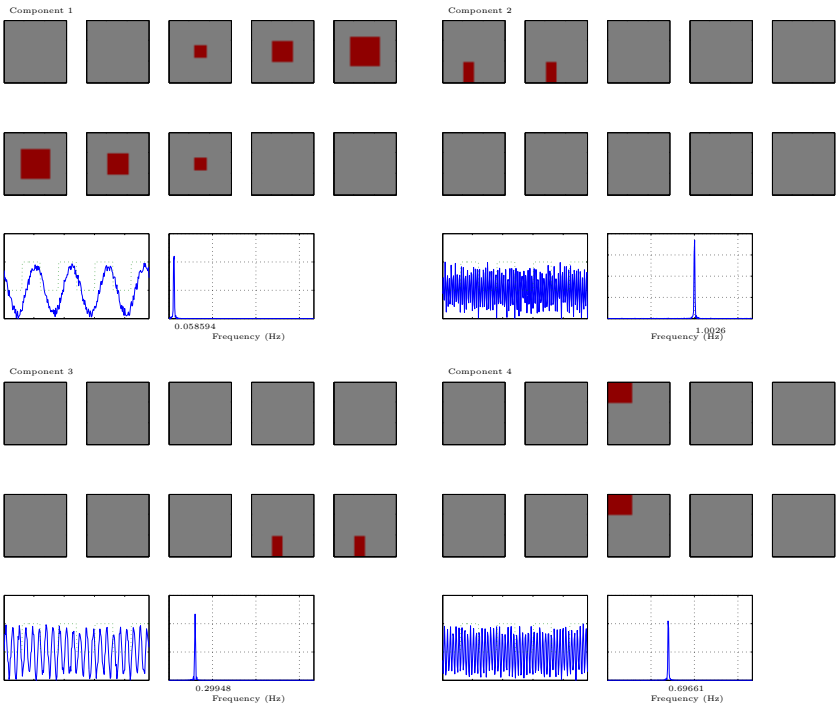

Figure 3. The first four components used in the simulation. In each panel, the first 10 images are the spatial component maps (one column of $\mathbf{A}$), and the dark red areas stand for activated voxels. The solid line in the subsequent plot is the corresponding time series (one row of §). The dotted line stands for the rest-activation block design, 0 for "rest" and 1 for "active". The spectrum plot for each time series is given at the end of each panel highlighting the frequencies. In this simulation study, Component 1 can be viewed as the one related to the experiment stimulus; Components 2 and 3 stand for heart beat and breath, respectively; Component 4 could be an artifact effect; Component 5 is not shown here since it is pure noise.

4.2. Analysis

Following the standard practice in ICA, we first normalized the contaminated data matrix $\mathbf{X}$ by column centering and row standardization (Hastie and Tibshirani (2002)). Both ICA and SSVD-ICA were then applied to the normalized matrix. For computing, we employed the fastICA algorithm (Hyvärinen, Karhunen and Oja (2001)), because of its fast computation and popularity.

To effectively display the activated voxels in the spatial maps, the values in each map were standardized to z-scores (McKeown et al. (1998a)) by subtracting the component mean and then dividing the component standard deviation. The voxels with $|z| \geq 1$ were then identified as those activated, and assigned value 1 when plotting, while the other voxels (with $|z| < 1$) were assigned 0.

4.3. Results

The results from SSVD-ICA are shown in the first column of Figure 4. All four components can be recovered reasonably well, although with some noise. Again, dark red areas indicate the activated voxels. The time course plots in each panel show the corresponding time components and their spectral plots.

The second column of Figure 4 shows the results from the conventional ICA that uses SVD for dimension reduction. Only the first component could be recovered. Even so, the result is more blurred than the corresponding one from SSVD-ICA. Thus the conventional ICA had trouble identifying the other components.

4.4. A second simulation

Our SSVD procedure models the periodic time components as sinusoidal functions, which is sometimes an approximation. The experimental condition in the above simulation was consistent with the sinusoidal assumption. To further investigate the performance of our approach, we designed a second simulation, where the true time components were boxcar-type stepwise functions, instead of sinusoidal. To save space, the details of the study are reported in our online supplementary report (Bai et al. (2007)).

Again, the SSVD-ICA recovered all spatial components reasonably well, even though it approximated the temporal components using sine curves. We think the reason is that the frequencies of the temporal components are the key factor, which is utilized by SSVD-ICA. On the other hand, the conventional ICA only detected half of the components, and the detected regions were more blurred.

5. An fMRI Data Analysis

5.1 Experiment paradigm and data description

To study brain regions that are related to different finger tapping movements, an fMRI data set was obtained from one human subject performing three different

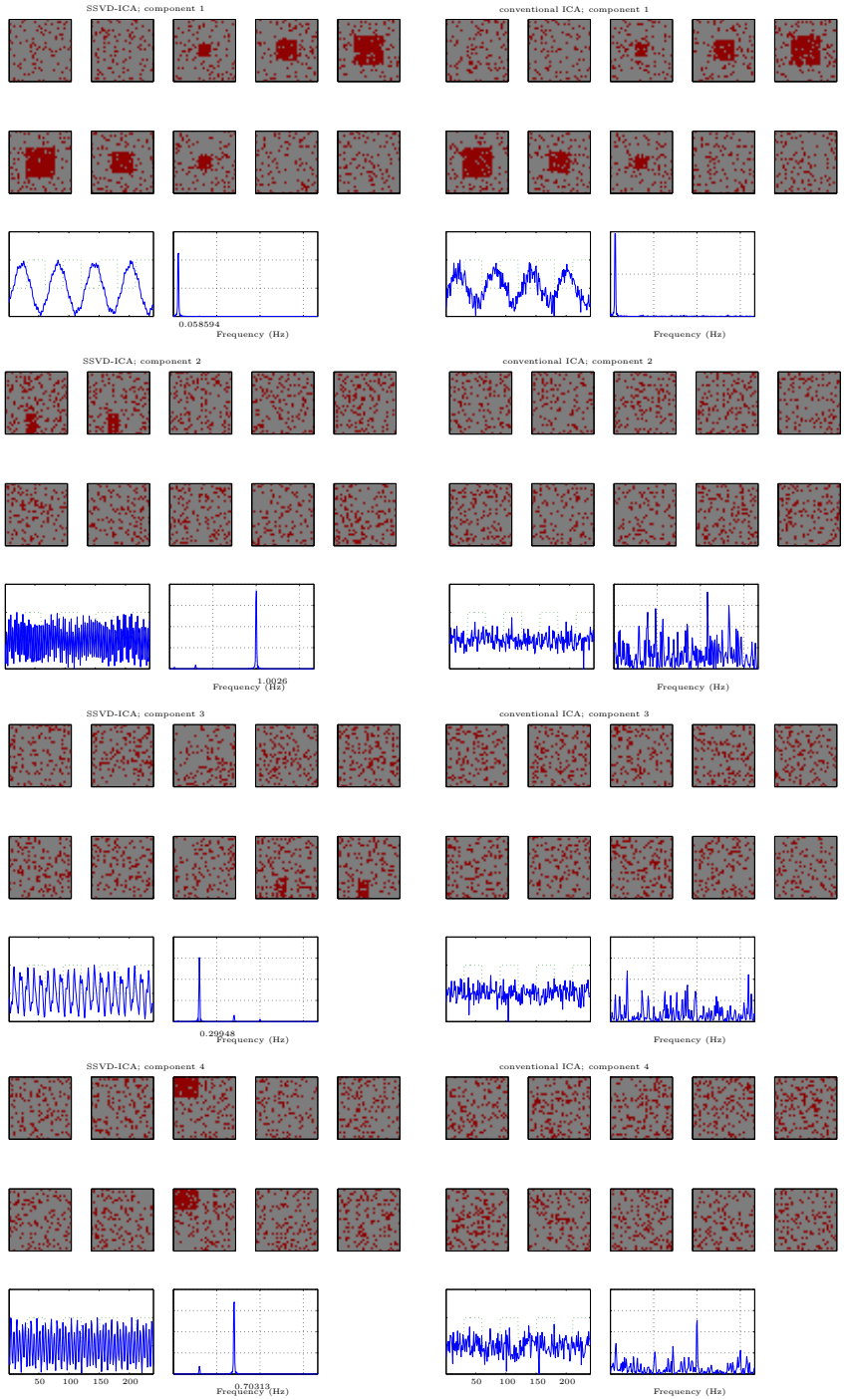

Figure 4. Comparison of the results from the proposed SSVD-ICA (the left column) and the conventional ICA that uses SVD (the right column).

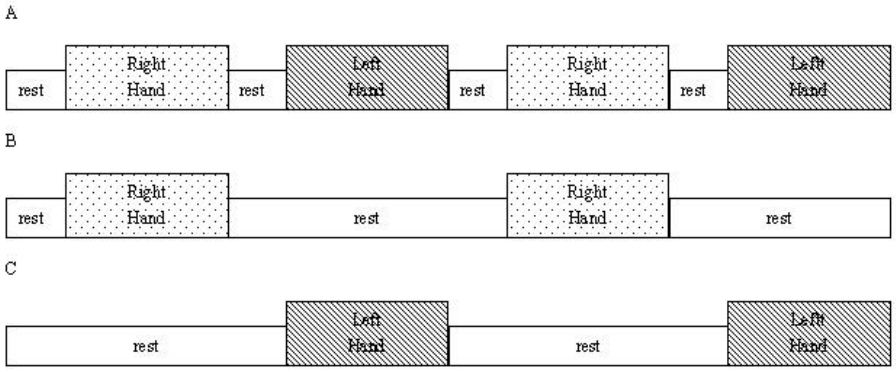

Figure 5. The experimental design used in acquiring the fMRI data. Panel (A) shows the complete design sequence. Panels (B) and (C) show the paradigms for right-hand and left-hand movements separately. Each rest block took 30 seconds (10 scans when TR = 3 seconds); each activation block took 120 seconds (40 scans).

tasks alternately: rest, right-hand movement, and left-hand movement. Each rest period lasted 30 seconds and each activation period lasted 120 seconds. The experimental paradigm is shown in Figure 5. The study uses block design, because the research interest is to robustly identify all the regions of interest (ROIs) involving finger tapping.

Two hundred MR scans were acquired on a modified 3T Siemens MAGNETOM Vision system. Each acquisition consisted of 49 contiguous slices. Each slice contained 64×64 voxels. Hence there were $64 \times 64 \times 49$ voxels from each scan. The size of each voxel is 3mm $\times$ 3mm $\times$ 3mm. Each acquisition took 2.9388 seconds, with the scan to scan repetition time (TR) set to be 3 seconds.

5.2. Analysis

The data set was preprocessed using SPM5. The preprocessing included realignment, coregistration, segmentation, spatial normalization, and spatial smoothing. We then used the MATLAB function **showsrs** developed by the Duke-UNC Brain Imaging and Analysis Center (BIAC) to visually check the processed data. Many spikes can be easily seen, three of which are plotted in Figure 1.

We then applied both the conventional ICA and our proposed SSVD-ICA to the processed data after column centering and row standardization. When applying ICA, we employed an existing analyzer of fMRI data, *Group ICA of fMRI Toolbox* (GIFT) (Calhoun et al. (2001)). With SSVD-ICA, we performed separate analyses using both the true frequencies and the estimated ones. For all three methods, spatial ICA was carried out using the fastICA algorithm.

5.3. Results

The results from the three methods are displayed in Figure 6. The goal of this study is to identify the activated brain regions corresponding to right-hand and left-hand movements. Four typical slices in cortical, subcortical and cerebellum regions are shown.

Within each row, the four image slices on the left represent the spatial maps, where the activated brain regions are highlighted in red and blue colors. Brighter color indicates higher activation intensity. All methods demonstrate the classic brain activation patterns during hand movement (Buhmann et al. (2003), Elsinger, Harrington and Rao (2006) and Taniwaki (2006)):

- *contralateral more than ipsilateral in the Primary Motor Cortex* (PMC) (slice 1 and slice 2), *Supplementary Motor Area* (SMA) (the bright area in the middle of slice 2), and *basal ganglia* (slice 3);
- *ipsilateral more than contralateral in the cerebellum* (slice 4).

The rightmost plot is the corresponding time component (solid line) with the stimulus sequence (dotted line) overlayed.

When applying GIFT, 30 independent components are needed in order to obtain the interesting ones related to both hand movements. The two components are shown in the two rows of Panel (A) respectively. The components identified by SSVD-ICA using estimated frequencies are presented in Panel (B). Our method clearly results in much less noise and higher intensity in all activated regions. In addition, the time components recovered by our method are better correlated with the experimental stimulus than those obtained by GIFT.

We also applied SSVD-ICA by specifying the true experimental frequency (0.0033Hz). Note that both right-hand and left-hand movements share the same frequency in this study. As a result, the brain regions related to both movements are recovered in a single component as shown in Panel (C), where red areas indicate voxels activated by the right-hand movement and blue areas indicate voxels activated by the left-hand movement. Again, the activated areas are less noisy and more highlighted when compared with GIFT. Another advantage of specifying the frequency is that we only need to extract one component that is relevant to the experimental paradigm. By comparison, GIFT acquired 30 components before identifying the interesting components via visual inspection or correlation analysis.

6. Conclusion and Discussion

In this paper, we introduced an experiment-guided dimension reduction technique into the ICA procedure for fMRI studies, one that improves ICA's robustness against spikes that are common in fMRI data. We modified the conventional

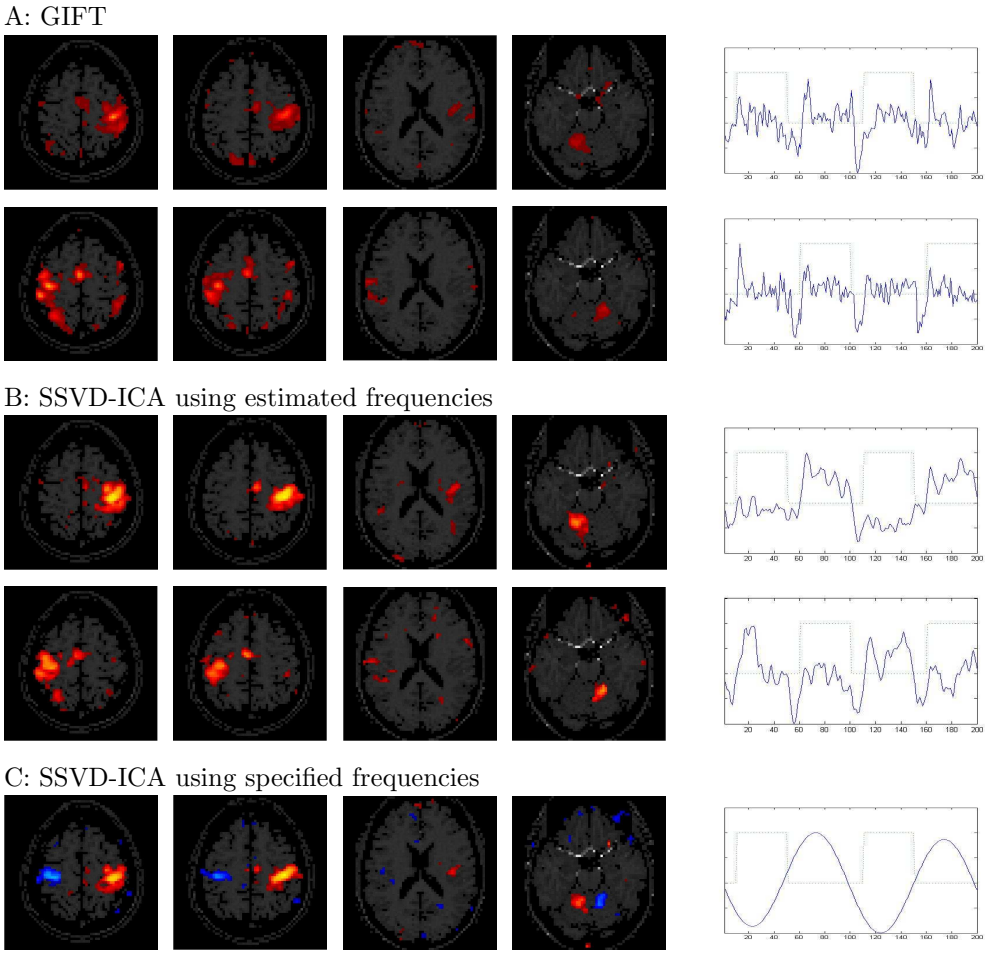

Figure 6. Activated brain regions and their corresponding temporal components detected by three methods. Panel (A): GIFT, right-hand (first row) and left-hand (second row); Panel (B): SSVD-ICA using estimated frequencies, right-hand (first row) and left-hand (second row); Panel (C): SSVD-ICA using specified frequencies, both right-hand and left-hand. The orientation of the brian slices is such that the right side of the brain is on the left side, and the left side of the brain is on the right side. Within each row, the first slice shows the *primary motor cortex* (PMC), the second slice contains both PMC and *supplementary motor area* (SMA), the third slice shows *basal ganglia*, and the fourth slice shows *cerebellum*. Red and blue areas illustrate the activated voxels. Brighter color indicates higher intensity. The last plot shows the corresponding time course with the right-hand-movement stimulus sequence overlayed.

SVD via basis expansion guided by the experimental design. The main benefits are two-fold: first, the extracted components are less sensitive to the spikes; second, the activated brain areas are identified with less noise and higher intensity. We illustrated such benefits via simulation and a comparison with GIFT using fMRI data. In addition, our method addresses a relevant component selection problem that is of immense importance in fMRI. The common practice of ICA is to acquire around 30 independent components, before finding the most relevant ones by visual inspection or correlation analysis. Our method allows researchers to decide on the number of components by specifying the frequencies of interest. This hybrid of data and model hypothesis makes the analysis more efficient, considering the usually huge size of fMRI data.

Our proposal achieves robustness by constraining SVD in the pre-processing step, which is direct, intuitive and effective. An alternative is to constrain ICA directly, which is technically more challenging. A common approach to ICA uses the marginal density information through high-order kurtosis as in fast ICA. More recently, nonparametric density estimation procedures have been applied to ICA by Chen and Bickel (2006) and Kawaguchi and Truong (2007). It is however a much harder problem to constrain the ICA in fMRI, because it is similar to identifying outliers for a large number of voxel time series, and so far there have not been good kernel methods for density estimation involving outliers. Furthermore, it is known in spatial ICA that the distributions are heavy tailed (Stone et al. (2002)), which means that treating outliers too early can lead to loss of information. It is important to note that our present results can be improved by combining our method with KDICA of Chen and Bickel (2006) and the polynomial spline ICA of Kawaguchi and Truong (2007).

Our approach approximates the periodic temporal components using sinusoidal functions at the interesting frequencies. As shown by the simulation studies and the application, such approximation works well for identifying spatial regions of interest, as it utilizes the key factor of frequency. Nevertheless, it is of interest to let the data determine the shape of the temporal components nonparametrically after specifying the frequencies. Periodic splines seems to offer a nice tool, the applicability of which is currently being investigated.

Neuroimage researchers are also interested in analyzing fMRI across groups of subjects, and Group ICA has been developed for that purpose (Calhoun et al. (2001)). Encouraged by the performance of our method in analyzing one subject, we are interested in extending it to Group ICA.

Acknowledgement

The authors want to thank an associate editor and one reviewer for their helpful comments that have greatly improved the scope and presentation of the

paper. Thanks also go to Josh Bizzell, Mechelle Lewis, and Andrew Smith for helpful comments and assistance with the analysis. The research was partially supported by NSF grants DMS-0606577 and CMMI-0800575 (Haipeng Shen) and DMS-0707090 (Young Truong), and NIH grants AG-021491 and GCRC 000046 (Xuemei Huang).

Appendix. Proof of Theorem 1 and Justification for Algorithm 1

Below we first provide the proof of Theorem 1, using matrix algebra.

Note that (3.4) is equivalent to minimizing

$$\mathrm{tr}(\mathbf{X}\mathbf{X}^T) - 2d\boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{u} + d^2(\boldsymbol{\psi}^T\mathbf{B}^T\mathbf{B}\boldsymbol{\psi})(\mathbf{u}^T\mathbf{u}), \tag{A.1}$$

subject to $\mathbf{u}^T\mathbf{u} = 1$ and $\mathbf{v}^T\mathbf{v} = 1$. For fixed d and $\mathbf{u}$, the minimizer is

$$\boldsymbol{\psi} = (d\mathbf{u}^T\mathbf{u}\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}^T\mathbf{u} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}^T\frac{\mathbf{u}}{d},$$

which we can plug back into (A.1) to obtain $\mathrm{tr}(\mathbf{X}\mathbf{X}^T) - \mathbf{u}^T\mathbf{X}\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}^T\mathbf{u}$. Hence the minimization problem (3.4) is reduced to

$$\max_{\mathbf{u}} \ \mathbf{u}^T\mathbf{X}\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X}^T\mathbf{u} \qquad \text{subject to } \mathbf{u}^T\mathbf{u} = 1, \tag{A.2}$$

which is a generalized eigen-problem. Similarly, the $\boldsymbol{\psi}$ that solves (3.4) is the solution to the generalized eigen-problem

$$\max_{\boldsymbol{\psi}} \ \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{X}\mathbf{B}\boldsymbol{\psi} \qquad \text{subject to } \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{B}\boldsymbol{\psi} = 1. \tag{A.3}$$

According to (A.1), the d that solves (3.4) is given by $d = \boldsymbol{\psi}^T\mathbf{B}^T\mathbf{X}^T\mathbf{u}$.

Given the above proof of Theorem 1, the numerical implementation (Algorithm 1) can be justified as follows.

To solve (A.2), consider the Cholesky decomposition $\mathbf{B}^T\mathbf{B} = \mathbf{R}_\mathbf{B}^T\mathbf{R}_\mathbf{B}$ where $\mathbf{R}_\mathbf{B}$ is a 2×2 upper triangular matrix. Then (A.2) is equivalent to

$$\max_{\mathbf{u}} \ \mathbf{u}^T\mathbf{X}\mathbf{B}\mathbf{R}_\mathbf{B}^{-1}(\mathbf{R}_\mathbf{B}^{-1})^T\mathbf{B}^T\mathbf{X}^T\mathbf{u} \qquad \text{subject to } \mathbf{u}^T\mathbf{u} = 1.$$

The solution to this problem, denoted as $\mathbf{u}^*$, is actually the first left eigenvector of the matrix $\mathbf{X}\mathbf{B}\mathbf{R}_\mathbf{B}^{-1}$ (Harville (1997)).

One can solve (A.3) similarly. Let $\tilde{\boldsymbol{\psi}} = \mathbf{R}_\mathbf{B}\boldsymbol{\psi}$, then (A.3) is equivalent to

$$\max_{\tilde{\boldsymbol{\psi}}} \ \tilde{\boldsymbol{\psi}}^T(\mathbf{R}_\mathbf{B}^{-1})^T\mathbf{B}^T\mathbf{X}^T\mathbf{X}\mathbf{B}\mathbf{R}_\mathbf{B}^{-1}\tilde{\boldsymbol{\psi}} \qquad \text{subject to } \tilde{\boldsymbol{\psi}}^T\tilde{\boldsymbol{\psi}} = 1.$$

The maximizer $\tilde{\boldsymbol{\psi}}^*$ is the first right eigenvector of $\mathbf{X}\mathbf{B}\mathbf{R}_\mathbf{B}^{-1}$. Consequently, the maximizer of the original problem (A.3) is $\boldsymbol{\psi}^* = \mathbf{R}_\mathbf{B}^{-1}\tilde{\boldsymbol{\psi}}^*$.

The above derivation suggests that one only needs to perform a single SVD of $\mathbf{X}\mathbf{B}\mathbf{R}_\mathbf{B}^{-1}$ to obtain both $\boldsymbol{\psi}^*$ and $\mathbf{u}^*$. Then the scale parameter d can be estimated as $d^* = \boldsymbol{\psi}^{*T}\mathbf{B}^T\mathbf{X}^T\mathbf{u}^*$.

References

Bai, P., Shen, H., Huang, X. and Truong, Y. (2007). A supervised singular value decomposition for independent component analysis of fMRI. Supplementary Report. [http://www.unc.edu/~haipeng/research/fmri/ssvd-ica.pdf.]

Buhmann, C., Glauche, V., Sturenburg, H. J., Oechsner, M., Weiller, C. and Büchel, C. (2003). Parmacologically modulated fMRI-cortical responsiveness to levodopa in drug-naive hemiparkinsonian patients. *Brain* **126**, 451-461.

Calhoun, V. D., Adali, T., Hansen, L. K., Larsen, J. and Pekar, J. J. (2003). ICA of functional MRI data: an overview. *In 4th International Symposium on Independent Component Analysis and Blind Signal Separation* (ICA 2003), 281-288.

Calhoun, V. D., Adali, T., Pearlson, G. D. and Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping* **14**, 140-151.

Chen, A. and Bickel, P. J. (2006). Efficient independent component analysis. *Ann. Statist.* **34**, 2825-2855.

Elsinger, C. L., Harrington, D. L. and Rao, S. M. (2006). From preparation to online control: reappraisal of neural circuitry mediating internally generated and externally guided actions. *NeuroImage* **31**, 1177-1187.

Friston, K. J., Holmes, A. P. and Worsley, K. J. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2**, 189-210.

Friston, K. J., Jezzard, P. and Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping* **1**, 153-171.

Genovese, C. R. (2000). A Bayesian time-course model for functional magnetic resonance imaging data (with discussion). *J. Amer. Statist. Assoc.* **95**, 691-703.

Gössl, C., Auer, D. P. and Fahrmeir, L. (2001). Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics* **57**, 554-562.

Gössl, C., Fahrmeir, L. and Auer, D. P. (2001). Bayesian modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage* **14**, 140-148.

Goutte, C., Toft, P., Rostrup, E., Nielsen, F. A. and Hansen, L. K. (1999). On clustering fMRI time series. *NeuroImage* **9**, 298-310.

Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer.

Hastie, T. and Tibshirani, R. (2002). Independent component analysis through product density estimation. Technical report, Department of Statistics at Stanford University.

Hu, D., Yan, L., Liu, Y., Zhou, Z., Friston, K. J., Tan, C. and Wu, D. (2005). Unified SPM-ICA for fMRI analysis. *NeuroImage* **25**, 746-755.

Huettel, S. A., Song, A. W. and McCarthey, G. (2004). *Functional Magnetic Resonance Imaging*. Sinauer Associates, Inc.

Hyvärinen, A., Karhunen, J. and Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sons, New York.

Jezzard, P., Matthews, P. M. and Smith, S. M., editors (2001). *Functional MRI: An Introduction to Methods*. Oxford University Press.

Kao, Y. and MacFall, J. R. (2000). Correction of MRI κ-space data corrupted by spike noise. *IEEE Trans. Medical Imaging* **19**, 671-680.

Kawaguchi, A. and Truong, Y. (2007). Spline independent component analysis. Manuscript. [http://www.unc.edu/~haipeng/research/fmri/sica.pdf.]

Kherif, F., Poline, J-B., Flandin, G., Benali, H., Simon, O., Dehaene, S. and Worsley, K. J. (2002). Multivariate model specification for fMRI data. *NeuroImage* **16**, 1068-1083.

Lange, N., Strother, S. C., Anderson, J. R., Nielsen, F. A., Holmes, A. P., Kolenda, T., Savory R. and Hansen, L. K. (1999). Plurality and resemblance in fMRI data analysis. *NeuroImage* **10**, 282-303.

Luo, W. and Nichols, T. E. (2003). Diagnosis and exploration of massively univariate neuroimaging models. *NeuroImage* **19**, 1014-1032.

McKeown, M. J., Hansen, L. K. and Sejnowski, T. J. (2003). Independent component analysis of functional MRI: what is signal and what is noise? *Current Opinion in Neurobiology* **13**, 620-629.

McKeown, M. J., Jung, T. P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T.-W. and Sejnowski, T. J. (1998a). Spatially independent activity patterns in functional MRI data during the stroop color-naming task. *PNAS* **95**, 803-810.

McKeown, M. J., Makeig, S., Brown, G. G., Jung, T., Kindermann, S. S., Bell, A. J. and Sejnowski, T. J. (1998b). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* **6**, 160-188.

Petersen, K. S., Hansen, L. K., Kolenda, T., Rostrup, E. and Strother, S. C. (2000). On the independent components of functional neuroimages. *Proceedings of the Third International Conference on Independent Component Analysis and Blind Source Separation* (ICA 2000), 615-620.

Stone, J. V., Porrill, J., Porter, N. R. and Wilkinson, I. D. (2002). Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions, *NeuroImage* **15**, 407-421.

Stone, J. V. (2004). *Independent Component Analysis: A Tutorial Introduction*. The MIT Press.

Taniwaki, T., Okayama, A., Yoshiura, T.,Togao, O., Nakamura, Y., Yamasaki, T., Ogata, K., Shigeto, H., Ohyagi, Y., Kira, J. and Tobimatsu, S. (2006). Functional network of the basal ganglia and cerebellum motor loops in vivo: different activation patterns between self-initiated and externally triggered movements. *NeuroImage* **31**, 745-753.

PayPal Inc., San Jose, CA 95131, U.S.A.

E-mail: pingbai@gmail.com

Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

E-mail: haipeng@email.unc.edu

Departments of Neurology, Radiology, Neurosurgery, Kinesiology and Bioengineering, Penn State-Milton's Hershey Medical Center, Hershey, PA 17033, U.S.A.

E-mail: xhuang3@hmc.psu.edu

Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

E-mail: truong@bios.unc.edu

(Received April 2007; accepted December 2007)