Journal of Statistical Software

October 2011, Volume 44, Issue 9. http://www.jstatsoft.org/

# Temporal and Spatial Independent Component Analysis for fMRI Data Sets Embedded in the AnalyzeFMRI R Package

**Cécile Bordier**
INSERM U836

**Michel Dojat**
INSERM U836

**Pierre Lafaye de Micheaux**
Université de Montréal

**Abstract**

For statistical analysis of functional magnetic resonance imaging (fMRI) data sets, we propose a data-driven approach based on independent component analysis (ICA) implemented in a new version of the **AnalyzeFMRI** R package. For fMRI data sets, spatial dimension being much greater than temporal dimension, spatial ICA is the computationally tractable approach generally proposed. However, for some neuroscientific applications, temporal independence of source signals can be assumed and temporal ICA becomes then an attractive exploratory technique. In this work, we use a classical linear algebra result ensuring the tractability of temporal ICA. We report several experiments on synthetic data and real MRI data sets that demonstrate the potential interest of our R package.

*Keywords*: multivariate analysis, temporal ICA, spatial ICA, magnetic resonance imaging, neuroimaging.

## 1. Introduction

Magnetic resonance imaging (MRI) is now a prominent non-invasive neuroimaging technique largely used in clinical routine and advanced brain research. Its success is largely due to a combination of at least three factors: (1) sensitivity of MR signal to various physiological parameters that characterize normal or pathological living tissues (such as diffusion properties of water molecules, relaxation time of proton magnetization or blood oxygenation) leading to a vast set of MRI modalities (respectively restricted in our example to diffusion MRI, weighted structural and functional MRI); (2) constant hardware improvements (e.g., mastering high field homogeneous magnets and high linear magnetic field gradients respectively allow an increase of spatial resolution or a reduction of acquisition time); and (3) sustained efforts in various laboratories to develop methods: for image processing (to de-noise, segment, realign, fusion or visualize MR brain images), for computational anatomy leading to the exploration of brain structure modifications during learning, brain development or pathology evolution and for time course analysis of functional MRI data. Statisticians play a key role in this last factor since data produced are complex: noisy, highly variable between subjects, massive and, for functional data, highly correlated both spatially and temporally, see Lange (2003).

Functional MRI (fMRI) allows to detect the variations of cerebral blood oxygen level induced by the brain activity of a subject, lying inside a MRI scanner, in response to various sensory-motor or cognitive tasks, see Chen and Ogawa (1999). The fMRI signal is based on changes in magnetic susceptibility of the blood during brain activation. It is a non-invasive and indirect detection of brain activity: the signal detected is filtered by the hemodynamic response function (HRF) and the neuro-vascular coupling is only partially explained, see Logothetis and Pfeuffer (2004). The main goal of fMRI experiments is to explore, in a reproducible way, the cortical networks implicated in pre-defined stimulation tasks in a cohort of normal or pathological subjects. The low signal to noise ratio obtained in functional images requires to repeat the sequence of stimuli several times (Henson 2004) and to enroll a sufficient number of subjects (Thirion *et al.* 2007). In general, the data resulting from an fMRI experiment consist in a set indexed with time (typically many hundred) of 3D dimensional functional images with a $3 \times 3 \times 3$ mm$^3$ spatial resolution, and in a structural (or anatomical) image with a $1 \times 1 \times 1$ mm$^3$ resolution used to accurately localize functional activations. Note that a 3D image is in fact an array of many voxel intensities. Various pre-processing steps are required to correct functional images from possible head subject movement, to realign functional and anatomical individual images and, for group studies, all individual data sets in a common referential. Spatial smoothing (e.g., using a Gaussian kernel) is generally applied to functional images to compensate for potential mis-realignment and enhance the signal-to-noise ratio.

Several frameworks have been proposed to date for the statistical analysis of these pre-processed sets of functional data; see Lazard (2008) for a recent review. The commonly used statistical approach, massively univariate, considers each voxel independently from each other using regression techniques; see Friston *et al.* (1995) and Bullmore *et al.* (1996). It is available in freeware packages such as **FSL** (see `http://www.fmrib.ox.ac.uk/fsl/` and Smith *et al.* 2004), **SPM** (see `http://www.fil.ion.ucl.ac.uk/spm/` and Friston *et al.* 2007), **BrainVISA** (see `http://brainvisa.info/` and Rivière *et al.* 2009) or **NIPY** (see `http://nipy.sourceforge.net/` and Millman and Brett 2007). The time series response at each voxel is modeled as a stationary linear filter where the finite impulse response corresponds to a model of the HRF. This leads to the specification of a general linear model (noted GLM thereafter; not to be confounded with the generalized linear model) where the columns of the design matrix yield the regressors, i.e., a sampling of the convolution of the stimulus time courses with an HRF model on a grid consisting of the acquisition time points. Other regressors can be seamlessly introduced to model possible confounds. Many refinements to this approach have been proposed; see Nichols and Holmes (2002), Friston *et al.* (2005) and Roche *et al.* (2007). Spatial smoothness of the activated areas, normal distribution, serially correlated errors modeled in general as an AR(1) process and a predefined form of the HRF used as a convolution kernel are the main *a priori* incorporated into the GLM. This model-driven approach allows to test, using standard Student or Fisher tests, the activated regions against a desired hypothesis by specifying linear combinations of regressors. It is largely used essentially because of its flexibility in model specification allowing to test various hypothesis represented in corresponding statistical parametric maps. Clearly, the validity of the interpretation of these maps depends on the accuracy of the specified model.

An alternative exploratory (data-driven) approach relies on multivariate analysis based on independent component analysis (ICA). ICA performs a blind separation of independent sources from a complex mixture of many sources of signal and noise. In this approach, relying on the intrinsic structure of the data, no assumptions about the form of the HRF or the possible causes of responses are introduced. Only the number of sources or components to search for could eventually be specified. To identify a number of unknown sources of signal, ICA assumes that these sources are mutually and statistically independent in space (sICA) or time (tICA). This assumption is particularly relevant to biological time-series, see Friston (1998). For fMRI data set analyses, sICA is preferred because temporal points (few hundreds, corresponding to each occurrence of a functional image acquisition) are small compared to spatial ones (more than $10^5$, corresponding to the number of voxels contained in a functional image) leading for tICA to a potentially computationnaly intractable problem. However, temporal ICA could be relevant for some neuroscientific applications where temporal independence of sources can be assumed, see Calhoun *et al.* (2001). It is thus surprising that tICA has not been applied to study fMRI data, but occasionally on very reduced portions of the brain. A possible explanation of this fact might be that a first step usually done in most, if not all, IC analysis is to begin with a PCA decomposition, which gives rise to heavy computations (in the temporal case). For example, Calhoun *et al.* (2001) wrote *"Note that tICA is typically much more computationally demanding than sICA for functional MRI applications because of a higher spatial than temporal dimension and can grow quickly beyond practical feasibility. Thus a covariance matrix on the order of $N^2$ (where $N$ is the number of spatial voxels of interest) must be calculated. A combination of increased hardware capacity as well as more advanced methods for calculating and storing the covariance matrix may provide a solution in the future."*. This being said, the two objectives of our paper are: (1) to describe an R package for temporal and spatial independent component analysis in the context of neuroimaging and (2) to propose the use of a classical linear algebra result, implemented in our package, to alleviate the intrinsic computational burden linked to temporal ICA.

The paper is structured as follows. First, in Section 2 we briefly describe the principle of temporal and spatial ICA in the context of fMRI data set analysis and detail our mathematical proposition for ensuring temporal ICA tractability. In Section 3, we describe the current version of the R package **AnalyzeFMRI** (Marchini and Lafaye de Micheaux 2010), which has been the first R package designed for the processing and analysis of large anatomical and functional MRI data sets. In Section 4, we report results using synthetic data and real MRI data sets coming from human visual experiments, obtained using temporal or spatial ICA (TS-ICA). Finally, we conclude about the interest of the **AnalyzeFMRI** package and our extensions for the exploration of MRI data and outline our plans for future extensions.

## 2. Spatial and temporal independent component analysis

Independent component analysis (ICA) is a statistical technique, well developped in the signal processing community, whose aim is to recover hidden underlying source signals from an observed mixture of these sources (blind source separation problem). In standard ICA, one considers the mixture as linear and the sources as statistically mutually independent and non Gaussian.

Hereafter, we will adopt the standard notation available in the statistical literature for data matrices, where each row consists in a set of measurements of several statistical variables. Note that this is non standard in the ICA community where data matrices are transposed, see e.g., Hyvärinen *et al.* (2001). Moreover, capital slanted letters denote matrices ($A$ say), vectors are specified as a one column matrix written in bold face type ($\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ say), and upright unslanted characters (X or $\mathbf{X}$ say) denote random variables or vectors.

The generative *linear instantaneous noise-free mixing ICA parametric model* is generally written under the form

$$\mathbf{X} = A\mathbf{S} \tag{1}$$

where $\mathbf{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_m)^\top$ is the $m \times 1$ continuous-valued random vector of the $m$ observable signals, $A = (a_{ij})$ is the unknown constant (i.e., non random) and invertible square *mixing matrix* of size $m \times m$ and $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)^\top$ is the $m \times 1$ continuous-valued random vector of the $m$ unknown *source signals* to be recovered, hereinafter referred to as *the sources*.

Note that, if we denote by $B$ the inverse of matrix $A$, then we can write $\mathbf{S} = B\mathbf{X}$. The term "recover" here means that we want, based on an observed sample $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ (possibly organized in a matrix $X$ of size $n \times m$) of the random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$, to estimate the densities $f_{\mathrm{S}_j}$ of the $m$ sources $\mathrm{S}_j$, or to build a "pseudo-observed" sample of size $n$ of each one of these $m$ sources, which are usually called the *(maximally) independent (extracted) components*. For example, this sample could be computed, if one has obtained an estimate $\widehat{B}$ of the unknown *separating matrix* $B$, as $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ where $\boldsymbol{s}_i = \widehat{B}\boldsymbol{x}_i$, $1 \leq i \leq n$.

Using the independence hypothesis about the $m$ sources, note that the density $f_{\mathbf{X}}$ of the random vector $\mathbf{X}$ can be expressed as

$$f_{\mathbf{X}}(\boldsymbol{x}) = |A^{-1}| f_{\mathbf{S}}(\boldsymbol{s}) = |B| \prod_{j=1}^{m} f_{\mathrm{S}_j}(s_j),$$

where $|\cdot|$ is the determinant operator. It then follows that one can write the log-likelihood of the observed sample as

$$\log\ \mathcal{L}(B) = \log \prod_{i=1}^{n} f_{\mathbf{X}}(\mathbf{X}_i) = n \log|B| + \sum_{i=1}^{n}\sum_{j=1}^{m} \log\ f_{\mathrm{S}_j}(\boldsymbol{b}_j^\top \mathbf{X}_i)$$

where $\boldsymbol{b}_j$ denotes the $j$-th column of $B$. This is easy to prove when one notices that $\mathrm{S}_j = \boldsymbol{b}_j^\top \mathbf{X}$.

Optimization algorithms are then used for $\widehat{B} = \operatorname{argmax} \log \mathcal{L}(B)$ computation. To perform this operation, prior densities for the sources, or a simple parametrization of the sources may be considered; see details in Hyvärinen *et al.* (2001, p. 205–206).

Alternative approaches, not necessarily based on the likelihood function, are available to estimate $B$ (and thus $\mathbf{S}$). For instance, there is a relation between independence and non gaussianity; see e.g., Cardoso (2003) or the extensive literature on projection pursuit. In our package, we used the **FastICA** algorithm that searches for (maximally) non Gaussian sources, where non gaussianity is measured using a kurtosis measure, see Hyvärinen *et al.* (2001).

To apply standard ICA techniques on fMRI data sets, the first step is to obtain 2D data matrices $X$ from 4D data arrays (concatenation in time of several 3D functional volumes). This can be performed in two (dual) ways:

(a) one may consider that the data consist in the realization of $t_l$ random variables, each one measured (sampled) on $v_l$ voxels. This results in $t_l$ 3D spatial maps of activation. Each 3D map is then unrolled (in an arbitrary order) to get a matrix $X$ of size $v_l \times t_l$. The mixing matrix $A$ is in this case of size $t_l \times t_l$.

(b) one may consider that the data consist in the realization of $v_l$ random variables, each one measured at $t_l$ time points. This results in $v_l$ time courses each one of length $t_l$, collected into a matrix $X$ of size $t_l \times v_l$ (here again, the order of the $v_l$ time courses in the resulting matrix is arbitrary). The mixing matrix $A$ is in this case of size $v_l \times v_l$.

Using these data, the empirical counterpart of the noise-free model 1 can then be written as

$$\dot{X}^\top = AS^\top \tag{2}$$

where $\dot{X}$ is the centered version of matrix $X = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix}$, and where $S = \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_n \end{bmatrix}$.

Case (a) corresponds to *spatial ICA* (sICA) and the rows of matrix $S^\top$ contain pseudo-observations of spatially independent source signals of length $n = v_l$ (unrolled source spatial maps). Case (b) corresponds to *temporal ICA* (tICA) and the rows of matrix $S^\top$ contain here pseudo-observations of temporally independent source signals of length $n = t_l$ (source time courses). Recall that the row-dimension of matrices $X$ and $S$ above corresponds to sample size, which is not standard in the neuroimaging or signal processing community where those matrices would be transposed.

At this point, because the mixing matrix $A$ is square in standard ICA (Hyvärinen *et al.* 2001, p. 267), in writing (2) we have implicitly supposed that the number of sources $m$ is equal to $t_l$ in case (a) and $v_l$ in case (b). This is not necessarily the case. Data pre-processing based on PCA is generally used to overcome this problem. Doing this, model 2 should be re-written as

$$\Lambda_{red}^{-1/2} E_{red}^\top \dot{X}^\top = \widetilde{A} S^\top \tag{3}$$

where $\Lambda_{red}$ (resp. $E_{red}$) is the (*reduced*) matrix whose diagonal elements (resp. columns) consist of the $m$ largest (non null) eigenvalues (resp. eigenvectors) of the empirical covariance (or eventually correlation) matrix $\dot{X}^\top \dot{X}/n$ (note that the mixing matrix into $\dot{X}^\top$ is then given by $A = E_{red}\Lambda_{red}^{1/2}\widetilde{A}$). The size of the matrix $E_{red}$ is respectively $t_l \times m$ in case (a) and $v_l \times m$ in case (b). Note also that the Singular Value Decomposition (SVD) of matrix $\dot{X}$ can be used:

$$\dot{X} = UDV^\top$$

Then, one can replace in Equation 3, $\Lambda_{red}$ with $D_{red}^2/n$ where $D_{red}$ is the diagonal matrix consisting of the $m$ largest singular values of $D$, and $E_{red}$ with $V_{red}$ refering to the associated singular vectors. Equation 3 then leads to the following decomposition:

$$\dot{X}^\top = \frac{1}{\sqrt{n}} V_{red} D_{red} \widetilde{A} S^\top = \sum_{j=1}^{m} A_{\bullet j} \otimes S_{\bullet j}, \tag{4}$$

where $S_{\bullet j}$ denotes the $j$-th column of a matrix $S$. Note that the pair $(A_{\bullet j}, S_{\bullet j})$ is sometimes (abusively) called the $j$-th independent (estimated) component, although this term should be

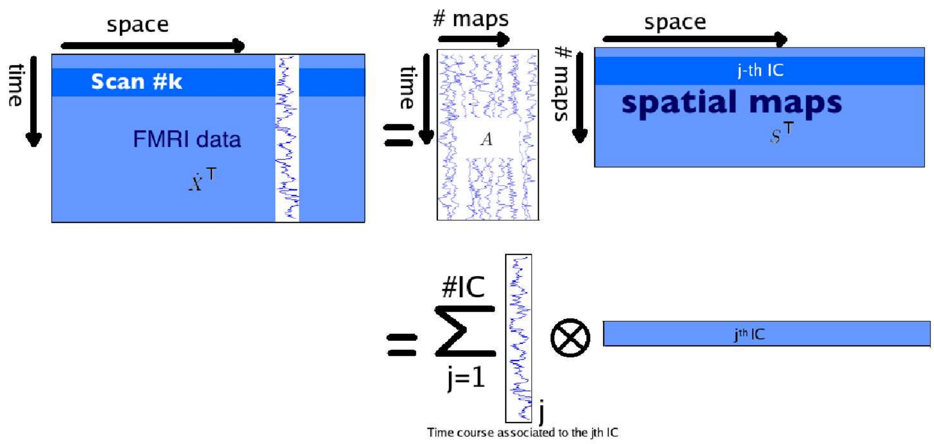


Figure 1: Illustration of the sICA decomposition after a PCA pre-processing step.

used solely for $S_{\bullet j}$, whereas $A_{\bullet j}$ refers to the weighting coefficients (degree of expression) of the $j$-th spatial component over time (for sICA) or of the $j$-th temporal source over space, i.e. over the voxels (for tICA).

Figure 1 below is an illustration of Equation 4 for sICA.

Unfortunately, due to the large number of voxels in fMRI experiments, it is not computationally tractable to fully diagonalize the correlation matrix in the temporal case (which is in this case of size $v_l \times v_l$) to perform the aforementionned PCA pre-processing step. *This may explain why tICA, as far as we know, has never been applied on the entire brain volume but only on parts of it*, see Calhoun *et al.* (2001), Seifritz *et al.* (2002) and Hu *et al.* (2005).

Our extension to the R package **AnalyzeFMRI** for TS-ICA uses a nice property of the SVD decomposition that allows to ensure tractability, see Jolliffe (1986) for other applications of this decomposition. It essentially states that the largest eigenvalues of the (huge) covariance matrix in the temporal case, as well as their associated eigenvectors, can be obtained from the same quantities computed from the (small) covariance matrix in the spatial case.

**Proposition.** *We consider the temporal case, where the size of the matrix $\dot{X}$ is $t_l \times v_l$. Let $S_X = \dot{X}^\top \dot{X}/t_l$ be the (empirical) covariance matrix of $\dot{X}$, with (large) size $v_l \times v_l$. The $r$ nonzero largest eigenvalues $\lambda_1, \ldots, \lambda_r$ of $S_X$ and their associated eigenvectors $\boldsymbol{f}_k$, $k = 1, \ldots, r$, are given by*

$$\lambda_k = \frac{1}{t_k} d_k^2 \text{ and } \boldsymbol{f}_k = \frac{1}{d_k} \dot{X}^\top \boldsymbol{g}_k$$

*where $d_k^2$ is the nonzero $k$-th largest eigenvalue of the (small) matrix $\dot{X}\dot{X}^\top$ of size $t_l \times t_l$ and where $\boldsymbol{g}_k$ is its associated eigenvector.*

*Proof.* We want to find the $r$ nonzero largest eigenvalues of $S_X$ and their associated eigenvectors $\boldsymbol{f}_k$, $k = 1, \ldots, r$. SVD theory allows us to write

$$\dot{X}^\top \dot{X} \boldsymbol{f}_k = d_k^2 \boldsymbol{f}_k \qquad k = 1, 2, \ldots, r; \tag{5}$$

$$\dot{X}\dot{X}^\top \boldsymbol{g}_k = d_k^2 \boldsymbol{g}_k \qquad k = 1, 2, \ldots, r. \tag{6}$$

Pre-multiplying Equation 6 by $\dot{X}^\top$, one can see that $\dot{X}^\top \boldsymbol{g}_k$ is an eigenvector of $\dot{X}^\top \dot{X}$ associated with the eigenvalue $d_k^2$. Thus, $\boldsymbol{f}_k$ is proportional to $\dot{X}^\top \boldsymbol{g}_k$.

The idea is thus to compute the $t_l$ eigenvalues $\{d_1^2, \ldots, d_{t_l}^2\}$ and the $t_l$ eigenvectors $\boldsymbol{g}_k$ of the (small) matrix $\dot{X}\dot{X}^\top$ of size $t_l \times t_l$. From this point, we get the $t_l$ first eigenvectors $\boldsymbol{f}_k$ (among the $v_l$ ones) of $S_X$ using this formula:

$$\boldsymbol{f}_k = \frac{1}{d_k} \dot{X}^\top \boldsymbol{g}_k. \tag{7}$$

The $v_l$ eigenvalues of $S_X$ are given by $\frac{1}{t_l} d_1^2, \ldots, \frac{1}{t_l} d_{t_l}^2, 0, \ldots, 0$. Note that the last $v_l - t_l$ eigenvectors of $S_X$ cannot be obtained using this approach, but anyway, as $d_i^2 = 0$ $(i > t_l)$ they do not contain any useful information. $\square$

## 3. The AnalyzeFMRI package

R package **AnalyzeFMRI** is a package for the exploration and analysis of large 3D MR structural data sets and 3D or 4D MR functional data sets. It was initiated by Marchini (2002) and transfered in 2007 to the third author of this paper. Sources, as well as precompiled binaries of our package for Mac or Microsoft Windows system, can be obtained from the Comprehensive R Archive Network (CRAN) at `http://CRAN.R-project.org/package=AnalyzeFMRI`. Note that our package is briefly described in the CRAN task view "Medical Imaging" (Whitcher 2010), as well as several other R packages for fMRI data analysis such as **arf3DS4** (Weeda *et al.* 2011), **cudaBayesreg** (Ferreira da Silva 2011), **fmri** (Tabelow and Polzehl 2011), **oro.nifti** (Whitcher *et al.* 2011), **Rniftilib** (Granert 2010) and **tractor.base** (Clayden *et al.* 2011). To efficiently explore fMRI data sets using tICA and sICA we added several interesting extensions to the initial package (e.g., tICA, automatic choice of the number of components to extract or GUI (graphical user interface) visualization tool). Some of them are briefly described below, see Whitcher *et al.* (2010) for more details, and also Marchini (2002) for a description of initial functions. Table 1 describes seven important functions available in the package.

*Importing data:* The package now provides read and write capabilities for the new NIfTI (`nii` or `hdr/img` files) format. This format contains a header gathering all the volume information (image dimension, voxel dimension, data type, orientation, quaternions, ..., up to more than 40 parameters) and a data part that contains values corresponding to the MR signal intensity measured at each voxel of the image object.

*Data pre-processing:* Briefly, before doing any statistical analysis, functional MR data should be corrected from geometric distortions, realigned and smoothed. Only the latter step is embedded into the current (and initial) version of the package.

*Image operators:* Several operators can be applied on the images such as rotation, translation, scaling, shearing or cropping. These operations can be performed by changing quaternion parameter in the NIfTI header or by direct modification of the matrix values. The matrix indices (voxel position) can be translated to volume coordinates (in mm) to facilitate comparison between subjects.

*Data analysis using TS-ICA:* In the initial version of the package, it was only possible to analyze fMRI data using spatial ICA. We added temporal ICA and the automatic detection of the number of components to extract. Automatic detection is based on an heuristic strategy:

| R function | Description |
|---|---|
| `f.analyzeFMRI.gui()` | Start an R-Tcl/Tk-based GUI to explore, using the **AnalyzeFMRI** package functions, an fMRI data set stored in ANALYZE format. |
| `f.icast.fmri.gui()` | The GUI provides a quick and easy to use interface for applying spatial or temporal ICA to fMRI data sets in NIfTI format. |
| `f.plot.volume.gui()` | TclTk GUI to display functional or structural MR images. This GUI is useful for instance to display the results performed with `f.icast.fmri.gui()`. |
| `f.read.header(file)` | Read ANALYZE or NIfTI (`.hdr` or `.nii`) header file. The format type is automatically detected by first reading the magic field. |
| `f.read.volume(file)` | Read ANALYZE or NIfTI image file and puts it into an array. Automatic detection of the format type. |
| `f.write.analyze(mat,file,...,)` | Store the data in ANALYZE format: creation of the corresponding `.img/.hdr` pair of files. |
| `f.write.nifti(mat,file,size,...)` | Store the data in NIfTI format: creation of the corresponding `.img/.hdr` pair of files or single `.nii` file. |

Table 1: Seven main functions of our package with their description.

the computation of the eigenvalues of the empirical correlation matrix of the data, keeping only those greater than 1. This rule is known as the Kaiser's rule, see (Jolliffe 1986, p. 95). The automatic detection is useful when no *a priori* knowledge is available. Note also that the user can now insert *a priori* knowledge in selecting only a specific region of the brain to explore (*via* a mask image) or in searching for components correlated with a specific time course signal.

*Visualization:* Anatomical or functional volumes and statistical (parametric or not) maps can be displayed in two separate windows with linked cursors to localize a specific position (see Figure 2). Our visualization tool can be used in two ways. First, you can use it to visualize the results of a temporal or spatial ICA (as displayed in Figure 2 for sICA). The time slider here indicates the rank of the component currently visualized (among all those extracted) and the displayed time course represents the values of the spatial component for the selected voxel (blue circle). Second, you can use it to visualize raw fMRI data. In this case the time slider represents the time course of the selected voxel, i.e. the MR signal values across time measured at the voxel position.

## 4. Results

We evaluated the TS-ICA part of the **AnalyzeFMRI** package both on simulated data and real data sets coming from human visual fMRI experiments.

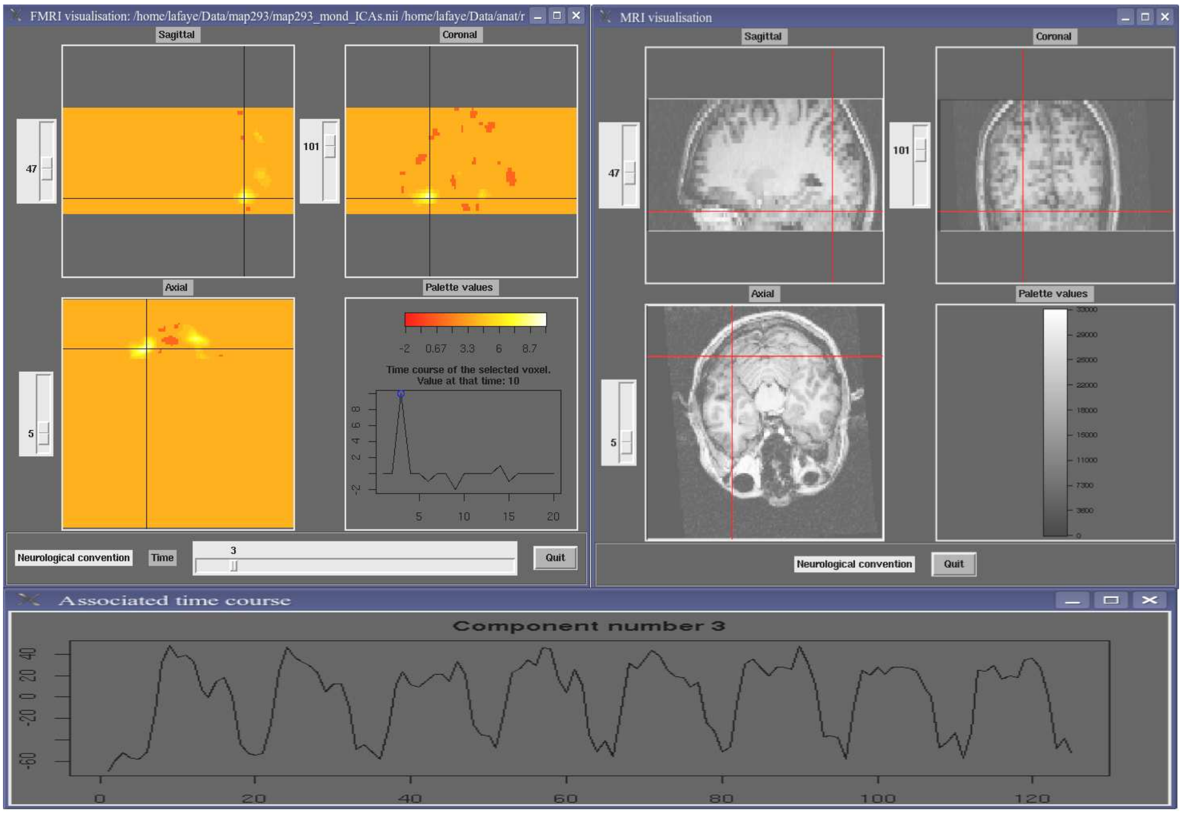


Figure 2: Image Display. Right top: Anatomical image (clockwise: sagittal, coronal and axial views). Left top: statistical map of activations obtained after spatial IC analysis of the functional data sets in the sagittal, coronal and axial orientation. The value of the selected extracted spatial component (here rank = 3) for the selected voxel (blue cross) is indicated in the right bottom quadrant (blue circle). The localization of the selected voxel is reported on the anatomical image (red cross). Bottom: Time course of the weighting coefficients of the third component (identical for all the voxels of this component).

## 4.1. Simulated data sets

In fMRI experiments, three standard paradigms are used. “Block design” which alternates, in a fixed order, stimuli that last few seconds; “event-related design” which alternates, in a random or pseudo-random order, stimuli that last few milliseconds and “phase-encoded paradigm” that generates traveling periodic waves of activation with different phases. In order to detect patterns of activation for the two former cases, we can use respectively a cross correlation with a square wave, or a binary cross correlation (to be defined later) with a sequence of 0 and $\pm 1$ representing the stimulation conditions. A Fourier analysis is more suitable for the latter.

Before testing our method on real data sets, we used three simulated cases: (1) a simple case to check that our method behaves correctly (this part is only provided and detailed as supplementary material; see supplementary file `simul1.zip` for details), and two cases simulating real conditions: (2) an event-related design simulation (supplementary file `simul2.zip`) and (3) a phase-encoded simulation (supplementary file `simul3.zip`). The former simulates the

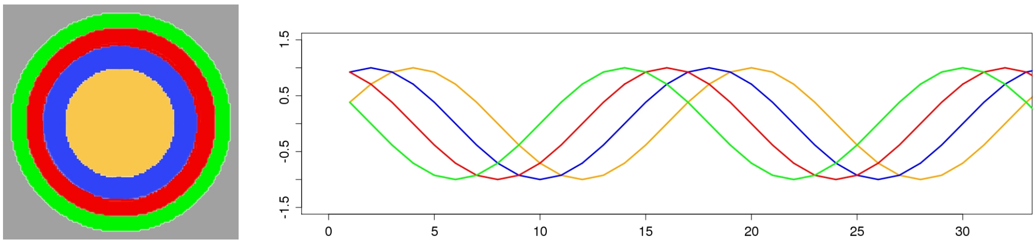


Figure 3: Simulated data set. Left: A transverse slice of the volume. Each color indicates the localization of each signal. Right: Time course of each temporal sequence of Bernoulli trials displayed in their corresponding color: source 1, orange, 9 events; source 2, blue, 17 events; source 3, red, 11 events; and source 4, green, 7 events.

"color center experiment" and the latter "retinotopic mapping experiment" reported respectively in Section 4.2.

R source code (including comments) for each one of the three aforementioned simulations is provided as supplementary material. Because our final results may change due to the use of random numbers (simulated data and initial conditions for ICA algorithm), we provided, in our R code, the seeds we used for the random generators. This allows the reader to obtain exactly the same results as those presented.

### *Event-related simulation*

With event-related paradigm, neuroscientists search for voxels activated specifically by each type of stimulus. To perform a simulation in this context, we used 100 3D-images ($128 \times 128 \times 3$ voxels) composed of four *non-overlapping* and concentric tubes. Each tube contained a temporal sequence of Bernoulli random variables with various probabilities of success (see Figure 3). The background, which surrounds the tube at the periphery, contained a Gaussian noise (sd $= 0.05$). To be realistic, we also added everywhere a Gaussian noise (sd $= 0.1$).

These simulated data were obtained using the following R code:

```
R> source("simul2.R")
R> set.seed(16)
R> simul2()
```

At this point, three files are created: `originalSignal.txt`, `mask.nii` and `simulEvent.nii`. We then applied temporal and spatial ICA to these simulated data, using the following R code. For tICA with automatic choice of the number of components:

```
R> set.seed(1)
R> f.icast.fmri.gui()
```

Select `simulEvent.nii` with maskfile `mask.nii` then select `Temporal ICA` and click on the `Start` button. For sICA with automatic choice of the number of components:

```
R> set.seed(16)
R> f.icast.fmri.gui()
```

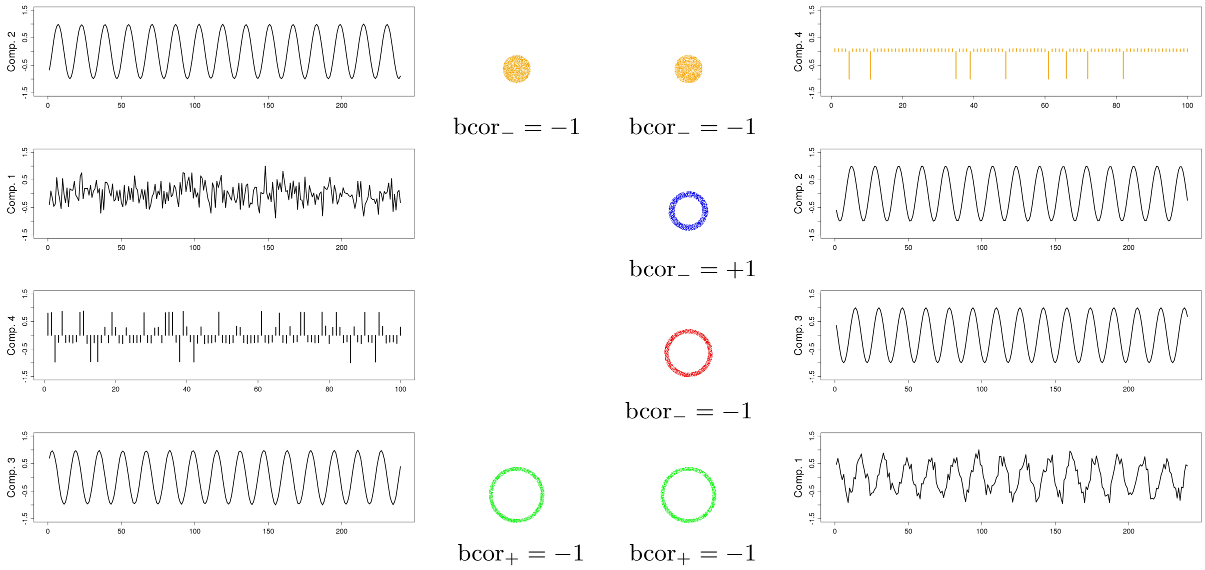


Figure 4: Time course and (thresholded) localization of the components detected using temporal ICA (left) and spatial ICA (right). Each ring indicates the thresholded localization of the component coded with the same color. For each component, the binary correlation coefficient of its time course with the corresponding initial signal is indicated. See Figure 3 for the correspondence with the exact position of the simulated signals. See text for the computation of the components localization. Each time course was normalized.

Select `simulEvent.nii` with maskfile `mask.nii` then select `Spatial ICA` and click on the `Start` button.

At this point, two pairs of files were created, namely:

- `simulEvent_ICAt.nii` and `simulEvent-ICAt-time-series.dat`,
- `simulEvent_ICAs.nii` and `simulEvent-ICAs-time-series.dat`.

Note that one can inspect temporal courses and spatial localizations using the GUI function `f.plot.volume.gui()`, on these two pairs of files respectively. The R instructions needed to perform the remaining part of the analysis are provided in the file `simul2-analysis.R` whose contents can be copied and pasted line by line in the R console, or which can possibly be sourced.

Figure 4 shows the time course and (thresholded) spatial localization of the 4 extracted components. For the latter, we used the following procedure. For each extracted time course $C_i$ ($1 \leq i \leq 4$), we considered either its positive part or its negative part, selecting the one having the highest peak of amplitude (in absolute value). Let's note $\tilde{C}_i$ the selection. Then, we computed the *binary correlation* (see (8) below) between each one of the original temporal signals of sources $S_j$ ($1 \leq j \leq 4$) and a thresholded version $\tilde{C}_{i[j]}$ of $\tilde{C}_i$. The thresholds used to obtain $\tilde{C}_{i[j]}$, $1 \leq j \leq 4$ were respectively 0.91, 0.83, 0.89 and 0.93 for the sources from the center to the periphery (see Figure 3). Intuitively, these thresholds correspond to the number of peaks of each original temporal signal among 100, i.e. 9 for source 1 (orange), 17 for source 2 (blue), 11 for source 3 (red) and 7 for source 4 (green). We define the binary correlation (number in $[-1, 1]$) between two (non necessarily positive) binary random sequences $u = (u_t, 1 \leq t \leq T)$ and $v = (v_t, 1 \leq t \leq T)$ by:

$$\mathrm{bcor}(u, v) = \frac{\sum_{t=1}^{T} \mathrm{sign}(u_t \times v_t)}{\sum_{t=1}^{T} \left(\mathrm{sign}|u_t| + \mathrm{sign}|v_t| - \mathrm{sign}|u_t \times v_t|\right)}. \tag{8}$$

Note that $\mathrm{sign}(0) = 0$.

We then assigned each extracted component to the original signal corresponding to the computation of the highest absolute value of the binary correlation (see Figure 4). For tICA, we found the following results:

- Components 1 and 2 were assigned with the temporal signal of source 1, with a binary correlation equal respectively to +1 and -1. As there was a conflict between the spatial localization given by these two components, we computed an "energy" index as follows. Let $\tilde{C}_1$ and $\tilde{C}_2$ be the two parts selected from the components 1 and 2 respectively. We then divide $\tilde{C}_1$ and $\tilde{C}_2$ respectively by $\max_{1\leq t\leq T} \tilde{C}_1$ and $\max_{1\leq t\leq T} \tilde{C}_2$ to obtain $C_1^*$ and $C_2^*$. Then we threshold $C_1^*$ and $C_2^*$ using the threshold $t_{12[1]}$ which is equal to half the empirical quantile of order 0.91 of the temporal signal $|C_1^*| + |C_2^*|$. The "energy" of component 1 versus component 2 to explain the source 1 is then given by the sum of the values in $|C_1^*|$ above $t_{12[1]}$. Similarly, the "energy" associated with component 2 is given by the sum of the values in $|C_2^*|$ above $t_{12[1]}$. The "energy" index was higher for component 2 (ratio of 0.58) which was consequently assigned to the temporal signal of source 1.

- Components 3 and 4 were assigned with the temporal signal of source 4, with a binary correlation equal respectively to -1 and +1. Here again, we computed the "energy" index which was higher for component 3 (ratio of 2.5) thus assigned to the temporal signal of source 4.

- Components 1 and 4 were consequently not associated with any source.

For sICA, we found the following results:

- Component 1 was assigned with the temporal signal of source 4, with a binary correlation equal to -1.

- Component 2 was assigned with the temporal signal of source 2, with a binary correlation equal to +1.

- Component 3 was assigned with the temporal signal of source 3, with a binary correlation equal to -1.

- Component 4 was assigned with the temporal signal of source 1, with a binary correlation equal to -1.

Surprisingly, spatial ICA works better in this case as compared to temporal ICA. We checked, using the R package **IndependenceTests** version 0.2 (for more information, see Bilodeau and Lafaye de Micheaux 2010, Beran *et al.* 2007 or Bilodeau and Lafaye de Micheaux 2005), the independence of our original random sequences of Bernoulli trials (note that this package can also check the independence of variables that are singular with respect to the Lebesgue measure).

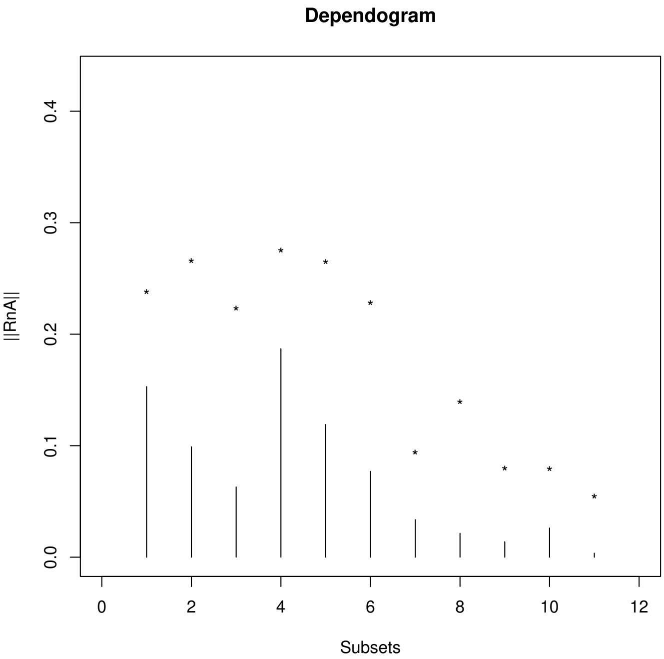


Figure 5: Test of the mutual independence between our four original signals.

```
R> library("IndependenceTests")
R> Orig <- read.table("originalSignal.txt", header = TRUE)
R> out <- dependogram(Orig, rep(1, 4), N = 10, B = 200)
```

There were no reason to significantly reject this independence hypothesis (at the 5% level), as can be seen on Figure 5. On the other side, using the same R instructions on the extracted temporal components, they were found significantly dependent.

A possible explanation to the tICA failure (notwithstanding the fact that standard ICA model is only defined for continuous random variables, since the unmixing and mixing matrix coefficients are real numbers and thus are not constrained in anyway to give binary values) may be the use of kurtosis in the FastICA algorithm, a quantity which is not optimal for sequences of Bernoulli trials, see Himberg and Hyvärinen (2001). More generally, regarding the power of ICA to extract independent *versus* sparse components, we refer the reader to Daubechies and Roussos (2009).

### *Traveling wave simulation*

We generated several sinusoids with the same fundamental frequency $f$=1/16 Hz and various phases to simulate traveling activation waves. The resulting data set consisted in a sequence comprising 240 3D-images. Each image ($128 \times 128 \times 3$ voxels) was composed of four *partially overlapping* and concentric tubes (each tube covers partly the tube that it surrounds). Each tube contained a pure sinusoidal signal (with a frequency $f$ equal to 1/16 Hz) in its non overlapping part and a sum of two pure sinusoidal signals in its parts that intersect with another tube. For pure signals, different phases were considered, namely $\phi_1 = 0$, $\phi_2 = \pi/4$,

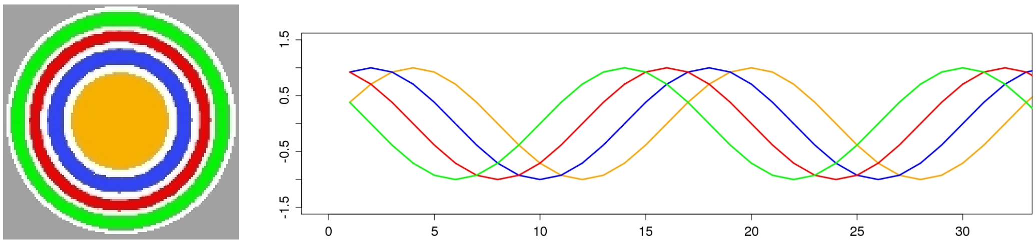


Figure 6: Simulated data set. Left: A transverse slice of the volume. Each color indicates the localization of each signal. Pure signals are represented with a color (orange, blue, red and green), Gaussian noise (sd = 0.2) is in grey and mixed signals are in white. Right: Temporal course of the pure single signals written in their corresponding color ($f$=1/16 Hz, phases = 0, $\pi/4$, $\pi/2$, $3\pi/4$ respectively from the center to the periphery).

$\phi_3 = \pi/2$ and $\phi_4 = 3\pi/4$ from the tube at the center to the one at the periphery respectively. The background, which overlaps the tube at the periphery, contained, in its non overlapping part, a Gaussian noise (sd = 0.2), see Figure 6. Thus, four pure signals and four mixed signals were present. To be realistic, we also added everywhere a Gaussian noise (sd = 0.1).

These simulated data were obtained using the following R code:

```
R> source("simul3.R")
R> set.seed(1)
R> simulSinusoid()
```

Here, three files are created: `originalSignal.txt`, `mask.nii` and `simulSinusoid.nii`.

Before going any further, it is convenient to think about a sinusoid waveform, which is deterministic in nature, as a sequence of different realizations of the same random variable $\mathrm{X} = \sin(2\pi \mathrm{U} f + \phi)$, where U is a continuous uniform random variable or, even better in the present case, a discrete uniform random variable on the sampled points. Note that, with standard algorithms, blind source separation is not concerned with the sequencing of the input signals. Indeed, changing the time ordering in which the mixtures are presented at the input leads to the same source separation (with the corresponding change in time indexing). This comment also applies to the two previous simulations. Note also that sinusoids with the same frequency but presenting different phases are in fact not independent. For example, correlation, which is a normalized version of the covariance, is not zero except for sinusoids with phase difference of $\pi/2$. Indeed, let $\mathrm{X} = \sin(2\pi \mathrm{U} f + \phi_1)$ and $\mathrm{Y} = \sin(2\pi \mathrm{U} f + \phi_2)$ be two random variables, where U is a discrete uniform random variable with support $\{a, a+1, \ldots, b-1, b\}$, i.e. with characteristic function $\varphi_{\mathrm{U}}(t) = \frac{e^{iat}}{n} \sum_{k=0}^{n-1} e^{ikt}$ where $n = b - a + 1$. We computed an explicit expression for the covariance $\mathbb{C}\mathrm{ov}(\mathrm{X}, \mathrm{Y}) = \mathbb{E}(\mathrm{XY}) - \mathbb{E}(\mathrm{X})\mathbb{E}(\mathrm{Y})$ between X and Y by showing that

$$\mathbb{E}(\mathrm{X}) = \mathrm{Im}\left[e^{i\phi_1} \frac{e^{ia2\pi f}}{n} \sum_{k=0}^{n-1} e^{ik2\pi f}\right] = \frac{1}{n} \sum_{k=0}^{n-1} \sin\left(\phi_1 + 2\pi f(a+k)\right)$$

and

$$\mathbb{E}(\mathrm{XY}) = \frac{1}{2}\left[\cos(\phi_1 - \phi_2) - \frac{1}{n} \sum_{k=0}^{n-1} \cos\left(\phi_1 + \phi_2 + 4\pi f(a+k)\right)\right].$$

We plotted $\mathbb{C}\mathrm{ov}(\mathrm{X}, \mathrm{Y})$ as a function of $\phi_1 - \phi_2$. It is not null but for $\pi/2$.

We then applied temporal and spatial ICA to these simulated data, using the following R code. For tICA with automatic choice of the number of components:

```
R> set.seed(1)
R> f.icast.fmri.gui()
```

Select `simulSinusoid.nii` with maskfile `mask.nii`, then select `Temporal ICA` and click on the `Start` button. For sICA with automatic choice of the number of components:

```
set.seed(1)
f.icast.fmri.gui()
```

Select `simulSinusoid.nii` with maskfile `mask.nii`, then select `Spatial ICA` and click on the `Start` button.

At this point, two pairs of files were created, namely:

- `simulSinusoid_ICAt.nii` and `simulSinusoid-ICAt-time-series.dat`,
- `simulSinusoid_ICAs.nii` and `simulSinusoid-ICAs-time-series.dat`.

Now, the R instructions needed to perform the remaining part of the analysis are provided in the file `simul3-analysis.R` whose content can be copied and pasted line by line in the R console, or which can eventually be sourced.

Three components were extracted for tICA as well as for sICA. As expected, temporal ICA extracted two components corresponding to sinusoids with phase difference of $\pi/2$. Spatial ICA does not impose any (independence) constraints on the extracted time courses. This is reflected in the results obtained for sICA.

## 4.2. Real data sets

We conducted two types of evaluation using real data sets coming from retinotopic mapping and color center mapping experiments. These data were part of a cognitive study investigating which color sensitive areas are specially involved with colors induced by synesthesia (Hupé *et al.* 2011). The real data sets used are provided as supplementary material.

### Experiment 1: Retinotopy mapping

Retinotopic mapping of human visual cortex using fMRI is a well established method (Sereno *et al.* 1995; Warnking *et al.* 2002) that allows to properly delineate low visual areas. It uses four separate experiments with 4 periodic stimuli (an expanding/contracting ring and a rotating counter or anti-counter clockwise wedge) to measure respectively eccentricity and polar angle maps. For this study, we only used functional MRI data corresponding to the expanding ring experiment (240 volumes acquired each 2 seconds). The periodic visual stimulus expanded from 0.2 to 3 degrees in the visual field during 32 seconds and was repeated fifteen times. This periodic stimulation generated a wave of activation in the retinotopic visual areas (Engel *et al.* 1994), located in the occipital lobe, at the frequency of 1/32 Hz measured at a discrete temporal sampling of 2 seconds (equivalent to 1/16 temporal bins).

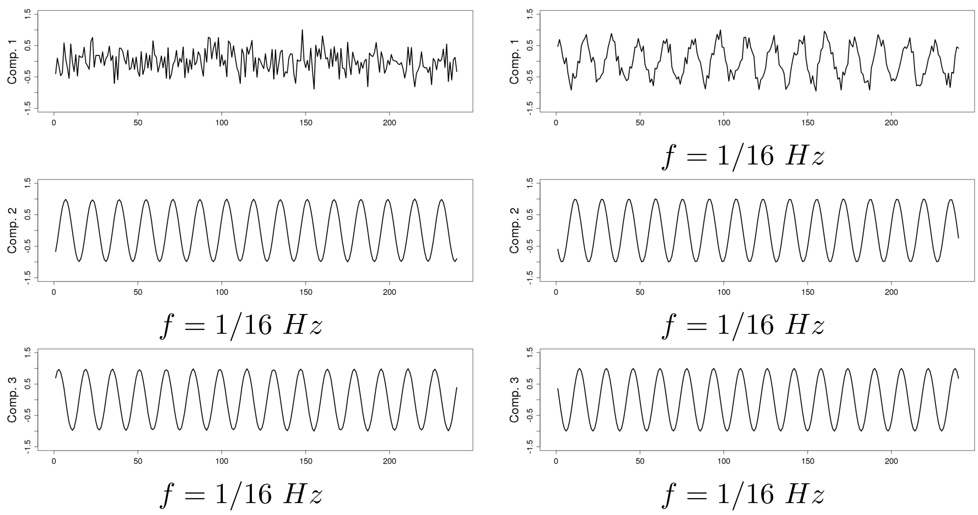


Figure 7: Time course of the components detected using temporal ICA (left) and spatial ICA (right). For tICA, the phase difference between components 3 and 2 is $\pi/2$. The first component (top left) represents a noise signal. For sICA, the phase difference between the temporal signal of source 1 and, respectively, the time courses of components 1, 2 and 3 are: 1.037, 3.377 and 2.385. ICA cannot recover the sign of the sources, so the phases of extracted sinusoidal signals are only defined modulo $\pi$. Each component was normalized.

With such a periodic stimulation paradigm, there are *per se* no specific advantages in using ICA over Fourier analysis. Phase encoded paradigm is used in our paper as a paradigmatic example to test the capabilities of our framework to detect known spatial and temporal components. After IC analysis of these functional data, using our R function `f.icast.fmri()` on the files `Retino4D.nii` and `RetinoMask.nii`, 18 and 13 components were automatically extracted respectively with tICA and sICA. In this experiment, we searched for components corresponding to cortical activation at the frequency of the visual stimulation. tICA and sICA extracted more (noisy) components than the ones specific to the stimulus. Indeed, the main problem with fMRI data is that each activated voxel of each volume contains a mixture of the signal of interest (BOLD effect) with several confound signals with several origins: ocular movement, heart rate, respiratory cycle, or head movement. Figure 8 shows the temporal and spatial components at the frequency of the visual stimulation corresponding to the cortical activation of interest. The computed phases are (approximatively) respectively equal to 2.90, -2.99 and 1.42 for tICA (green, red and blue components) and -2.44 and 2.03 for sICA (green and red components). These results can be obtained with the R code given in the file `retino-analysis.R` within the supplementary file `realDataProgs.zip`. Figure 8 displays on the corresponding anatomical image, the cortical localization of these extracted components. This was obtained with the **MRIcron** software (Rorden 2011). For tICA and for each temporal component, this was obtained by selecting in the associated column of the estimated mixing matrix (see Equation 4) the most active voxels, defined arbitrarily (see Beckmann and Smith 2004 for another approach) as those with a value above the 95% quantile (in absolute value). For sICA, we also thresholded arbitrarily each component at the 95% quantile. Based on the retinotopy property of the visual system, the expanding ring generates a cortical activation wave moving from the posterior part to the anterior part of the occipital lobe. As

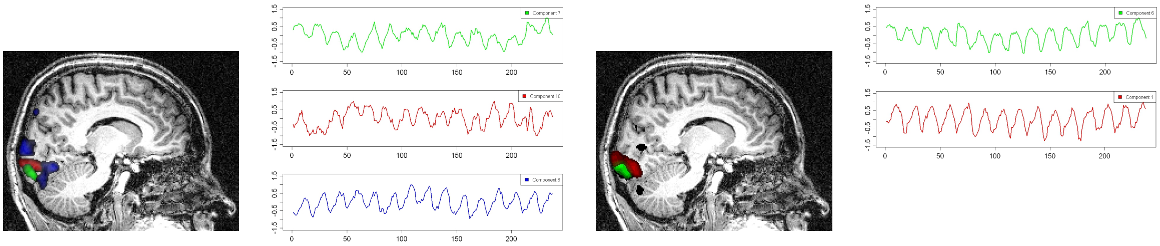

Figure 8: Time course of extracted components and their spatial localization. Left: extracted components using tICA. The phases of these components are found to be (approximatively) 2.90, -2.99 and 1.42 for green, red and blue components respectively. Right: extracted components using sICA. The phases of these components are found to be (approximatively) -2.44 and 2.03 for green and red components respectively. On the anatomical MR scan (sagittal view) is indicated the localization of the most activated voxels for each component displayed with the corresponding color. The sequencing of the activated voxels follows the retinotopic property of the visual system: the periodic visual stimulation, an expanding ring, generates a periodic cortical activation moving from the posterior to the anterior part of the occipital lobe.

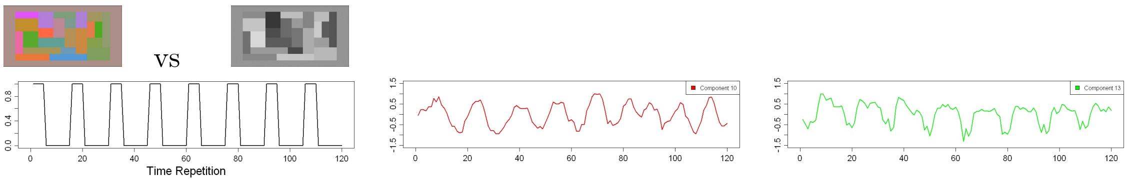


Figure 9: Color center mapping. Left: Visual stimulation alternated blocks presenting chromatic and achromatic version of Mondrian like patterns. Middle: Component at the frequency of the stimulation of the visual system extracted using temporal ICA. Right: Component at the frequency of the stimulation of the visual system extracted using spatial ICA.

indicated in Figure 8 the computed phases of the extracted components increase as expected from the posterior to the anterior part of the occipital lobe.

### Experiment 2: Color center mapping

In this experiment, we presented to the subject two stimuli, a set of chromatic rectangles (Mondrian like patterns) and the same patterns in an achromatic version (Figure 9). Each chromatic and achromatic sets of rectangles were periodically presented during successive blocks of 10 seconds. Our analysis was made on 120 functional volumes acquired each 2 seconds. Because we were only interested in visual areas, we used a mask to select only the occipital part of each volume. Using our R function `f.icast.fmri()` on the files `Mond4D.nii` and `MondMask.nii`, 21 and 20 components were automatically extracted with tICA and sICA respectively. As shown in Figure 9, we found both with tICA and sICA one periodic component at the frequency of the stimulus. These results can be obtained with the R code given in the file `mondrian-analysis.R` within the supplementary file `realDataProgs.zip`.

As shown in Figure 10, obtained with the MRIcron software, the most activated voxels containing the components extracted using temporal and spatial ICA were localized in a ventral

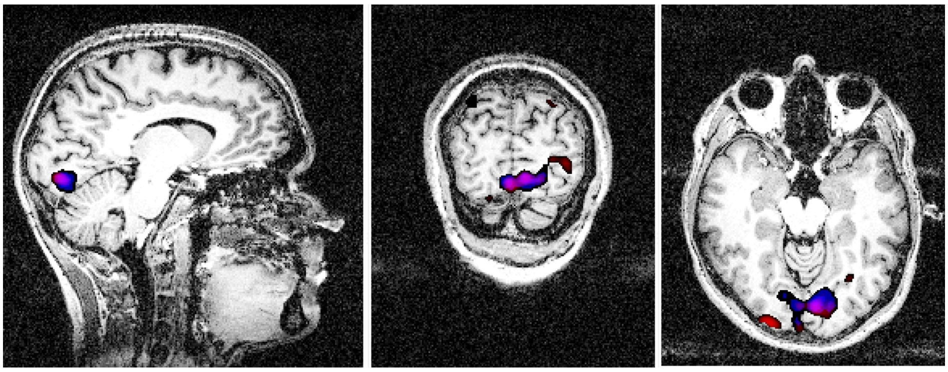

Figure 10: Localization of the most active voxels for the extracted component at the stimulus frequency using temporal (red) and spatial ICA (blue). Anatomical view: Left: sagittal; Middle: Coronal; Right: Transverse. There is a strong overlap between clusters corresponding to temporal and spatial analysis with a more accurate localization for the latter.

cortical part commonly called V4, V4H or V8 in Humans and known to be sensitive to color perception, see Wade *et al.* (2002).

## 5. Discussion

As an alternative to the standard model-driven approach (GLM), where the time course of the stimulus pattern is convolved with a hemodynamic response function and used as a predictor to detect brain activation, data-driven approaches such as ICA can reveal brain activation patterns with a good temporally and spatially accuracy or to extract noise components from the data, see McKeown *et al.* (2006). The strength of ICA is its ability to reveal hidden spatio-temporal structure when the definition of a specified *a priori* model is problematic. Since its first application to fMRI data analysis (McKeown *et al.* 1998), ICA have been used in various brain function studies. For example, ICA was successfully applied to investigate the cortical networks related to natural multimodal stimulation (Malinen *et al.* 2007) or natural viewing conditions (Bartels and Zeki 2004); situations in which activity is present in various brain sites and no *a priori* knowledge about the spatial location or about the activity waveforms were available. In Bartels and Zeki (2004), sICA allowed to segregate a multitude of functionally specialized cortical and subcortical regions because they exhibit specific differences in the activity time course of the voxels belonging to them. In Seifritz *et al.* (2002), tICA revealed un-predicted and un-modeled responses in the auditory system. Following Calhoun *et al.* (2001) or Malinen *et al.* (2007), GLM-derived activations are spatially less extensive and comprised only sub-areas of the ICA detected activations.

A number of ICA approaches have been proposed for fMRI data analysis. A comparison of some algorithms for fMRI analysis can be found in Correa *et al.* (2007). There are two largely used MATLAB toolboxes, **GIFT** (Rachakonda *et al.* 2010) implementing the FastICA algorithm (Hyvärinen 1999), which maximizes the non-gaussianity of estimated sources and **JADE**, which relies on a joint approximate diagonalization of eigenmatrices (Cardoso and Souloumiac 1993). Probabilistic ICA (PICA) is embedded in the **FSL** package (Beckmann and Smith 2004), a library of tools for neuroimaging data analysis (`http://www.fmrib.`

ox.ac.uk/fsl/). In this paper, we propose a new version of the R package **AnalyzeFMRI** including temporal and spatial IC analysis. We reused, with some memory improvements, the implementation of the FastICA algorithm proposed in the R package **fastICA**. Essentially for tractability considerations, spatial ICA is generally used in the context of neuroimaging. However, the temporal independence of sources can be supposed in some applications. In this case, only a small part of the brain is considered, see Calhoun *et al.* (2001); Seifritz *et al.* (2002). We have shown using a classical linear algebra result that temporal ICA can be tractable on large fMRI data sets. We have proposed an automatic selection of the extracted components by thresholding the eignevalues of the correlation matrix above 1 (rule known as the Kaiser's rule, see Jolliffe 1986, p. 95). More sophisticated strategies such as probabilistic ICA could be envisaged. Based on simulated data and real functional data sets, we have demonstrated the applicability of the package proposed for spatial and temporal ICA. As we have seen with the traveling wave case, sinusoids with the same frequency but presenting different phases are not independent and then can not be extracted using ICA. A possible solution to this problem would be to use a least square approach by imposing a strong *a priori* on the sources: the $i$-th source is $\mathrm{S}_i(t) = \sin(2\pi \mathrm{U}_t f + \phi_i)$ where the frequency $f$ is supposed to be known. We then search estimated sources $Y_1, ..., Y_m$ under this specific form that can be written as a linear combination of the observed signals: $Y_i = a_{i1}X_1(t) + a_{i2}X_2(t) + \ldots + a_{im}X_m(t), t = 1, \ldots, n$. The least square problem to optimize (numerically) is then

$$\sum_{t=1}^{n} \left(\sin(2\pi U_t f + \phi_i) - a_{i1}X_1(t) - a_{i2}X_2(t) - \ldots - a_{im}X_m(t)\right)^2, i = 1, \ldots, m.$$

Note that we could also differentiate with respect to the $\phi_i$'s and the $a_{ij}$'s to simplify the computation.

Several extensions should be inserted in the future. In order to select the more informative components we used a simple thresholding procedure. More refined strategies could be used such as this recently proposed in Varoquaux *et al.* (2010a). The package should be extended for dealing with group studies. Indeed, ICA generates a large number of components for each subject and obviously larger for a cohort of subjects. Several methods have been proposed for dealing specifically with group studies (Svensen *et al.* 2002; Esposito *et al.* 2005; Varoquaux *et al.* 2010b) and to facilitate the identification of components which are spurious or reproducible, see Himberg *et al.* (2004), Cordes and Nandy (2007), Ylipaavalniemi and Vigario (2008) and Wang and Peterson (2008). The sorting of relevant components can be performed using several indexes such as correlation coefficient with a reference function (Hu *et al.* 2005) or power spectrum (Moritz *et al.* 2003). A possible improvement would be to use a more general measure of dependence as the one provided in Beran *et al.* (2007).

## 6. Conclusion

To resume, ICA is a powerful data-driven technique that allows neuroscientists to explore the intrinsic structure of data and alleviate the need for explicit *a priori* about the neural responses. We propose with the TS-ICA extension to the R package **AnalyzeFMRI** a robust tool for the application both of spatial and temporal ICA to fMRI data.

## Acknowledgments

Cécile Bordier is recipient of a grant from Institut National de la Santé et de la Recherche Scientifique (INSERM). Pierre Lafaye de Micheaux is recipient of a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The authors would also like to thank Professor Christian Jutten for many helpful comments, as well as the referees and associate editors.

## References

Bartels A, Zeki S (2004). "The Chronoarchitecture of the Human Brain–Natural Viewing Conditions Reveal a Time-Based Anatomy of the Brain." *NeuroImage*, **22**(1), 419–433.

Beckmann CF, Smith SM (2004). "Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging." *IEEE Transactions on Medical Imaging*, **23**(2), 137–152.

Beran R, Bilodeau M, Lafaye de Micheaux P (2007). "Nonparametric Tests of Independence between Random Vectors." *Journal of Multivariate Analysis*, **98**(9), 1805–1824.

Bilodeau M, Lafaye de Micheaux P (2005). "A Multivariate Empirical Characteristic Function Test of Independence with Normal Marginals." *Journal of Multivariate Analysis*, **95**(2), 345–369.

Bilodeau M, Lafaye de Micheaux P (2010). ***IndependenceTests**: Nonparametric Tests of Independence between Random Vectors.* R package version 0.1, URL http://CRAN.R-project.org/package=IndependenceTests.

Bullmore E, Brammer M, Williams SC, Rabe-Hesketh S, Janot N, David A, Mellers J, Howard R, Sham P (1996). "Statistical Methods of Estimation and Inference for Functional MR Image Analysis." *Magnetic Resonance in Medicine*, **35**(2), 261–277.

Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001). "Spatial and Temporal Independent Component Analysis of Functional MRI Data Containing a Pair of Task-Related Waveforms." *Human Brain Mapping*, **13**(1), 43–53.

Cardoso JF (2003). "Dependence, Correlation and Gaussianity in Independent Component Analysis." *Journal of Machine Learning Research*, **4**, 1177–1203.

Cardoso JF, Souloumiac A (1993). "Blind Beamforming for Non Gausssian Signals." *Proceedings of the Institution of Electrical Engineers - Part F: Communications, Radar and Signal Processing*, **140**(6), 362–370.

Chen W, Ogawa S (1999). "Principles of BOLD Functional MRI." In C Moonen, PA Bandettini (eds.), *Functional MRI*, pp. 103–113. Springer-Verlag, Berlin.

Clayden JD, Muñoz Maniega S, Storkey AJ, King MD, Bastin ME, Clark CA (2011). "**TractoR**: Magnetic Resonance Imaging and Tractography with R." *Journal of Statistical Software*, **44**(8), 1–18. URL http://www.jstatsoft.org/v44/i08/.

Cordes D, Nandy R (2007). "Independent Component Analysis in the Presence of Noise in fMRI." *Magnetic Resonance Imaging*, **25**(9), 1237–1248.

Correa N, Adali T, Calhoun VD (2007). "Performance of Blind Source Separation Algorithms for fMRI Analysis Using a Group ICA Method." *Magnetic Resonance Imaging*, **25**(5), 684–694.

Daubechies I, Roussos Eand Takerkart S (2009). "Independent Component Analysis for Brain fMRI Does Not Select for Independence." *Proceedings of the National Academy of Sciences of the United States of America*, **106**(26), 10415–10422.

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994). "fMRI of Human Visual Cortex." *Nature*, **369**(6481), 525.

Esposito F, Scarabino T, Hyvärinen A, Himberg J, Formisano E, Comani S, Tedeschi G, Goebel R, Seifritz E, Di Salle F (2005). "Independent Component Analysis of fMRI Group Studies by Self-Organizing Clustering." *NeuroImage*, **25**(1), 193–205.

Ferreira da Silva AR (2011). "**cudaBayesreg**: Parallel Implementation of a Bayesian Multilevel Model for fMRI Data Analysis." *Journal of Statistical Software*, **44**(4), 1–24. URL http://www.jstatsoft.org/v44/i04/.

Friston KJ (1998). "Modes or Models: A Critique on Independent Component Analysis for fMRI." *Trends in Cognitive Sciences*, **2**, 373–375.

Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, Penny WD (eds.) (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, New York.

Friston KJ, Holmes AP, Poline JB, Frith CD, Frackowiak RSJ (1995). "Statistical Parametric Maps in Functional Imaging: A General Linear Approach." *Human Brain Mapping*, **2**, 189–210.

Friston KJ, Stephan KE, Lund TE, Morcom A, Kiebel S (2005). "Mixed-Effects and fMRI Studies." *NeuroImage*, **24**(1), 244–252.

Granert O (2010). ***Rniftilib**: R Interface to NIFTICLIB (V1.1.0)*. R package version 0.0-29, URL http://CRAN.R-project.org/package=Rniftilib.

Henson RN (2004). "Analysis of fMRI Timeseries: Linear Time-Invariant Models, Event-Related fMRI and Optimal Experimental Design." In RSJ Frackowiak, KJ Friston, C Frith, R Dolan, CJ Price, S Zeki, J Ashburner, WD Penny (eds.), *Human Brain Function*, 2nd edition, chapter 40, pp. 793–822. Elsevier.

Himberg J, Hyvärinen A (2001). "Independent Component Analysis For Binary Data: An Experimental Study." In *Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation (ICA2001), San Diego, California*, pp. 552–556.

Himberg J, Hyvärinen A, Esposito F (2004). "Validating the Independent Components of Neuroimaging Time Series via Clustering and Visualization." *NeuroImage*, **22**(3), 1214–1222.

Hu D, Yan L, Liu Y, Zhou Z, Friston KJ, Tan C, Wu D (2005). "Unified SPM-ICA for fMRI Analysis." *NeuroImage*, **25**(3), 746–755.

Hupé JM, Bordier C, Dojat M (2011). "The Neural Bases of Grapheme-Color Synesthesia Are not Localized in Real Color Sensitive Areas." *Cerebral Cortex*. Forthcoming.

Hyvärinen A (1999). "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis." *IEEE Transactions on Neural Network*, **10**(3), 626–634.

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

Jolliffe IT (1986). *Principal Component Analysis*. Springer-Verlag.

Lange N (2003). "What Can Modern Statistics Offer Imaging Neuroscience?" *Statistics and Methods in Medical Research*, **12**(5), 447–469.

Lazard NA (2008). *The Statistical Analysis of Functional MRI Data*. Springer-Verlag, Berlin.

Logothetis NK, Pfeuffer J (2004). "On the Nature of the BOLD fMRI Contrast Mechanism." *Magnetic Resonance Imaging*, **22**(10), 1517–1531.

Malinen S, Hlushchuk Y, Hari R (2007). "Towards Natural Stimulation in fMRI–Issues of Data Analysis." *NeuroImage*, **35**(1), 131–139.

Marchini J (2002). "**AnalyzeFMRI**: An R Package for the Exploration and Analysis of MRI and fMRI Datasets." *R News*, **2**(1), 17–23. URL http://CRAN.R-project.org/doc/Rnews/.

Marchini JL, Lafaye de Micheaux P (2010). ***AnalyzeFMRI**: Functions for Analysis of fMRI Datasets Stored in the ANALYZE or NIfTI Format*. R package version 1.1-13, URL http://CRAN.R-project.org/package=AnalyzeFMRI.

McKeown MJ, Makeig S, Brown GG, Jung T, Jindermann SS, Bell AJ, Sejnowski TJ (1998). "Analysis of fMRI Data by Blind Separation into Independent Spatial Components." *Human Brain Mapping*, **6**, 160–188.

McKeown MJ, Wang ZJ, Abugharbieh R, Handy TC (2006). "Increasing the Effect Size in Event-Related fMRI Studies." *IEEE Engineering in Medicine and Biology Magazine*, **25**(2), 91–101.

Millman KJ, Brett M (2007). "Analysis of Functional Magnetic Resonance Imaging in Python." *Computing in Science and Engineering*, **9**(3), 52–55.

Moritz CH, Rogers BP, Meyerand ME (2003). "Power Spectrum Ranked Independent Component Analysis of a Periodic fMRI Complex Motor Paradigm." *Human Brain Mapping*, **18**(2), 111–122.

Nichols TE, Holmes AP (2002). "Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples." *Human Brain Mapping*, **15**(1), 1–25.

Rachakonda S, Egolf E, Correa N, Calhoun V (2010). ***GIFT**: Group ICA of fMRI Toolbox*. MATLAB toolbox version 1.3, URL http://www.nitrc.org/projects/gift/.

Rivière D, Geffroy D, Denghien I, Souedet N, Cointepas Y (2009). "**BrainVISA**: An Extensible Software Environment for Sharing Multimodal Neuroimaging Data and Processing Tools." In *Proceedings of the 15th Human Brain Mapping Conference*.

Roche A, Meriaux S, Keller M, Thirion B (2007). "Mixed-Effect Statistics for Group Analysis in fMRI: A Nonparametric Maximum Likelihood Approach." *NeuroImage*, **38**(3), 501–510.

Rorden C (2011). ***MRIcro** Software Guide*. Version 1.40 build 1, URL http://www.MRIcro.com/.

Seifritz E, Esposito F, Hennel F, Mustovic H, Neuhoff JG, Bilecen D, Tedeschi G, Scheffler K, Di Salle F (2002). "Spatiotemporal Pattern of Neural Processing in the Human Auditory Cortex." *Science*, **297**(5587), 1706–1708.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995). "Borders of Multiple Visual Areas in Humans Revealed by Functional Magnetic Resonance Imaging." *Science*, **268**(5212), 889–893.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004). "Advances in Functional and Structural MR Image Analysis and Implementation as **FSL**." *NeuroImage*, **23**(Supplement 1), S208–S219.

Svensen M, Kruggel F, Benali H (2002). "ICA of fMRI Group Study Data." *NeuroImage*, **16**(3), 551–563.

Tabelow K, Polzehl J (2011). "Statistical Parametric Maps for Functional MRI Experiments in R: The Package **fmri**." *Journal of Statistical Software*, **44**(11), 1–21. URL http://www.jstatsoft.org/v44/i11/.

Thirion B, Pinel P, Meriaux S, Roche A, Dehaene S, Poline JB (2007). "Analysis of a Large fMRI Cohort: Statistical and Methodological Issues for Group Analyses." *NeuroImage*, **35**(1), 105–120.

Varoquaux G, Merlin K, Poline JB, Ciuciu P, Thirion B (2010a). "ICA-Based Sparse Feature Recovery from fMRI Datasets." In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1177–1180.

Varoquaux G, Sadaghiani S, Pinel P, Kleinschmidt A, Poline JB, Thirion B (2010b). "A Group Model for Stable Multi-subject ICA on fMRI Datasets." *NeuroImage*, **51**(1), 288–299.

Wade AR, Brewer AA, Rieger JW, Wandell AB (2002). "Functional Measurements of Human Ventral Occipital Cortex: Retinotopy and Color." *Philosophical Transactions of the Royal Society of London B*, **357**(1424), 963–973.

Wang Z, Peterson BS (2008). "Partner-Matching for the Automated Identification of Reproducible ICA Components from fMRI Datasets: Algorithm and Validation." *Human Brain Mapping*, **29**(8), 875–893.

Warnking J, Dojat M, Guerin-Dugue A, Delon-Martin C, Olympieff S, Richard N, Chehikian A, Segebarth C (2002). "fMRI Retinotopic Mapping – Step by Step." *NeuroImage*, **17**(4), 1665–1683.

Weeda WD, de Vos F, Waldorp L, Grasman R, Huizenga H (2011). "**arf3DS4**: An Integrated Framework for Localization and Connectivity Analysis of fMRI Data." *Journal of Statistical Software*, **44**(14), 1–33. URL http://www.jstatsoft.org/v44/i14/.

Whitcher B (2010). "CRAN Task View: Medical Image Analysis." Version 2010-11-30, URL http://CRAN.R-project.org/view=MedicalImaging.

Whitcher B, Lafaye de Micheaux P, Buchsbaum B, Polzehl J (2010). "Medical Image Analysis for Structural and Functional MRI." UseR! 2010 tutorial, URL http://useR2010.org/tutorials/Whitcher.html.

Whitcher B, Schmid VJ, Thornton A (2011). "Working with the DICOM and NIfTI Data Standards in R." *Journal of Statistical Software*, **44**(6), 1–28. URL http://www.jstatsoft.org/v44/i06/.

Ylipaavalniemi J, Vigario R (2008). "Analyzing Consistency of Independent Components: An fMRI Illustration." *NeuroImage*, **39**(1), 169–180.

**Affiliation:**

Cécile Bordier, Michel Dojat
Université Joseph Fourier, Grenoble-Institut des Neurosciences
(INSERM U836 and Université de Grenoble)
Bâtiment: Edmond J. Safra
Site Santé 38706 La Tronche, France
E-mail: Michel.Dojat@ujf-grenoble.fr, cecile.bordier@ujf-grenoble.fr
URL: http://nifm.ujf-grenoble.fr/~dojatm/

Pierre Lafaye de Micheaux
Department of Mathematics and Statistics
Université de Montréal
Montréal, Qc, Canada
E-mail: lafaye@dms.umontreal.ca
URL: http://www.neurostatistician.eu/

*Journal of Statistical Software* http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 44, Issue 9 *Submitted:* 2010-10-29
October 2011 *Accepted:* 2011-06-16