NeuroImage 169 (2018) 363–373

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

LEICA: Laplacian eigenmaps for group ICA decomposition of fMRI data

Chihuang Liu [a,*], Joseph JaJa [a], Luiz Pessoa [b]

[a] Institute for Advanced Computer Studies and Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
[b] Department of Psychology and Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Keywords:
fMRI
Unsupervised learning
Laplacian eigenmaps
ICA
Group inference
Functional networks

ABSTRACT

Independent component analysis (ICA) is a data-driven method that has been increasingly used for analyzing functional Magnetic Resonance Imaging (fMRI) data. However, generalizing ICA to multi-subject studies is non-trivial due to the high-dimensionality of the data, the complexity of the underlying neuronal processes, the presence of various noise sources, and inter-subject variability. Current group ICA based approaches typically use several forms of the Principal Component Analysis (PCA) method to extend ICA for generating group inferences. However, linear dimensionality reduction techniques have serious limitations including the fact that the underlying BOLD signal is a complex function of several nonlinear processes. In this paper, we propose an effective nonlinear ICA-based model for extracting group-level spatial maps from multi-subject fMRI datasets. We use a nonlinear dimensionality reduction algorithm based on Laplacian eigenmaps to identify a manifold subspace common to the group, such that this mapping preserves the correlation among voxels' time series as much as possible. These eigenmaps are modeled as linear mixtures of a set of group-level spatial features, which are then extracted using ICA. The resulting algorithm is called LEICA (Laplacian Eigenmaps for group ICA decomposition). We introduce a number of methods to evaluate LEICA using 100-subject resting state and 100-subject working memory task fMRI datasets from the Human Connectome Project (HCP). The test results show that the extracted spatial maps from LEICA are meaningful functional networks similar to those produced by some of the best known methods. Importantly, relative to state-of-the-art methods, our algorithm compares favorably in terms of the functional cohesiveness of the spatial maps generated, as well as in terms of the reproducibility of the results.

Introduction

Functional magnetic resonance imaging (fMRI) studies have shed light on the overall functional organization of the brain. For example, resting state studies have revealed the existence of a number of intrinsic functional networks, including the "default mode" and "salience" networks Biswal et al. (1995); Cole et al. (2014); Greicius et al. (2003). These studies have also led to the discovery of specific patterns related to brain disorders Jafri et al. (2008); Greicius et al. (2004); Garrity et al. (2007), age Damoiseaux et al. (2008), and gender Weissman-Fogel et al. (2010).

The determination of large-scale patterns of brain activity is a challenging problem due to several factors, including the high-dimensionality of the data, various noise sources, and inter-subject variability. The most widely used approach to capture group-level functional connectivity is based on Independent Component Analysis (ICA) McKeown et al. (1998), which is a data-driven approach that assumes the existence of statistically independent latent features, called sources, which can generate the data through linear mixtures. However, ICA does not naturally provide a suitable method for drawing inferences about groups of subjects Erhardt et al. (2011), which has led to the development of several multi-subject approaches that are partly based on ICA. These approaches revolve around a few strategies. One method is to concatenate the data from multiple subjects along the temporal dimension Calhoun et al. (2001) followed by the use of Principal Component Analysis (PCA) Jolliffe (2002) to reduce the dimensionality of the data, and ending with the application of ICA to extract spatial features common at the group level. The tensorial extension to ICA Beckmann and Smith (2005) uses a three-dimensional tensor to estimate shared spatial patterns and time courses between subjects, and subject-specific loadings of the components to capture the multidimensional structure of the data. In Guo and Pagnoni (2008), a general framework is proposed for using the expectation maximization algorithm to estimate an unconstrained mixing matrix for the group-level independent components. Finally, canonical ICA Varoquaux et al. (2010) uses Canonical Correlation Analysis (CCA) to find components that maximize cross-correlation before

* Corresponding author. Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA.
E-mail address: chliu@umd.edu (C. Liu).

https://doi.org/10.1016/j.neuroimage.2017.12.018
Received 21 July 2017; Received in revised form 5 December 2017; Accepted 9 December 2017
Available online 13 December 2017

applying the ICA step.

Two methods involving temporal concatenation of the data include GIFT Calhoun et al. (2001) (http://icatb.sourceforge.net/) and MELODIC Beckmann and Smith (2004) (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC/). GIFT approximates the PCA of the group temporally concatenated data by first reducing each subject's dataset to dimension m, and then running the PCA on the concatenation of the principal components across all subjects to further reduce temporal dimensionality. MELODIC, however, uses an incremental approach called MELODIC's Incremental Group PCA (MIGP) Smith et al. (2014) to implement the approximation. Although MELODIC and GIFT have been broadly used for group level fMRI analysis, they are not without limitations as will be explained next.

These ICA-based strategies involve the use of PCA to reduce the dimensionality of the data. PCA is basically a linear projection of the original data onto a low-dimensional subspace in such a way as to retain most of the variance of the data. However, PCA does not distinguish between variance due to noise versus variance due to inherent underlying signal variations. Moreover, there is no principled justification for using the variance as the basis for data reduction based solely on a linear transformation. The BOLD signal is a complex function of neural activity, oxygen metabolism, cerebral blood volume, cerebral blood flow, and other physiological parameters Wang et al. (2003). The dynamics underlying neural activity and hemodynamic physiology involve multiple nonlinearities Birn et al. (2001); Miller et al. (2001); Xie et al. (2008), a fact that has motivated the exploration of nonlinear approaches for fMRI analysis Thirion and Faugeras (2004); Mannfolk et al. (2010). To the best of our knowledge, nonlinear models have not been proposed to extract common spatial patterns for multi-subject fMRI analysis.

Several nonlinear dimensionality reduction approaches have been successfully used to address a wide range of applications Van Der Maaten et al. (2009). Kernel PCA Schölkopf et al. (1998) maps data to a feature space using a nonlinear transformation (kernel) and performs linear PCA in the feature space. Isomap Tenenbaum et al. (2000) is based on computing a low dimensional representation that tries to preserve the pairwise geodesic distances, which are approximated by the pairwise shortest paths between corresponding points. Maximum variance unfolding (MVU) Weinberger and Saul (2006) computes a low-dimensional manifold such that local distances and angles are preserved. Locally linear embedding (LLE) Roweis and Saul (2000) is based on the geometric intuition of local linearity, which assumes that each point and its neighbors lie on an approximately linear patch of a low-dimensional manifold. The mapping is computed by trying to preserve linearity locally, such that each point has the same neighborhood structure as in the high-dimensional space. A technique of a different flavor is the t-distributed stochastic neighbor embedding (t-SNE) Maaten and Hinton (2008), which captures similarities between pairs of high-dimensional points through a probability distribution, followed by determining a mapping to a lower-dimensional space that achieves a similar distribution. Laplacian eigenmaps Belkin and Niyogi (2003) project data into a low dimensional Euclidean space such that local proximity relations are preserved as much as possible, mapping close points in the original space to close points in the low-dimensional space. Of these techniques, Laplacian eigenmaps appear to be the most suitable for fMRI data since its overall objective function can be expressed as minimizing

$$\sum_{ij} \left\| \boldsymbol{y}_i - \boldsymbol{y}_j \right\|^2 W_{ij} \tag{1}$$

where $\boldsymbol{y}_i$, $\boldsymbol{y}_j$ are the low-dimensional vectors corresponding to points i, j in the original space, and W_{ij} is a proximity measure of points i and j in the original space. If we choose W_{ij} to be the zero-thresholded correlation between voxels i and j in the original space, this nonlinear projection maps voxels with strong correlations to nearby points in the low-dimensional Euclidean space. Correlations between voxels in fMRI data reveal functional relationships and hence contain the information that can be used to construct functional networks.

We note that Laplacian eigenmaps have been successfully used in analyzing fMRI data Marquand et al. (2017); Haak et al. (2016). Our approach differs from earlier studies in that they are ROI-based methods focusing on specific networks using prior brain-related knowledge, while our framework makes use of ICA to extract intrinsic networks for general fMRI datasets.

In this paper, we propose a group-level model that uses Laplacian eigenmaps as the main data reduction step, which preserves the correlation information in the original data as best as possible. The nonlinear map is robust relative to noise in the data and to inter-subject variability. After the Laplacian eigenmaps transformation, ICA is applied to the reduced data to estimate spatial features, followed by thresholding to extract group-level independent spatial maps. We call our algorithm LEICA (Laplacian Eigenmaps for group ICA decomposition).

The rest of the paper is organized as follows. In section 2, we present the details of our algorithm and inference method, and describe several tests used to validate our method and compare the results with state-of-the-art methods. In section 3, we show the corresponding test results on both resting state and working memory task fMRI datasets for 100 subjects from the Human Connectome Project. Finally, in section 4, we discuss our results.

Materials and methods

Group ICA model

Spatial ICA assumes that the observed fMRI data are generated by a linear mixture of spatially independent components. The underlying model can be expressed as:

$$X = MA \tag{2}$$

where X is the observation matrix whose columns represent the time series corresponding to the voxels, A consists of the corresponding independent sources, and M is the mixing matrix. We note that (2) does not include a noise model, but fMRI data is inevitably confounded by noise. Moreover, as stated, the model is not suitable for multi-subject fMRI analysis.

Noisy group level ICA models have been proposed in the literature. Let $\{X_n;\ n = 1,\ 2,\ \ldots,\ N\}$ be the fMRI data of N subjects with n being the subject index. X_n has dimensionality $T \times V$, where V is the number of voxels and T is the number of time frames. To handle noise, group ICA models have the following form:

$$X = UA + E \tag{3}$$

where $X = [X_1^T\ X_2^T\ \ldots\ X_N^T]^T$ is the $NT \times V$ data matrix formed by temporally concatenating data across N individuals, A consists of the group-level independent sources, U is the group mixing matrix, and E is the noise. Existing methods differ in the assumptions and procedures used to estimate the group mixing matrix U and the noise E.

In temporal concatenation ICA Calhoun et al. (2001); Smith et al. (2004), the group mixing matrix has form $U = PM$ and the group model becomes

$$X = PMA + E \tag{4}$$

where P is estimated by PCA on the concatenated group data matrix and MA consists of the group-level principal components, which are assumed to be a linear subspace of the sources. Low variance components are considered as noise. The classic noise-free ICA is then used to estimate the mixing matrix M and the independent sources A.

In the tensor ICA approach Beckmann and Smith (2005), the group mixing matrix is defined to have the Khatri-Rao product structure $U = M| \otimes |R$ where M is the subject-specific loading of the group components

and R is the time course associated with the independent components. The group model can now be expressed as

$$X = (M| \otimes |R)A + E \tag{5}$$

where the data are modeled as a trilinear product of group-level independent sources A, associated with common time courses R and subject-specific loadings contained in M, corrupted with additional noise E.

In the unified framework Guo and Pagnoni (2008), no assumption about the structure of U is made but instead it is estimated using the expectation maximization algorithm that maximizes the expected likelihood that the observed data is generated by this model. The noise is modeled by a zero-mean multivariate Gaussian distribution.

Canonical ICA Varoquaux et al. (2010) tries to handle noise in a hierarchical fashion in the sense that individual observation noise and subject variability noise are estimated successively. The group mixing matrix has the form:

$$U = \begin{bmatrix} P_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & P_N \end{bmatrix} \begin{bmatrix} Q_1 \\ \vdots \\ Q_N \end{bmatrix} M \tag{6}$$

where the P_i's are the loading matrices for individual principal components estimated by PCA, and the Q_i's are the loading matrices for the common group principal components estimated by CCA, which determines components that maximize cross-correlation.

We propose in this work a new model, called LEICA, based on the assumption that the linear mixture of independent sources lies in a low-dimensional manifold that cannot be captured by linear models. We do not make explicit assumptions about the group mixing matrix either, but address the noise problem through the construction of a similarity graph. Our overall model can be expressed as:

$$\begin{aligned} Y &= MA \\ X &= \phi^{-1}(Y) + E \end{aligned} \tag{7}$$

where ϕ is a nonlinear embedding. We propose to use the Laplacian eigenmaps algorithm to learn this embedding as explained next.

Dimensionality reduction using Laplacian eigenmaps

Laplacian eigenmaps constitute a nonlinear dimensionality reduction technique that is based on determining a low dimensional representation of a high dimensional dataset, which preserves locality information as much as possible. It maps similar data points to nearby points in Euclidean space.

We start by normalizing the data of each subject. Then we let $X = [\boldsymbol{x_1}\ \boldsymbol{x_2}\ \ldots\ \boldsymbol{x_V}]$, where $\boldsymbol{x_v} \in \mathbb{R}^{NT}$ is the concatenated time series for voxel v across the subjects. The procedure of using Laplacian eigenmaps to derive a d-dimensional embedding of the voxels is as follows.

First we construct a group-level weighted similarity graph $G = (V, E)$ with the nodes V representing the voxels and the weights of the edges E are pairwise similarities chosen to be the Pearson's correlation coefficient

$$r_{ij} = \frac{\left(\boldsymbol{x_i} - \overline{x_i}\right) \cdot \left(\boldsymbol{x_j} - \overline{x_j}\right)}{NT\sigma_i\sigma_j} \tag{8}$$

where $\overline{x_i}$, $\overline{x_j}$ are the sample means of $\boldsymbol{x_i}$, $\boldsymbol{x_j}$ and σ_i, σ_j are the sample standard deviations of $\boldsymbol{x_i}$, $\boldsymbol{x_j}$. To avoid explicitly considering the group-concatenated data, the group-level correlation matrix is computed by averaging the subject-level correlation matrices, as the data of each subject have already been normalized.

For each voxel $v \in [1, \ldots, V]$, we determine the k nearest neighbors set $\mathcal{N}(v)$ of v consisting of the k voxels with the largest correlations with v. We then compute the following weight function

$$W_{ij} = \begin{cases} r_{ij} & \text{if } i \in \mathcal{N}(j) \text{ or } j \in \mathcal{N}(i) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where r_{ij} is the Pearson's correlation coefficient of voxel i and j. By considering only positive correlations r_{ij}, negative correlations are not included in a voxel's k nearest graph, which also reflects the intuition that anti-correlated voxels are less likely to participate in the same network Fox et al. (2005).

The choice of k needs to be carefully considered. If k is chosen to be too small, the graph will not contain sufficient local connectivity strengths for all voxels, and thus the low-dimensional embedding will not capture enough information. In contrast, if k is too large, many weak connections will be included in the graph and the graph will be more subjective to noise and inter-subject variation. Picking a suitable k is a trade-off between connectivity information and noise, and the best value depends on the data. We will describe our approach later, which always leads to a sparse graph.

Once the graph G is constructed, we compute the normalized graph Laplacian

$$L = I - D^{-1/2}WD^{-1/2} \tag{10}$$

where D is diagonal degree matrix whose entries are defined by $D_{ii} = \sum_j W_{ij}$. The matrix L is also called the normalized Laplacian. Then we compute the smallest few eigenvalues and corresponding eigenvectors of L

$$L\boldsymbol{f} = \lambda\boldsymbol{f} \tag{11}$$

Let $\boldsymbol{f_0}, \boldsymbol{f_1}, \ldots, \boldsymbol{f_d}$ be the $d+1$ eigenvectors corresponding to the eigenvalues ordered as $0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_d$. The eigenvector $\boldsymbol{f_0}$ corresponding to eigenvalue 0 is a constant vector, so we use the other d eigenvectors to form embeddings of the original data. Vector $\boldsymbol{y_v} = [\boldsymbol{f_1}(v), \boldsymbol{f_2}(v), \ldots, \boldsymbol{f_d}(v)]^T \in \mathbb{R}^d$ is the d-dimensional embedding for $\boldsymbol{x_v}$.

Group-level spatial maps

After embedding the data into a lower dimensional manifold, we obtain a $d \times V$ matrix $Y = [\boldsymbol{y_1}\ \boldsymbol{y_2}\ \ldots\ \boldsymbol{y_V}]$ in which each row corresponds to an eigenvector as described in the previous section. We now model this matrix Y as a linear mixture of a number of spatial independent features as indicated in equation (7). The common independent spatial features A can be extracted by applying ICA.

The overall LEICA workflow is illustrated in Fig. 1.

The resulting spatial components extracted by the ICA are continuous values at each voxel, and hence some form of thresholding is employed to determine the final spatial maps. Group t-test based and mixture-model based thresholding methods have been proposed in the literature and used in GIFT and MELODIC; these methods are applicable to our model as well. But it has been shown that a simple thresholding approach based on the absolute values of the voxels also works well Varoquaux et al. (2010) and is consistent with the FastICA algorithm that we use. Therefore our thresholding procedure consists of converting the spatial components to Z-scores (zero mean and unit variance) and then thresholding based on absolute Z values.

Evaluation strategies

In this section, we introduce a number of approaches to evaluate our model, which will be applied to resting state and working memory data sets of the Human Connectome project Van Essen et al. (2013). Given the extracted group-level spatial maps, we show that thresholded spatial maps capture important qualitative properties of resting-state networks, as well as functional organization compatible with working memory task

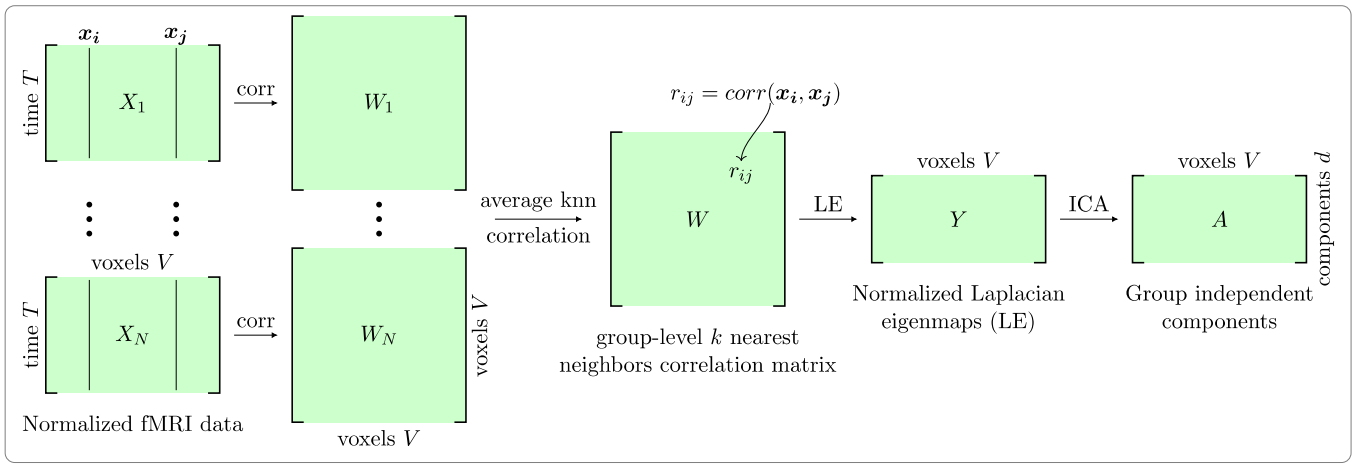


Fig. 1. Processing steps used by LEICA. The correlation matrices corresponding to individual subjects (1 through *N*) are computed first, followed by averaging across subjects, and constructing the group-level *k* nearest neighbors (*k*NN) correlation matrix (the latter is always a sparse matrix). This matrix is used to compute normalized Laplacian eigenmaps followed by the application of the ICA algorithm.

execution. Furthermore, the results are comparable to those obtained with state-of-the-art methods. Next, we show that our model is able to extract more functionally cohesive and stable spatial components than other methods, and exhibits better generalizability from smaller subgroups.

Similarity with other methods

The current state-of-the-art and most widely used tools to tackle the problem addressed here are GIFT and MELODIC. The two toolboxes implement the same concatenated ICA model with slight variations on how the group-level PCA estimation and thresholding are conducted. While their results are comparable, MELODIC seems to have better reproducibility due to the instability of GIFT's *t*-test based thresholding Varoquaux et al. (2010). Hence, we focus our comparison with MELODIC (but also show results for GIFT). Our model employs different estimation and inferencing procedures than MELODIC. In addition, the final values of the spatial maps in our model are *Z*-scores based on the absolute strengths of the ICA components, while the values in MELODIC are *Z*-statistics based on mixture modeling Beckmann and Smith (2004). Thus, we compare thresholded maps which are positive with different magnitude ranges in the two models. We use the cosine similarity measure to compare the corresponding thresholded maps, a commonly used similarity measure in high-dimensional positive spaces.

Specifically, after generating the spatial maps from both models, we apply the corresponding thresholding procedures (the threshold values are adjusted such that the spatial maps have approximately the same number of voxels on average). Let a_1, a_2 be the two thresholded spatial maps from the two models; the cosine similarity is defined as:

$$c = \frac{a_1 \cdot a_2}{\|a_1\|_2 \|a_2\|_2} \tag{12}$$

Functional connectivity structure

Functional MRI data have been used to capture large-scale functional connectivity structure in the brain. Voxels in the same intrinsic functional network are expected to have similar activity patterns, which translates into correlated voxel time series. In fact, most seed-based analysis Yeo et al. (2011) and clustering approaches Craddock et al. (2012) are based on this assumption. We develop an approach to capture the functional connectivity profile of each spatial map as follows.

Let the rows of the matrix A be a set of thresholded spatial maps, and let W be the full correlation matrix of all pairs of voxels, averaged over all subjects. The i-th row (or column) of W is the correlation between voxel i and the rest of the voxels. We define the correlation maps corresponding to the given thresholded spatial maps by:

$$A_s = A \cdot W \tag{13}$$

The i-th row of A_s defines in a certain sense the correlation profile of the i-th spatial map in A. That is, each component of this profile is the weighted sum of the correlations of all the voxels for the spatial component, weighted by the magnitude of the corresponding voxel. Note that our approach differs from commonly used measures such as the average correlation of the voxels in a region of interest (ROI) Di Martino et al. (2008); Jiang et al. (2004). Note also that the magnitude of a voxel in a spatial component indicates the importance of this voxel for this component. This weighted sum provides a normalized correlation map for each spatial component over all voxels. It should be pointed out, however, that using average and weighted average correlations generates very similar results.

After generating A_s, we can threshold A_s to maintain only the strongest correlations, followed by comparing the thresholded correlation map $\overline{A_s}$ with A. If A represents inherent functional connectivity networks of the brain, we expect the voxels in each spatial component to be highly correlated with the voxels in the same component, and we expect the correlations with the rest of the voxels to be relatively weak. Therefore, the thresholded correlation map $\overline{A_s}$ should be very similar to A. If $\overline{A_s}$ is not similar with A, this would indicate that there are voxels outside the component that are very correlated with the voxels in the component. Finally, the similarity measure we use to compare $\overline{A_s}$ and A is the average cosine similarity over all the corresponding pairs of spatial maps between $\overline{A_s}$ and A.

Reproducibility and consistency

We introduce two tests to evaluate the reproducibility and consistency of our method. Similar tests have been proposed in Varoquaux et al. (2010). Let A_1, A_2 be the matrices defining d extracted spatial maps using two different sets of subjects from the same dataset ("population"). Each row corresponds to a non-thresholded spatial map that has been converted to zero mean and unit norm. We define the cross correlation of the spatial maps to be the matrix $C = A_1 A_2^T$. Each element C_{ij} is the Pearson correlation coefficient of the i-th row of A_1 and the j-th row of A_2; hence, $C_{ij} = 1$ if and only if the two components are identical. As in Varoquaux et al. (2010), we use the the parameter to estimate the overlap of the two subspaces spanned by the rows of A_1 and A_2. The value of p quantifies how similar the two subspaces are, although the individual spatial maps may differ significantly even when the two subspaces are similar.

$$p = \frac{1}{d} tr(CC^T) \tag{14}$$

Another measure corresponds to the one-to-one matching of the components. Let A_1, A_2 be two matrices such that each row is a thresholded or non-thresholded spatial component with zero mean and unit norm, and $\tilde{A}_1$, $\tilde{A}_2$ be respectively row permutations of A_1 and A_2. We define the average one-to-one matching score to be corresponding to all possible row permutations of the two matrices. In this case, we are matching the rows of A_1 and the rows of A_2 such that the overall correlation score is maximized.

$$q = \max\left(\frac{1}{d} tr\left(\tilde{A}_1 \tilde{A}_2^T\right)\right) \tag{15}$$

To evaluate our model, we evaluate the p and q measures for several scenarios. Let S be the complete dataset with N subjects, and let S_1, S_2 be two sets of subjects selected at random from the complete dataset. We run our algorithm and both MELODIC and GIFT on the two groups S_1 and S_2. We then compute the p scores on the non-thresholded spatial components and the q scores on both thresholded and non-thresholded spatial components. The two groups are selected as follows.

1. S_1, S_2 are two subgroups randomly selected from S each with $\{N/5, 3N/10, 2N/5, N/2\}$ subjects.
2. S_1 is a subgroup randomly selected from S with $\{N/5, 3N/10, 2N/5, N/2\}$ subjects, $S_2 = S$.

Both experiments are repeated with random selection of subgroups and the averaged scores are reported. Test results from 1) illustrate how reproducible the model is, with varying number of available subjects. Results from 2) illustrate the consistency between a small subgroup and the whole group.

fMRI datasets

We use two fMRI datasets to evaluate the methods described above: a resting state fMRI dataset and a working memory task fMRI dataset.

Resting state fMRI dataset

The resting state fMRI dataset we use is the 100 healthy unrelated subject dataset from the WU-Minn Human Connectome Project Van Essen et al. (2013). Each subject was involved in four 15-minute runs with TR = 0.72 s totaling 1200 frames per run. All frames were sampled into 91,282 grayordinates (a combination of cortex vertices and subcortical voxels). The data was then temporally preprocessed and de-noised using the FIX approach Griffanti et al. (2014); Salimi-Khorshidi et al. (2014). The resulting images were then aligned using MSM registration Robinson et al. (2014). Full details of this dataset can be found in publications from the project Van Essen et al. (2013); Smith et al. (2013). Each voxel's time series was normalized for each run. Note that, out of the 91,282 grayordinates in the resting state fMRI data, we employed cortical data only (59,412 vertices of the left and right hemispheres).

Working memory task fMRI dataset

We employed the working memory task dataset of the 100 healthy unrelated subjects from the WU-Minn Human Connectome Project Van Essen et al. (2013). Participants were presented with blocks of trials that consisted of pictures of places, tools, faces and body parts. Within each run, the 4 different stimulus types were presented in separate blocks such that half the blocks used a 2-back memory task and half used a 0-back memory task. Individual subject data includes 2 runs of 405 frames each, preprocessed using FEAT Woolrich et al. (2001). Full details of this dataset can be found in Van Essen et al. (2013); Smith et al. (2013). Here, we analyzed only data from 2-back memory blocks as follows. To account for head motion, motion parameters were regressed out of the time series with AFNI's Cox (1996) 3dDeconvolve, which was also used to removed slow baseline drifts. Subsequently, time series data were z scored. To account for the cue stimulus indicating block type at the beginning of each block, data corresponding to the first 12 s of each block were removed. The remaining data segment of the block, plus 2 s due to hemodynamic lag, were considered to reflect working memory related data. In this way, 25 frames for each block were used, totaling 200 frames once the data were concatenated. Each voxel's time series was normalized for each run.

Model parameters choice

Choosing an appropriate model order d (the number of non-constant eigenvectors; see equation (11)) depends on the data. Since model order selection is not the main focus of this paper, we use an empirical value. In Smith et al. (2013) a model order of 30 was used to conduct group analysis on both cortical and subcortical data; 23 of the extracted spatial maps were suggested to be meaningful based on visual inspection of the maps. As we are only using cortical data, a model order of $d = 20$ was employed here.

The determination of the number k of nearest neighbors used in constructing the similarity graph (see equation (9)) should also be based on the data. The value of k needs to be large enough to include the vertices that are closely correlated with a reference voxel, but not so large as to include too many weak correlations in which case the model becomes noisier. The nearest neighbors of a voxel v are those voxels that are strongly correlated with it. To select a consistent threshold cut-off, we Z-scored the correlations and utilized a threshold of $Z > 2$. Specifically, for each voxel, its correlation values with all other voxels (corresponding to a single row or column of the group-level correlation matrix) were converted to Z-scores and thresholded. This procedure selects the neighbors that are the most functionally coupled to the voxel in question. In this manner, we determine the number of nonzero values obtained after thresholding for each voxel, and average this count across all voxels. This average value is then taken as the value of k.

Results

Resting state dataset

Group-level spatial maps

As described in section 2.6, an important parameter that is estimated from the data by the model is the number k of nearest neighbors used in constructing the similarity graph. Based on the resting state dataset, $k = 1742$. Fig. 2 shows the 20 group-level thresholded spatial maps extracted by LEICA on the resting state dataset. The extracted independent spatial components were converted to Z scores and thresholded at $|Z| > 2$.

Similarity score

We first visually compared the spatial maps from LEICA and MELODIC. Subsequently, we computed the similarity scores of pairs putatively capturing the same brain networks. Out of the 20 LEICA spatial maps, 18 of them exhibited good matches with the corresponding MELODIC spatial maps. The cosine similarity scores are presented in Table 1.

In Fig. 3 we show sample pairs with varying similarity scores. In all cases, the pairs are visually very similar, and even the least similar pair ($c = 0.53$, Fig. 3(f)), shows good correspondence. Furthermore, LEICA spatial maps are qualitatively smoother than those produced by MELODIC.

Two LEICA components (1 and 14) did not match components generated by MELODIC. Nevertheless, the two maps appear to be related to meaningful functional networks. We tentatively identify spatial map 1 to be related to the auditory network, and spatial map 14 to be related to the fronto-parietal network important for attention and executive function.

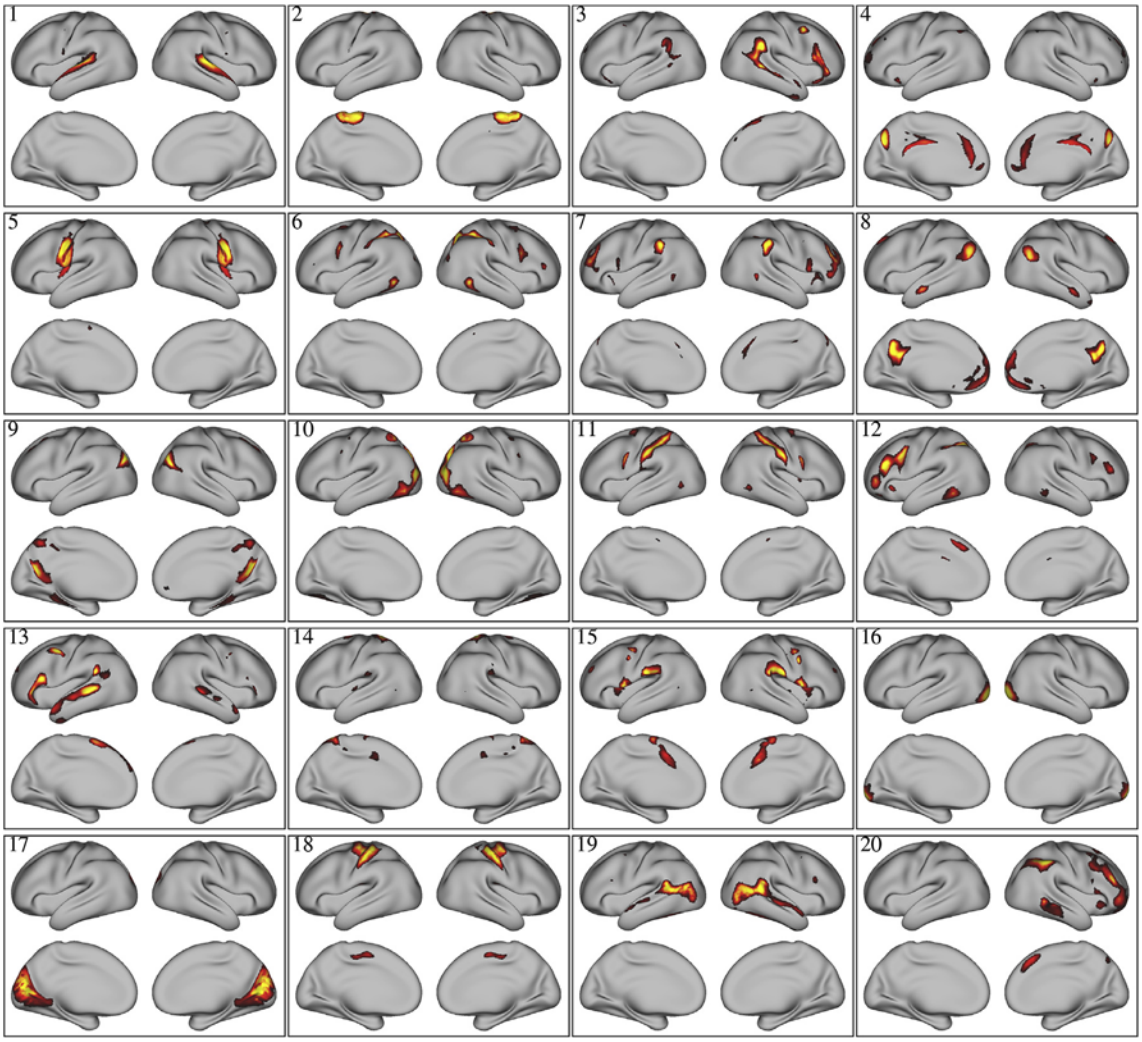


Fig. 2. Twenty group spatial maps generated by LEICA on the 100-subject HCP preprocessed resting state fMRI dataset. Spatial maps were thresholded at $|Z| > 2$.

Table 1
Cosine similarity scores for MELODIC and LEICA spatial maps (SM) for the resting state dataset. The numbering of the spatial maps is the same as in Fig. 2.

SM #	c
16	0.96
5	0.90
10	0.87
18	0.83
17	0.80
8	0.78
2	0.77
13	0.72
15	0.72
4	0.71
9	0.71
20	0.68
19	0.61
7	0.60
3	0.59
12	0.57
11	0.56
6	0.53

Next, we describe a quantitative comparison between LEICA and both MELODIC and GIFT.

Recovery of functional connectivity structure

We start with an example to illustrate the validation strategy described in section 2.4.2. In Fig. 4 we show two spatial maps (left in each frame) and the corresponding correlation maps (right). In (a), to generate an artificial map, a sphere with a geodesic radius of 40 mm was generated around a vertex near the motor network and weighted by a Gaussian kernel. In (b), the spatial map corresponded to a motor-related network generated by LEICA. In both cases, correlation maps were created by computing the weighted sum of the correlation maps for each individual vertex (in the spatial maps), as described in section 2.4.2. When a spatial map does not closely correspond to a meaningful intrinsic network, such as the case in (a), the similarity score was low ($c = 0.42$), and its correlation map is relatively poor (in (a), compare the two left-most and the rightmost surfaces). When a spatial map better captures an intrinsic network, such as the case in (b), both images matched well and the cosine similarity score was high ($c = 0.84$).

To evaluate the model, we followed the approach above and compared the individual spatial maps generated by LEICA, MELODIC, and GIFT with the corresponding correlation maps (Table 2), just as done in Fig. 4 (b). LEICA exhibited cosine similarity scores (between component maps and correlation maps) that were similar or higher than the two other algorithms, indicating that the method uncovered meaningful component maps that represent intrinsic functional networks.

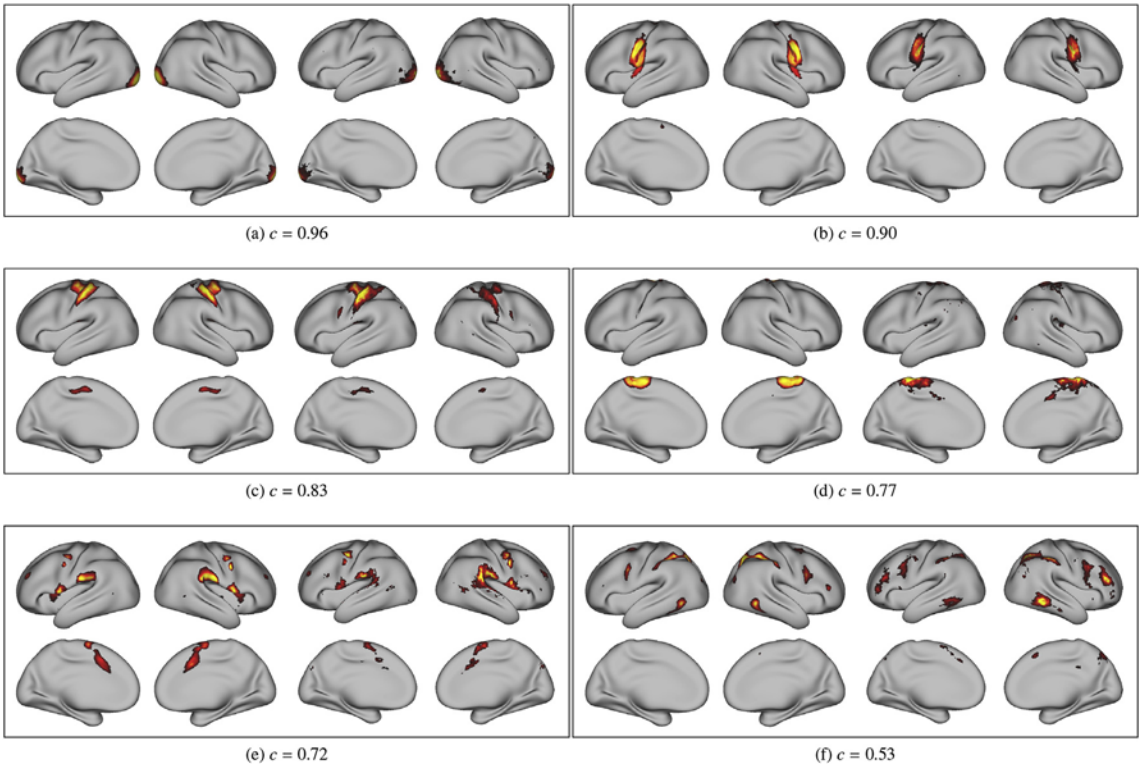


Fig. 3. Six sample pairs of spatial maps with different cosine similarity scores c. In each frame, the spatial maps on the left are generated by LEICA and those on the right are generated by MELODIC. Both models appear to capture the same components (even in the case with least similarity, $c = 0.53$). The components from our model appear to be smoother than those by MELODIC.

Table 2
Cosine similarity between 20 spatial maps generated by LEICA, MELODIC, and GIFT and their correlation maps for the resting state dataset. The numbering of spatial maps (SM #) applies to LEICA; the values for MELODIC and GIFT are sorted in decreasing order.

SM #	LEICA	MELODIC	GIFT
16	0.96	0.95	0.94
5	0.93	0.93	0.93
18	0.84	0.86	0.86
11	0.82	0.83	0.82
17	0.81	0.79	0.81
6	0.79	0.79	0.81
9	0.79	0.78	0.79
10	0.79	0.78	0.78
13	0.78	0.78	0.78
2	0.78	0.78	0.77
8	0.77	0.77	0.76
7	0.76	0.77	0.76
12	0.74	0.76	0.75
3	0.74	0.74	0.73
14	0.73	0.70	0.67
20	0.73	0.61	0.62
1	0.73	0.54	0.56
19	0.67	0.51	0.50
4	0.66	0.44	0.48
15	0.66	0.18	0.43
average	**0.77**	0.71	0.73

Reproducibility results

We first describe how the results for GIFT and MELODIC were generated. For GIFT, the subject-level dimensionality was first reduced to $m = 150$ using PCA, followed by the temporal concatenation across all reduced data. As stated in Calhoun et al. (2001); Smith et al. (2014) the subject-specific PCA should not over-reduce the data ($m > d$); usually the larger the m the better, and we chose m to be large for best performance. PCA was then applied on the concatenated data to reduce the dimensionality to $d = 20$, followed by ICA to extract the independent components. For MELODIC, the temporal concatenation PCA was implemented using the MIGP algorithm to compute group-level principal components with dimension of 4500, as used in the original paper Smith et al. (2014) for the same dataset. The MELODIC toolbox was then used to extract 20 independent components.

The reproducibility results generated by cross validating a pair of subgroups and by cross validating one subgroup relative to the complete group are shown in Table 3. Overall LEICA achieved the most reproducible and stable results among the models. From Table 3(a), we see that given any pair of distinct subgroups, LEICA was able to extract a set of group spatial features that was most stable most often. These results are important because they indicate that the inter-subject variability was modeled better using Laplacian eigenmaps than PCA. In Table 3(b), the values from LEICA are also significantly higher, which indicates that LEICA was able to extract more consistent group spatial features when fewer subjects are used.

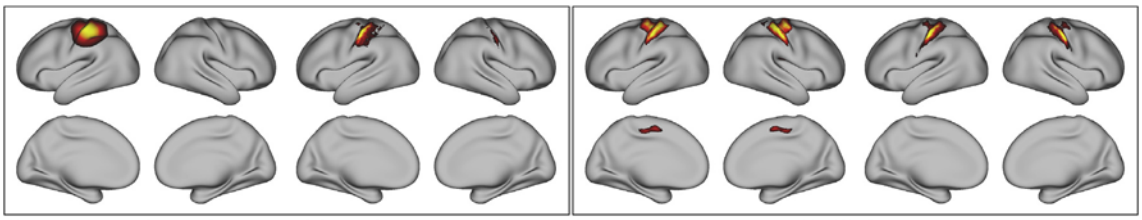
Working memory dataset

Group-level spatial maps

The method estimated the number of nearest neighbors used in constructing the similarity graph to be $k = 938$ for the working memory dataset. Fig. 5 shows the 20 group-level independent thresholded spatial maps extracted by LEICA on the 2-back working memory fMRI dataset. The extracted spatial components were also converted to Z-scores and thresholded at $|Z| > 2$.

Similarity score

Out of the 20 LEICA spatial maps, 18 exhibited relatively good matches with those from MELODIC (cosine similarity scores are



(a) An artificial spatial map (left) generated by drawing a circle around a vertex (b) The motor network generated by LEICA (left) and its corresponding correlation map (right). Cosine similarity score: $c = 0.42$.

near the left motor network (with Gaussian weights) and its corresponding corre- tion map (right). Cosine similarity score: $c = 0.84$.

Fig. 4. Two spatial maps and their corresponding correlation maps. The correlation maps were created by computing the weighted sum of the correlation maps for each individual vertex as described in section 2.4.2.

Table 3
Reproducibility results based on pairs of subgroups each with 20, 30, 40, 50 subjects for the resting state dataset. p quantifies the similarity of the two subspaces and q measures the one-to-one matching similarity.

(a) Reproducibility results based on pairs of subgroups

Resting State	GIFT	MELODIC	LEICA
20 subjects			
p	0.82	0.80	**0.88**
q - non-thresholded	0.79	0.83	**0.83**
q - thresholded	0.48	**0.82**	0.77
30 subjects			
p	0.84	0.85	**0.91**
q - non-thresholded	0.81	0.83	**0.85**
q - thresholded	0.60	**0.84**	0.80
40 subjects			
p	0.89	0.87	**0.92**
q - non-thresholded	0.82	0.84	**0.86**
q - thresholded	0.76	**0.84**	0.81
50 subjects			
p	0.91	0.89	**0.93**
q - non-thresholded	0.87	0.85	**0.89**
q - thresholded	0.82	0.85	**0.85**

(b) Reproducibility results of one subgroup relative to the whole group

Resting State	GIFT	MELODIC	LEICA
20 subjects			
p	0.73	0.71	**0.92**
q - non-thresholded	0.73	0.68	**0.86**
q - thresholded	0.46	0.65	**0.82**
30 subjects			
p	0.75	0.74	**0.94**
q - non-thresholded	0.74	0.69	**0.90**
q - thresholded	0.52	0.68	**0.86**
40 subjects			
p	0.79	0.76	**0.95**
q - non-thresholded	0.75	0.72	**0.90**
q - thresholded	0.61	0.70	**0.87**
50 subjects			
p	0.81	0.77	**0.96**
q - non-thresholded	0.79	0.74	**0.91**
q - thresholded	0.68	0.72	**0.87**

presented in Table 4). Thus, the two models captured similar sets of components.

Recovery of functional connectivity structure

As in the case of the resting state dataset, we compared spatial maps obtained by LEICA, MELODIC, and GIFT with the corresponding correlation maps (Table 5). For the task dataset, LEICA exhibited cosine similarity scores (between component maps and correlation maps) that were consistently higher than the two other algorithms, indicating that the method better uncovered functional networks during working memory performance.

Reproducibility results

Along the same lines of section 3.1.4, we chose the subject-level PCA dimension $m = 90$ for GIFT, and a dimension of 200 for MIGP in MELODIC. The reproducibility results for the three models are shown in Table 6. Compared with the results for resting state (Table 3), all reproducibility scores dropped considerably, reflecting the rather limited amount of task data (200 frames per subject vs. 4800 frames per subject). However, the performance of LEICA did not decline as much as observed for the other two models, highlighting the advantage of the method when more limited data are available, as is typical in most fMRI research.

Discussion

In the present paper, we propose a nonlinear ICA-based model for extracting group-level spatial maps from multi-subject fMRI datasets. There are two key elements to our method. The first element is the construction of a k-nearest neighbor graph based on the group-average correlation matrix, which makes the final results less noisy and captures intrinsic functional connectivity. The second element is the use of a nonlinear dimensionality reduction stage based on Laplacian eigenmaps, which seeks to identify a manifold subspace common to the group. The mapping is such that it preserves the correlation structure among the voxels as much as possible, which we propose is important to capture functional brain networks during resting and task states.

Model flexibility

LEICA has more parameters than traditional PCA-ICA models. In particular, the construction of the similarity graph requires the specification of the edge weights and the number k of nearest neighbors. A number of techniques exist for specifying edge weights, including Pearson's correlation coefficient, the heat kernel, and binarized weights. As noted before, we believe that the thresholded Pearson's correlation coefficient is appropriate for our application. The number k of nearest neighbors needs to be determined, as we illustrated earlier. Also, the final similarity graph can be fine tuned. In this paper, we have provided practical and simple ways to select the parameters that are particular to our model. These methods seem to work well in practice and may offer more flexibility than found in other models.

Number of nearest neighbors k

LEICA provides a principled method for determining the value of k based on z scoring, which produces good results on both the resting state and working memory datasets. However, we note that the model is not very sensitive to the exact value of k. Fig. 6 shows motor-related networks extracted by the method applied to resting state data with different values of k. When k is rather small, say $k = 50$ or even $k = 200$, the resulting motor networks do not represent the true network well. For $k = 1000$, the motor networks are accurately extracted but the left and right hemispheres are separated. When $k = 2500$, the results are nearly identical to what was extracted automatically (Fig. 2). For $k = 5000$,

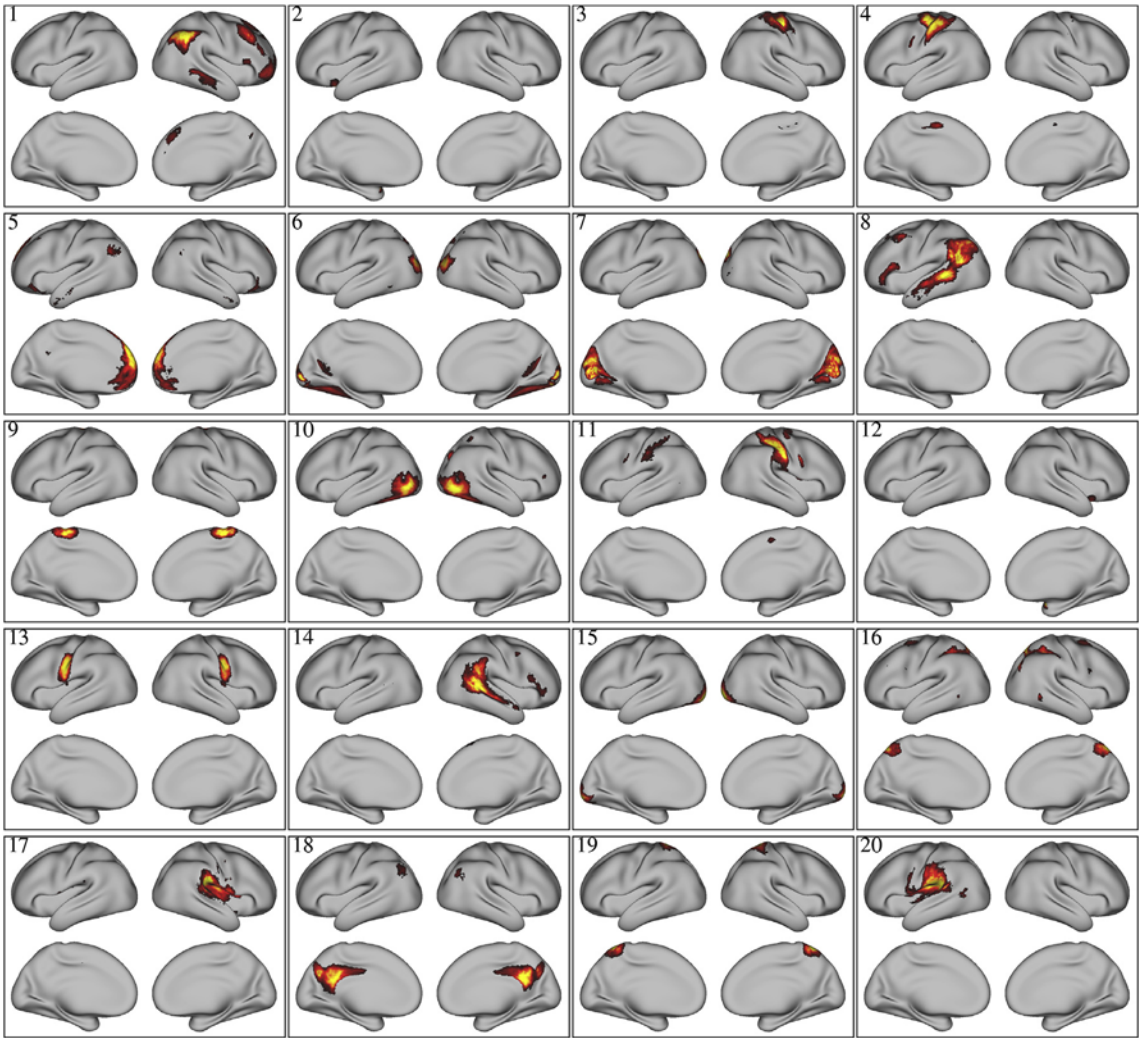


Fig. 5. Twenty group spatial maps generated by LEICA on the 100-subject HCP preprocessed working memory task dataset. Spatial maps were thresholded at $|Z| > 2$.

Table 4
Cosine similarity scores for spatial maps (SM) from MELODIC and LEICA on the working memory dataset. The numbering of the spatial maps is the same as in Fig. 5.

SM #	*c*
13	0.97
7	0.95
10	0.93
15	0.86
4	0.81
1	0.77
9	0.76
6	0.75
18	0.71
8	0.67
11	0.63
20	0.58
16	0.56
19	0.44
5	0.41
14	0.40
17	0.38
3	0.35

which we consider as "large" (nearly three times the value of $k = 1742$ obtained by the method), the motor network becomes somewhat noisy

Table 5
Cosine similarity between 20 spatial maps and their correlation maps for LEICA, MELODIC, GIFT on the working memory dataset. The numbering of spatial maps (SM #) only applies to LEICA, and the values for MELODIC and GIFT are sorted in decreasing order.

SM #	LEICA	MELODIC	GIFT
2	0.99	0.97	0.99
12	0.99	0.84	0.91
13	0.97	0.80	0.84
9	0.96	0.78	0.81
3	0.93	0.77	0.80
4	0.93	0.74	0.78
11	0.85	0.74	0.77
19	0.85	0.73	0.74
17	0.79	0.72	0.73
15	0.76	0.70	0.71
6	0.74	0.70	0.69
18	0.72	0.64	0.68
5	0.70	0.62	0.67
16	0.69	0.59	0.63
10	0.69	0.55	0.61
20	0.68	0.51	0.60
14	0.68	0.45	0.53
1	0.67	0.44	0.47
8	0.65	0.40	0.34
7	0.61	0.26	0.29
average	**0.79**	0.65	0.68

Table 6
Reproducibility results based on subgroups each with 20, 30, 40, 50 subjects on the working memory dataset.

(a) Reproducibility results based on pairs of subgroups			
Working Memory	GIFT	MELODIC	LEICA
20 subjects			
p	0.35	0.36	**0.68**
q - non-thresholded	0.41	0.41	**0.70**
q - thresholded	0.27	0.41	**0.68**
30 subjects			
p	0.47	0.45	**0.76**
q - non-thresholded	0.51	0.49	**0.77**
q - thresholded	0.39	0.49	**0.74**
40 subjects			
p	0.51	0.52	**0.81**
q - non-thresholded	0.55	0.55	**0.81**
q - thresholded	0.49	0.54	**0.78**
50 subjects			
p	0.56	0.55	**0.84**
q - non-thresholded	0.59	0.57	**0.84**
q - thresholded	0.55	0.56	**0.80**
(b) Reproducibility results of one subgroup relative to the whole group			
Working Memory	GIFT	MELODIC	LEICA
20 subjects			
p	0.49	0.52	**0.73**
q - non-thresholded	0.53	0.52	**0.72**
q - thresholded	0.33	0.50	**0.69**
30 subjects			
p	0.60	0.58	**0.79**
q - non-thresholded	0.63	0.57	**0.79**
q - thresholded	0.42	0.55	**0.76**
40 subjects			
p	0.66	0.65	**0.83**
q - non-thresholded	0.60	0.63	**0.82**
q - thresholded	0.52	0.60	**0.79**
50 subjects			
p	0.69	0.67	**0.86**
q - non-thresholded	0.64	0.65	**0.84**
q - thresholded	0.61	0.62	**0.81**

but the result is still quite reasonable. We note that for the range $k \in [1200, 2500]$ there seems to be no appreciable differences in the results obtained by the method.

Model order d

The general problem of estimating the model order (here, the number of components) amounts to estimating the intrinsic dimensionality of a nonlinear manifold from a given data sample, a notoriously difficult problem Camastra and Staiano (2016). Moreover, in the context of our application, we note that it is not clear that there is a single "correct" value that is needed. A technique used to estimate the number of components for fMRI data is based on Akaikes information criterion (AIC) and the minimum description length (MDL), both of which make significant technical assumptions about the model Wax and Kailath (1985). For fMRI datasets, MELODIC and GIFT use this technique for their model parameter estimation but the accuracy is not guaranteed Varoquaux et al. (2010). Another technique is based on a statistical shape model Varoquaux et al. (2010). In the present investigation, we decided against using such techniques due to their considerable assumptions. Previous empirical results for fMRI data indicate that the appropriate model parameter lies between 20 and 100, depending on the dataset. Here, we found that the model order of 20 was adequate for the two datasets investigated, as shown by the high quality of the spatial components generated, but we note that the exploration of methods to estimate model order was beyond the scope of the paper.

Computational complexity and execution time

When the number of subjects is very large, it becomes computationally difficult to concatenate the data across all subjects along the temporal dimension to perform the typical estimation procedures (such as PCA) due to memory limitations. This issue is handled in GIFT by using successive PCAs to reduce the dimensionality of each individual's data before concatenation across subjects, which is then followed by running PCA again on the concatenated data. In MELODIC, a recently proposed algorithm called MIGP is used to compute an approximation of PCA on the temporally concatenated data. In contrast, LEICA computes the correlation matrix of each individual subject, followed by averaging these correlation matrices over the subjects. We then build a k-nearest neighbor graph based on the resulting correlation matrix. This is followed by computing the Laplacian eigenmaps on the sparse correlation matrix, which includes a sparse eigenvector decomposition that has a fast implementation, and can be easily boosted by GPU platforms as shown in Jin and Jaja (2016).

The execution times of MELODIC, GIFT, and our model on the 100 subject HCP resting state and working memory datasets are shown in Table 7. All programs were run on the same machine with Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz CPU. Overall, LEICA ran much faster than both algorithms on the resting state dataset, and slower than MELODIC and faster than GIFT on the working memory dataset. This can be explained by the fact that the execution time of MELODIC is dominated by the MIGP step, which is quadratic in terms of the length T of the time series, while the execution time of LEICA is linear in T ($T = 4800$ for the resting state dataset and $T = 200$ for the working memory dataset). The MIGP step of MELODIC was implemented in MATLAB (we used the scripts provided by HCP) and the ICA step was from the MELODIC toolbox. Both GIFT and our model were implemented in Python (as the original GIFT toolbox is not compatible with CIFTI data format used by HCP, we implemented the model in Python).

Conclusion

In this paper, we present a novel model – called LEICA – involving a non-linear dimensionality reduction procedure followed by ICA to

Table 7
Execution times for MELODIC, GIFT, and LEICA on the 100-subject resting state and working memory datasets from the Human Connectome Project.

	MELODIC	GIFT	LEICA
resting state	2d 5h 17m	8h 42m	5h 33m
working memory	8m 47s	20m 15s	15m 29s

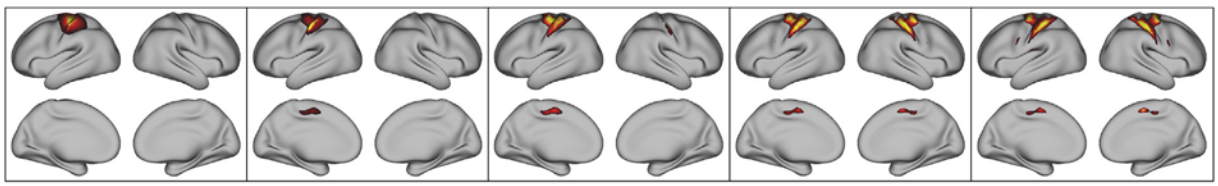


Fig. 6. The motor networks extracted by LEICA on the resting state dataset with different values of k.

identify group-level spatial features from multi-subject fMRI datasets, and applied it to resting state and working memory data. The spatial maps extracted by our model were shown to be related to meaningful functional networks, and are comparable or better than those generated by one of the current state-of-the-art models (MELODIC). Evaluation of the functional connectivity structure of spatial component maps revealed that LEICA detects functionally cohesive maps. Moreover, LEICA spatial maps are at least as functionally cohesive or better than those detected by MELODIC and GIFT for resting state data, and more functionally cohesive than by MELODIC and GIFT for working memory data. Tests of reproducibility showed that LEICA is at least as stable as MELODIC and GIFT when ample data are available (resting state dataset), and more stable than these methods when more limited data are available (working memory dataset). Finally, our method is computationally efficient. Our software is available at https://github.com/liuchihuang/LEICA.

Conflicts of interest

None.

Acknowledgment

The authors gratefully acknowledge the support provided by The University of Maryland/MPowering the State through the Center for Health-related Informatics and Bioimaging (CHIB) and by the NSF MRI Grant Number: CNS1429404. L.P. was supported by grants from the National Institute of Mental Health (MH071589 and MH112517). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience (U54MH091657) at Washington University.

References

Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imag. 23 (2), 137–152.
Beckmann, C.F., Smith, S.M., 2005. Tensorial extensions of independent component analysis for multisubject fmri analysis. Neuroimage 25 (1), 294–311.
Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. 15 (6), 1373–1396.
Birn, R.M., Saad, Z.S., Bandettini, P.A., 2001. Spatial heterogeneity of the nonlinear dynamics in the fmri bold response. Neuroimage 14 (4), 817–826.
Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magn. Reson. Med. 34 (4), 537–541.
Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J., 2001. A method for making group inferences from functional mri data using independent component analysis. Hum. Brain Mapp. 14 (3), 140–151.
Camastra, F., Staiano, A., 2016. Intrinsic dimension estimation: advances and open problems. Inf. Sci. 328, 26–41.
Cole, M.W., Bassett, D.S., Power, J.D., Braver, T.S., Petersen, S.E., 2014. Intrinsic and task-evoked network architectures of the human brain. Neuron 83 (1), 238–251.
Cox, R.W., 1996. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29 (3), 162–173.
Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fmri atlas generated via spatially constrained spectral clustering. Hum. Brain Mapp. 33 (8), 1914–1928.
Damoiseaux, J., Beckmann, C., Arigita, E.S., Barkhof, F., Scheltens, P., Stam, C., Smith, S., Rombouts, S., 2008. Reduced resting-state brain activity in the "default network" in normal aging. Cerebr. Cortex 18 (8), 1856–1864.
Di Martino, A., Scheres, A., Margulies, D.S., Kelly, A., Uddin, L.Q., Shehzad, Z., Biswal, B., Walters, J.R., Castellanos, F.X., Milham, M.P., 2008. Functional connectivity of human striatum: a resting state fmri study. Cerebr. Cortex 18 (12), 2735–2747.
Erhardt, E.B., Rachakonda, S., Bedrick, E.J., Allen, E.A., Adali, T., Calhoun, V.D., 2011. Comparison of multi-subject ica methods for analysis of fmri data. Hum. Brain Mapp. 32 (12), 2075–2095.
Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. Unit. States Am. 102 (27), 9673–9678.
Garrity, A.G., Pearlson, G.D., McKiernan, K., Lloyd, D., Kiehl, K.A., Calhoun, V.D., 2007. Aberrant "default mode" functional connectivity in schizophrenia. Am. J. Psychiatr. 164 (3), 450–457.
Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. Unit. States Am. 100 (1), 253–258.
Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. Proc. Natl. Acad. Sci. Unit. States Am. 101 (13), 4637–4642.
Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., et al., 2014. Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. Neuroimage 95, 232–247.
Guo, Y., Pagnoni, G., 2008. A unified framework for group independent component analysis for multi-subject fmri data. Neuroimage 42 (3), 1078–1093.
Haak, K.V., Marquand, A.F., Beckmann, C.F., 2016. Connectopic mapping with resting-state fmri. Neuroimage arXiv preprint arXiv:1602.07100.
Jafri, M.J., Pearlson, G.D., Stevens, M., Calhoun, V.D., 2008. A method for functional network connectivity among spatially independent resting-state components in schizophrenia. Neuroimage 39 (4), 1666–1681.
Jiang, T., He, Y., Zang, Y., Weng, X., 2004. Modulation of functional connectivity during the resting state and the motor task. Hum. Brain Mapp. 22 (1), 63–71.
Jin, Y., Jaja, J.F., 2016. A high performance implementation of spectral clustering on cpu-gpu platforms. In: Parallel and Distributed Processing Symposium Workshops, pp. 825–834, 2016 IEEE International. IEEE.
Jolliffe, I., 2002. Principal Component Analysis. Wiley Online Library.
Maaten, L. v. d., Hinton, G., 2008. Visualizing data using t-sne. J. Mach. Learn. Res. 9 (Nov), 2579–2605.
Mannfolk, P., Wirestam, R., Nilsson, M., Ståhlberg, F., Olsrud, J., 2010. Dimensionality reduction of fmri time series data using locally linear embedding. Magn. Reson. Mater. Phys. Biol. Med. 23 (5–6), 327–338.
Marquand, A.F., Haak, K.V., Beckmann, C.F., 2017. Functional corticostriatal connection topographies predict goal directed behaviour in humans. bioRxiv, 169151.
McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.-P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fmri data by blind separation into independent spatial components. Hum. Brain Mapp. 6, 160–188.
Miller, K.L., Luh, W.-M., Liu, T.T., Martinez, A., Obata, T., Wong, E.C., Frank, L.R., Buxton, R.B., 2001. Nonlinear temporal dynamics of the cerebral blood flow response. Hum. Brain Mapp. 13 (1), 1–12.
Robinson, E.C., Jbabdi, S., Glasser, M.F., Andersson, J., Burgess, G.C., Harms, M.P., Smith, S.M., Van Essen, D.C., Jenkinson, M., 2014. Msm: a new flexible framework for multimodal surface matching. Neuroimage 100, 414–426.
Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290 (5500), 2323–2326.
Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468.
Schölkopf, B., Smola, A., Müller, K.-R., 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. 10 (5), 1299–1319.
Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., et al., 2013. Resting-state fmri in the human connectome project. Neuroimage 80, 144–168.
Smith, S.M., Hyvärinen, A., Varoquaux, G., Miller, K.L., Beckmann, C.F., 2014. Group-pca for very large fmri datasets. Neuroimage 101, 738–749.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al., 2004. Advances in functional and structural mr image analysis and implementation as fsl. Neuroimage 23, S208–S219.
Tenenbaum, J.B., De Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290 (5500), 2319–2323.
Thirion, B., Faugeras, O., 2004. Nonlinear dimension reduction of fmri data: the laplacian embedding approach. In: Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on. IEEE, pp. 372–375.
Van Der Maaten, L., Postma, E., Van den Herik, J., 2009. Dimensionality reduction: a comparative review. J. Mach. Learn. Res. 10, 66–71.
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.-M.H., et al., 2013. The wu-minn human connectome project: an overview. Neuroimage 80, 62–79.
Varoquaux, G., Sadaghiani, S., Pinel, P., Kleinschmidt, A., Poline, J.-B., Thirion, B., 2010. A group model for stable multi-subject ica on fmri datasets. Neuroimage 51 (1), 288–299.
Wang, Y.M., Schultz, R.T., Constable, R.T., Staib, L.H., 2003. Nonlinear estimation and modeling of fmri data using spatio-temporal support vector regression. In: Biennial International Conference on Information Processing in Medical Imaging. Springer, pp. 647–659.
Wax, M., Kailath, T., 1985. Detection of signals by information theoretic criteria. IEEE Trans. Acoust. Speech Signal Process. 33 (2), 387–392.
Weinberger, K.Q., Saul, L.K., 2006. An introduction to nonlinear dimensionality reduction by maximum variance unfolding. In: AAAI, vol. 6, pp. 1683–1686.
Weissman-Fogel, I., Moayedi, M., Taylor, K.S., Pope, G., Davis, K.D., 2010. Cognitive and default-mode resting state networks: do male and female brains "rest" differently? Hum. Brain Mapp. 31 (11), 1713–1726.
Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of fmri data. Neuroimage 14 (6), 1370–1386.
Xie, X., Cao, Z., Weng, X., 2008. Spatiotemporal nonlinearity in resting-state fmri of the human brain. Neuroimage 40 (4), 1672–1685.
Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106 (3), 1125–1165.