Published in final edited form as:
Biometrics. 2013 December ; 69(4): 970–981. doi:10.1111/biom.12068.

A hierarchical model for probabilistic independent component analysis of multi-subject fMRI studies

Ying Guo[1] and **Li Tang**[2]

Ying Guo: yguo2@emory.edu

[1]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia 30322, USA

[2]Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

Summary

An important goal in fMRI studies is to decompose the observed series of brain images to identify and characterize underlying brain functional networks. Independent component analysis (ICA) has been shown to be a powerful computational tool for this purpose. Classic ICA has been successfully applied to single-subject fMRI data. The extension of ICA to group inferences in neuroimaging studies, however, is challenging due to the unavailability of a pre-specified group design matrix. Existing group ICA methods generally concatenate observed fMRI data across subjects on the temporal domain and then decompose multi-subject data in a similar manner to single-subject ICA. The major limitation of existing methods is that they ignore between-subject variability in spatial distributions of brain functional networks in group ICA. In this paper, we propose a new hierarchical probabilistic group ICA method to formally model subject-specific effects in both temporal and spatial domains when decomposing multi-subject fMRI data. The proposed method provides model-based estimation of brain functional networks at both the population and subject level. An important advantage of the hierarchical model is that it provides a formal statistical framework to investigate similarities and differences in brain functional networks across subjects, e.g., subjects with mental disorders or neurodegenerative diseases such as Parkinson's as compared to normal subjects. We develop an EM algorithm for model estimation where both the E-step and M-step have explicit forms. We compare the performance of the proposed hierarchical model with that of two popular group ICA methods via simulation studies. We illustrate our method with application to an fMRI study of Zen meditation.

Keywords

Independent component analysis; Multi-subject imaging data; Functional magnetic resonance imaging (fMRI); Hierarchical model; Group inferences; Maximum likelihood estimation; EM algorithm

Correspondence to: Ying Guo, yguo2@emory.edu.

Supplementary Materials

Web Appendices, Tables and Figures referenced in Sections 2, 3 and 4 are available with this paper at the Biometrics website on Wiley Online Library.

1. Introduction

Functional magnetic resonance imaging (fMRI) has become an increasingly important tool to investigate neural processing patterns in brain functioning. In fMRI studies, subjects' neural activity is captured by a series of three-D brain images obtained with an MRI scanner while the subjects are exposed to different experimental conditions. The observed fMRI blood-oxygen-level-dependent (BOLD) effects represent the combination of source signals generated by various underlying brain functional networks. A common objective in fMRI analysis is to separate these networks and characterize both their distributed patterns and temporal dynamics. Independent component analysis (ICA) has become an important computational tool for this purpose in neuorimaging studies. As a special case of blind source separation, ICA aims to recover independent sources given only sensor observations that are unknown linear mixtures of the unobserved independent source signals. The two key assumptions of ICA are that the source signals are statistically independent and follow non-Gaussian distributions. As a fully data-driven approach, ICA has important applications in identifying brain functional networks in fMRI studies where there is little prior information regarding the temporal and spatial characteristics of these networks.

The classic ICA is readily applicable to single-subject fMRI analysis (McKeown et al., 1998; Beckmann and Smith, 2004). Let $\boldsymbol{Y}$ be the $T \times V$ fMRI data observed from a subject, where T is the number of time points and V is the total number of voxels in a three-D brain image acquired at each time point. Here, the two-D matrix $\boldsymbol{Y}$ is constructed from the observed time series of three-D images by concatenating each image into a row of $\boldsymbol{Y}$. Classic noise-free ICA of single-subject fMRI (McKeown et al., 1998) can be expressed as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{S}, \quad (1)$$

where $\boldsymbol{S}$ is a $q \times V$ matrix with each row representing the concatenated spatial map of a non-Gaussian source signal and q denoting the number of source signals, $\boldsymbol{A}$ is a $T \times q$ matrix that mixes the independent q source signals to generate the observed fMRI images with each column of $\boldsymbol{A}$ representing the time course of a source signal. The q extracted source signals are assumed to be independent in the spatial domain and hence are called independent components (ICs).

Most fMRI studies involve multiple subjects. The generalization of the ICA for group inferences is more challenging than that of other statistical models such as the general linear model (GLM). The major difficulties are that ICA does not have a pre-specified design matrix, both time courses and spatial maps need to be estimated from the data, and there is considerable between-subject variability in fMRI signals. Among the existing group ICA approaches, the most commonly used method is temporal-concatenation group ICA (TC-GICA). In TC-GICA, the observed fMRI BOLD data are stacked across subjects on the temporal domain and an ICA decomposition is then performed on the group data. Specifically, let $\boldsymbol{Y}_i$ be a $T \times V$ matrix representing the observed fMRI data from subject i. TC-GICA temporally concatenates data across subjects and then decomposes the $TN \times V$ group data as,

$$(\boldsymbol{Y}_1', \ldots, \boldsymbol{Y}_N')' = \boldsymbol{M}\boldsymbol{S}, \quad (2)$$

where $\boldsymbol{M}$ is a $TN \times q$ group mixing matrix and $\boldsymbol{S}$ is a $q \times V$ group spatial map matrix of source signals where the rows of $\boldsymbol{S}$ are assumed to be independent. An error term can be added to (2) to represent residual variability in observed data, which leads to probabilistic ICA models. A key characteristic of TC-GICA is that distributed pattern of brain functional networks, represented by $\boldsymbol{S}$, is assumed to be the same across subjects and only the temporal dynamics of the networks, or the mixing processes, can be potentially different across subjects. Currently, the most commonly used group ICA methods such as the Group ICA of fMRI Toolbox (GIFT) by Calhoun et al. (2001) and the tensor probabilistic ICA (PICA) model by Beckmann and Smith (2005) are developed under the TC-GICA framework. Recently, Guo and Pagnoni (2008) and Guo (2011) have proposed general group PICA models which can flexibly model various types of between-subject variability in temporal dynamics of fMRI signals. However, the general group PICA models are also developed under the TC-GICA framework.

When comparing TC-GICA in (2) with the single subject ICA model in (1), we can see TC-GICA essentially performs group ICA in a similar manner to single-subject ICA by conducting ICA decomposition of a two-dimensional data matrix. A major limitation of TC-GICA is that it assumes brain functional networks have the same distributed patterns across subjects in ICA decomposition. This assumption is often questionable given the differences in subjects' functional networks due to biological and demographic factors (Meunier et al., 2008). Under TC-GICA, between-subject variability in spatial source signals can only be addressed by using heuristic post-ICA regressions to reconstruct subject-specific spatial maps and temporal responses. For example, the GIFT method (Calhoun et al., 2001) uses a back-construction method and the tensor PICA model uses a dual-regression approach (Beckmann et al., 2009). Therefore, TC-GICA only provides ad-hoc post-ICA solutions to capture between-subject variability but does not directly account for it in the group ICA, which can lead to loss of accuracy and efficiency in estimating brain functional networks.

We propose a hierarchical group PICA model to formally model subject-specific effects in both temporal and spatial domains in ICA decomposition of multi-subject fMRI data. The proposed model is composed of two stages where the first stage decomposes each subject's fMRI data into ICs characterized by subject-specific temporal dynamics and spatial sources signals and the second stage models each subject's spatial source signals in terms of population-level spatial signals modulated by subject-specific variability. By directly modeling between-subject variability in group ICA, our hierarchical PICA model is fundamentally different from TC-GICA which is ICA decomposition assuming homogeneous spatial source signals followed by ad-hoc adjustment for between-subject variability. In comparison to TC-GICA, our proposed hierarchical group PICA model has several advantages: 1) our method provides the first formal group ICA model for simultaneously estimating brain functional networks at both the population- and subject-level. 2) Unlike TC-GICA, the proposed hierarchical PICA decomposes group fMRI data in a multi-subject space as the outer-product of subject-specific time courses and spatial maps. 3) By systematically accounting for between-subject variability in both the temporal and

spatial domains, our method can potentially provide more accurate estimation of functional networks at both the population- and subject-level.

We propose a maximum likelihood (ML) approach for estimating the hierarchical group PICA model. We develop an EM algorithm which has explicit forms for both the E-step and M-step. Through simulation studies, we show that the proposed method provides more accurate estimation of fMRI source signals in both spatial and temporal domains than commonly used group ICA methods, especially when there is high between-subject variability in the source signals.

2. Methods

In this section, we first present the details of the hierarchical group PICA model. We then develop a maximum likelihood approach via the EM algorithm for model estimation.

2.1 A hierarchial group PICA model

We propose a hierarchical group PICA model for analyzing multi-subject fMRI data. To set notation, let $i = 1, \ldots, N$ index subjects, $t = 1, \ldots, T$ index time points, and $\upsilon = 1, \ldots, V$ index voxels. Let $\boldsymbol{y}_i(\upsilon)$ be a $T \times 1$ data vector representing the observed fMRI data from subject i at voxel υ. The first-level model of the hierarchical group PICA is defined as follows.

The first-level model—

$$\boldsymbol{y}_i(\upsilon) = \boldsymbol{A}_i \boldsymbol{\xi}_i(\upsilon) + \boldsymbol{\varepsilon}_i^{(1)}(\upsilon), \quad (3)$$

where $\boldsymbol{\xi}_i(\upsilon)$ is a $q \times 1$ vector with the ℓth($\ell = 1, \ldots, q$) element $\xi_{i\ell}$ representing the spatial source signal of the ℓth IC at voxel υ. We assume the q spatial source signals in $\boldsymbol{\xi}_i(\upsilon)$ are statistically independent and follow non-Gaussian distributions. $\boldsymbol{A}_i$ is the $T \times q$ mixing matrix for subject i since $\boldsymbol{A}_i$ mixes the q statistically independent source signals in $\boldsymbol{\xi}_i(\upsilon)$ to generate the observed fMRI data $\boldsymbol{y}_i(\upsilon)$. Furthermore, $\boldsymbol{A}_i$ characterizes temporal dynamics of the ICs where each column of $\boldsymbol{A}_i$ representing the time course of an IC. $\boldsymbol{\varepsilon}_i^{(1)}(\upsilon)$ is a $T \times 1$ vector representing the noise in the ith subjects' data at voxel υ and $\boldsymbol{\varepsilon}_i^{(1)}(\upsilon) \sim MVN(\boldsymbol{0}, \boldsymbol{\Sigma}_\upsilon^{(1)})$ where $\boldsymbol{\Sigma}_\upsilon^{(1)}$ is a $T \times T$ error covariance matrix for the υth voxel. Since the spatial correlation across voxels is modelled by the spatial source signals, the noise term is assumed to be independent across voxels (Beckmann and Smith, 2004). Prior to ICA, preliminary analysis such as pre-whitening (Bullmore et al., 1996) can be performed to remove temporal correlations in the noise term and to standardize the variability across voxels. Therefore, the noise term in PICA models is often assumed to have an isotropic covariance (Beckmann and Smith, 2004, 2005; Guo and Pagnoni, 2008; Guo, 2011), i.e., $\boldsymbol{\Sigma}_\upsilon^{(1)} = \sigma^2_{\varepsilon^{(1)}} \boldsymbol{I}_T$.

The first-level model decomposes each subject's observed fMRI data into a linear combination of neural source signals characterized by subject-specific distributed patterns and temporal dynamics. Compared to the classic noise-free ICA model in (1), the proposed PICA model includes an additional Gaussian noise term to account for additional variability in the observed fMRI BOLD effects that are not captured by the extracted q source signals.

The inclusion of the noise term can help address the issue of overfitting and also provide the possibility for formal statistical inference on the source signals (Beckmann and Smith, 2004). The ICA decomposition in the first-level model is known as spatial ICA since statistical independence is assumed for the spatial maps of source signals. For fMRI data, spatial ICA has become dominant because the spatial independence assumption is well suited to the sparse nature of spatial patterns for most cognitive activation paradigms (McKeown et al., 1998).

The second-level model—In the second stage of the hierarchical model, we model subject-specific spatial source signals in term of population-level source signals,

$$\boldsymbol{\xi}_i(\upsilon)=\boldsymbol{s}(\upsilon)+\boldsymbol{\varepsilon}_i^{(2)}(\upsilon), \quad (4)$$

where $\boldsymbol{s}(\upsilon)$ is a $q \times 1$ vector of population-level spatial source signals of the q statistically independent and non-Gaussian ICs, $\boldsymbol{\varepsilon}_i^{(2)}(\upsilon)$ is a $q \times 1$ vector of random variability representing the deviation of the ith subject's spatial source signals from the population mean. We assume $\boldsymbol{\varepsilon}_i^{(2)}(\upsilon) \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}^{(2)})$ where $\boldsymbol{\Sigma}^{(2)}=\mathrm{diag}(\sigma^2_{\varepsilon^{(2)},1}, \ldots, \sigma^2_{\varepsilon^{(2)},q})$ with $\sigma^2_{\varepsilon^{(2)},\ell}$ representing the between-subject variability for the ℓth ($\ell = 1, \ldots, q$) spatial source signal. By specifying IC-specific variances in $\boldsymbol{\Sigma}^{(2)}$, we allow the between-subject variability to vary across ICs.

Source Distribution Model—The population-level spatial source signals $\boldsymbol{s}(\upsilon)$ in the second-stage model (4) are unobserved and need to be estimated based on the observed fMRI data. We propose to model $\boldsymbol{s}(\upsilon)$ through probabilistic distribution functions to provide a statistical framework for estimation and inference of the hierarchical group PICA model. Following previous work (Guo and Pagnoni, 2008; Guo, 2011), we specify a mixture of Gaussian distributions as our source distribution model. The choice of the source distribution model is based on the fact that the source signals in functional brain images are generally attributed to a small percentage of voxels in the brain whereas most brain areas exhibit background fluctuations (Biswal and Ulmer, 1999). The mixture of Gaussians is well suited to model such mixed patterns by employing different Gaussian components to capture the distribution of the small proportion of activated voxels and the distribution of the majority of the brain areas that are not strongly related to the signal. In addition, the mixture of Gaussians is flexible enough to model various types of non-Gaussian source signals including both sub- and super-Gaussian signals in fMRI data (Xu et al., 1997). Another important advantage of the mixture of Gaussian distributions is that it offers tractable mathematical properties that facilitate the estimation of the proposed hierarchical group PICA model, as will be shown later.

Let $\boldsymbol{s}(\upsilon) = [s_1(\upsilon), \ldots, s_q(\upsilon)]'$ where $s_\ell(\upsilon)$ is the spatial source signal for the ℓth ($\ell = 1, \ldots, q$) IC at voxel υ. We assume each source signal follows a mixture of Gaussian distributions, i.e., for $\ell = 1, \ldots, q$,

$$f(s_\ell(\upsilon);\boldsymbol{\varphi}_\ell)=\sum_{j=1}^{m}\pi_{\ell j}\psi(s_\ell(\upsilon);\mu_{\ell j},\sigma_{\ell j}^2), \quad (5)$$

where m is the number of Gaussian density components in the mixture, $\psi(s_{\ell\upsilon};\mu_{\ell j},\sigma_{\ell j}^2)$ is the density function for the jth Gaussian component with mean $\mu_{\ell j}$ and variance $\sigma_{\ell j}^2$, $\pi_{\ell j}$ is the proportion of the jth Gaussian component which satisfies $0 \le \pi_{\ell j} \le 1$ and $\sum_{j=1}^{m}\pi_{\ell j}=1$. $\boldsymbol{\varphi}_\ell=\{\pi_{\ell j},\mu_{\ell j},\sigma_{\ell j}^2\}$ represents the parameters associated with the mixture of Gaussian distributions for the ℓth IC. It has been shown that a mixture of two to three Gaussian components is sufficient to capture the distribution of fMRI spatial signals (Beckmann and Smith, 2004; Guo and Pagnoni, 2008; Guo, 2011), with the different Gaussian components corresponding to background noise, negative or positive BOLD effects in fMRI signals. Given the statistical independence between the source signals, the probability density function (pdf) of $\boldsymbol{s}(\upsilon)$ is then $f(\boldsymbol{s}(\upsilon);\boldsymbol{\varphi})=\prod_{\ell=1}^{q}f(s_\ell(\upsilon);\boldsymbol{\varphi}_\ell)$ where $\boldsymbol{\varphi} = \{\boldsymbol{\varphi}_\ell\}$.

To facilitate derivations involving the mixture of Gaussian distributions, we define latent class variable $z_\ell(\upsilon)$ for the ℓth ($\ell = 1, \ldots, q$) IC at voxel υ where $z_\ell(\upsilon)$ takes value in [1, …, m] with probability $p\,(z_\ell(\upsilon) = j) = \pi_{\ell j}$ ($1 \le j \le m$) and that $\boldsymbol{z}(\upsilon) = (z_1(\upsilon), \ldots, z_q(\upsilon))'$. Conditional on $z_\ell(\upsilon)$, we have $p(s_\ell(\upsilon)|z_\ell(\upsilon)=j)=\psi(s_\ell(\upsilon);\mu_{\ell j},\sigma_{\ell j}^2)$. The latent class variable will greatly facilitate our estimation procedure later, since it allows us to deal with the non-Gaussian source signals under the convenient Gaussian framework.

2.2 Estimation

2.2.1 Maximum Likelihood Estimation—Classic noise-free ICA models are often estimated using algorithms that maximize statistical independence of the extracted ICs where the independence is usually measured by nonGaussianity (Delyon et al., 1999; Hyvrinen, 1999) or mutual information (Bell and Sejnowski, 1995) of the components. Specifically, by inverting the noise-free ICA model in (1), one can write the source signals as $\boldsymbol{S} = \boldsymbol{WY}$ where $\boldsymbol{W}$ is known as the unmixing matrix. Note that prior to applying an ICA algorithm, some preprocessing steps such as a dimension reduction and whitening of the observed data are usually performed so that the mixing matrix becomes a square and orthogonal matrix and hence can be inverted to obtain the unmixing matrix W. The unmixing matrix can be estimated as $\hat{\boldsymbol{W}} = \text{argmax}\, g(\boldsymbol{WY})$ where $g(\cdot)$ is an objective function that measures the statistical independence. The source signals are then estimated as $\hat{\boldsymbol{S}} = \hat{\boldsymbol{W}}\boldsymbol{Y}$. However, these algorithms are not appropriate for PICA models since the source signals in PICA can no longer be expressed as the product of the unmixing matrix and the data by inverting the model, due to the presence of the additive noise.

We propose to estimate the hierarchical group PICA model using a maximum likelihood (ML) approach. The ML approach is an effective estimation method for noise-free ICA and has been shown to be equivalent to the algorithm based on mutual information (Bell and Sejnowski, 1995) under certain circumstances. Compared to alternative algorithms, the ML method can be extended to PICA models and also offers a convenient way for statistical

inference and model comparisons in ICA. In our previous work (Guo and Pagnoni, 2008; Guo, 2011), we have developed ML methods for estimating group PICA models under the TC-GICA framework.

For the proposed hierarchical group PICA model, based on the models in (3)–(5), we can write the complete log-likelihood as,

$$\ln p(\mathsf{Y}, \eth\dot{\mathrm{I}}\check{\mathrm{c}}, \mathsf{S}, \mathsf{Z}; \boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{\upsilon=1}^{V}\left\{\ln\Psi[\boldsymbol{y}_i(\upsilon); \boldsymbol{A}_i\boldsymbol{\xi}_i(\upsilon), \boldsymbol{\Sigma}_{\upsilon}^{(1)}] + \ln\Psi[\boldsymbol{\xi}_i(\upsilon); \boldsymbol{s}(\upsilon), \boldsymbol{\Sigma}^{(2)}] + \ln\Psi[\boldsymbol{s}(\upsilon); \boldsymbol{\mu}_{\boldsymbol{z}(\upsilon)}, \Gamma_{\boldsymbol{z}(\upsilon)}] + \sum_{\ell=1}^{q}\ln\pi_{z_\ell(\upsilon)}\right\}, \quad (6)$$

where $\mathbf{Y} = \{\boldsymbol{y}_i(\upsilon)\}$, $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_i(\upsilon)\}$, $\mathbf{S} = \{\boldsymbol{s}(\upsilon)\}$ and $\mathbf{Z} = \{\boldsymbol{z}(\upsilon)\}$, $\boldsymbol{\theta} = \{\{\boldsymbol{A}_i\}, \boldsymbol{\Sigma}_{\upsilon}^{(1)}, \boldsymbol{\Sigma}^{(2)}, \boldsymbol{\varphi}\}$ represents the parameters in the likelihood function, Ψ is the pdf of a multivariate Gaussian distribution, $\boldsymbol{\mu}_{\boldsymbol{z}(\upsilon)} = [\mu_{1z_1(\upsilon)}, \ldots, \mu_{qz_q(\upsilon)}]'$ and $\Gamma_{\boldsymbol{z}(\upsilon)} = \mathrm{diag}(\sigma^2_{1z_1(\upsilon)}, \ldots, \sigma^2_{qz_q(\upsilon)})$.

2.2.2 The EM Algorithm—Since the complete log-likelihood in (6) involves unobserved spatial source signals and latent source states, we consider the EM algorithm to find the ML estimates for θ by maximizing the marginal log-likelihood of the observed data. We develop an EM algorithm which provides explicit forms for both E-step and M-step.

At the E-step, we find the conditional expectation of the complete log-likelihood function.

$$Q(\boldsymbol{\theta}|\hat{\theta}^{(\boldsymbol{k})}) = E_{\boldsymbol{s},\boldsymbol{\xi}_i,\boldsymbol{z}|\boldsymbol{y},\hat{\theta}^{(k)}}[\ln p(\mathsf{Y}, \eth\dot{\mathrm{I}}\check{\mathrm{c}}, \mathsf{S}, \mathsf{Z}; \boldsymbol{\theta})]. \quad (7)$$

The detailed definition of $Q(\theta|\hat{\theta^{(k)}})$ is presented in Web Appendix A. To evaluate the conditional expectation, we need to derive the conditional distributions $p(\boldsymbol{\xi}_i|\boldsymbol{y}, \theta)$, $p(\boldsymbol{s}|\boldsymbol{y}, \theta)$, $p(\boldsymbol{s}, \boldsymbol{\xi}_i|\boldsymbol{y}, \theta)$ and $p(\boldsymbol{z}|\boldsymbol{y}, \theta)$. Note that we omit the voxel argument in the conditional distributions for notation convenience. The direct approach to obtain these conditional distributions is by integration of products of pdfs from the first- and second-level models and the source distribution model, which involves lengthy and cumbersome derivations.

In this paper, we propose a fast and convenient approach for deriving the conditional distributions in the E-step. This approach is composed of three steps. In Step 1, we derive the conditional distributions $p(\boldsymbol{\xi}_i|\boldsymbol{z}, \boldsymbol{y}, \theta)$, $p(\boldsymbol{s}|\boldsymbol{z}, \boldsymbol{y}, \theta)$, and $p(\boldsymbol{s}, \boldsymbol{\xi}_i|\boldsymbol{z}, \boldsymbol{y}, \theta)$. Specifically, by conditioning on the latent state $\boldsymbol{z}$, we can rewrite the source distribution model as a linear model with Gaussian residuals. We then collapse the proposed hierarchical model and source distribution model into a nonhierarchical representation in the form of a classical linear model. We can show that both the subject-specific source signals $\boldsymbol{\xi}_i(\upsilon)$ and population-level spatial source signals $\boldsymbol{s}(\upsilon)$ can be written as linear functions of the parameters in this linear model. Therefore, the conditional distributions $p(\boldsymbol{\xi}_i|\boldsymbol{z}, \boldsymbol{y}, \theta)$, $p(\boldsymbol{s}|\boldsymbol{z}, \boldsymbol{y}, \theta)$ and $p(\boldsymbol{s}, \boldsymbol{\xi}_i|\boldsymbol{z}, \boldsymbol{y}, \theta)$ can be readily derived. The details of Step 1 are presented in Web Appendix B. Next, as Step 2, we derive $p(\boldsymbol{z}|\boldsymbol{y}, \theta)$ using the Bayes's theorem (Web Appendix B). Finally, in Step 3, we obtain the conditional distributions $p(\boldsymbol{\xi}_i|\boldsymbol{y}, \theta)$, $p(\boldsymbol{s}|\boldsymbol{y}, \theta)$ and $p(\boldsymbol{s}, \boldsymbol{\xi}_i|\boldsymbol{y}, \theta)$ by convolving the distributions derived in the previous two steps.

For the M-step, the EM algorithm obtains the updated estimates by maximizing $Q(\theta|\hat{\theta}^{(k)})$ w.r.t. θ. We derive explicit noniterative solutions for all the parameters. More specifically, the updated estimates for the subject-specific mixing matrix and the noise variance-covariance matrix in the first-level model is,

$$\hat{\boldsymbol{A}}_i^{(k+1)} = \left\{ \sum_{v=1}^{V} \boldsymbol{y}_i(v) E[\boldsymbol{\xi}_i'(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] \right\} \left\{ \sum_{v=1}^{V} E[\boldsymbol{\xi}_i(v)\boldsymbol{\xi}_i'(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] \right\}^{-1}, \quad (8)$$

$$\hat{\boldsymbol{\Sigma}}_v^{(1)(k+1)} = \frac{1}{N} \sum_{i=1}^{N} \left\{ \boldsymbol{y}_i(v)\boldsymbol{y}_i'(v) - 2\boldsymbol{y}_i(v) E[\boldsymbol{\xi}_i'(v) | \boldsymbol{Y}(v), \hat{\theta}^{(\boldsymbol{k})}] \hat{\boldsymbol{A}}_i^{(k+1)\prime} + \hat{\boldsymbol{A}}_i^{(k+1)} E[\boldsymbol{\xi}_i(v)\boldsymbol{\xi}_i'(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] \hat{\boldsymbol{A}}_i^{(k+1)\prime} \right\}. \quad (9)$$

Under the common assumption $\boldsymbol{\Sigma}_v^{(1)} = \sigma^2_{\varepsilon^{(1)}} \boldsymbol{I}$, we can show that,

$$\begin{aligned} \hat{\sigma}^{2(k+1)}_{\varepsilon^{(1)}} &= \frac{1}{TNV} \left\{ \sum_{i=1}^{N} \sum_{v=1}^{V} \boldsymbol{y}_i(v)' \boldsymbol{y}_i(v) \right. \\ &\left. - 2\boldsymbol{y}_i'(v) \hat{\boldsymbol{A}}_i^{(k+1)} E[\boldsymbol{\xi}_i(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] + tr\{ \hat{\boldsymbol{A}}_i^{(k+1)\prime} \hat{\mathbf{A}}_i^{(k+1)} E[\boldsymbol{\xi}_i(v)\boldsymbol{\xi}_i(v)' | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] \} \right\}. \end{aligned} \quad (10)$$

The variance estimator for the second-level model is

$$\hat{\sigma}^{2(k+1)}_{\varepsilon^{(2)},\ell} = \frac{1}{NV} \left\{ \sum_{i=1}^{N} \sum_{v=1}^{V} E[\xi^{(2)}_{i\ell}(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] - 2E[\xi_{i\ell}(v) s_\ell(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] + E[s^{(2)}_\ell(v) | \boldsymbol{Y}(v), \hat{\theta}^{(k)}] \right\}, \text{for} \ell = 1, \ldots q, \quad (11)$$

The parameters in the source distribution model can be estimated as follows,

$$\hat{\pi}^{(k+1)}_{\ell j} = \frac{1}{V} \sum_{v=1}^{V} p(z_\ell(v) = j | \boldsymbol{Y}(v), \hat{\theta}^{(k)}), \quad (12)$$

$$\hat{\mu}^{(k+1)}_{\ell j} = \frac{\sum_{v=1}^{V} \left[p(z_\ell(v) = j | \boldsymbol{Y}(v), \hat{\theta}^{(k)}), E(s_\ell(v) | z_\ell(v) = j, \boldsymbol{Y}(v), \hat{\theta}^{(k)}) \right]}{\sum_{v=1}^{V} p(z_\ell(v) = j | \boldsymbol{Y}(v), \hat{\theta}^{(k)})}, \quad (13)$$

$$\hat{\sigma}^{2(k+1)}_{\ell j} = \frac{\sum_{v=1}^{V} \left[p(z_\ell(v) = j | \boldsymbol{Y}(v), \hat{\theta}^{(k)}), E(s^2_\ell(v) | z_\ell(v) = j, \boldsymbol{Y}(v), \hat{\theta}^{(k)}) \right]}{\sum_{v=1}^{V} p(z_\ell(v) = j | \boldsymbol{Y}(v), \hat{\theta}^{(k)})} - [\hat{\mu}^{(k+1)}_{\ell j}]^2. \quad (14)$$

The conditional moments of the source signals and the conditional probabilities of the latent states in (8)–(14) are evaluated based on the conditional distributions derived in the E-step.

2.2.3 Spatial maps for brain functional networks—After obtaining $\hat{\theta}$, we can estimate the subject-specific and population-level spatial source signals and their variability based on the mean and variance of their conditional distributions. The spatial IC maps can

then be obtained by plotting the estimated source signals or source signals standardized by their standard deviations (Guo and Pagnoni, 2008; Guo, 2011) to evaluate the relative activity of various brain locations in each source signal. In fMRI analysis, researchers often are also interested in thresholded IC maps to identify "significantly active" voxels in each functional network. We propose a thresholding method based on the mixture distributions for this purpose (Web Appendix C).

2.2.4 Data dimension reduction and whitening—In general, ICA incorporates several pre-prossessing steps including centering, dimension reduction and whitening in order to reduce the complexity of ICA decomposition (Hyvrinen et al., 2001). In group ICA analysis of multi-subject fMRI data, dimension reduction is generally performed to reduce the computational load and avoid overfitting (Hyvrinen et al., 2001; Beckmann and Smith, 2005; Calhoun et al., 2001; Guo and Pagnoni, 2008; Guo, 2011).

For the proposed hierarchical group PICA model, we perform a dimension reduction and whitening procedure based on probabilistic PCA (Beckmann and Smith, 2004) on the observed fMRI data from each subject, i.e.,

$$\tilde{\boldsymbol{y}}_i(v)=(\boldsymbol{\Lambda}_{i,q}-\tilde{\sigma}^2_{i,q}\boldsymbol{I}_q)^{-1/2}\boldsymbol{U}'_{i,q}\boldsymbol{y}_i(v), \quad (15)$$

where $\boldsymbol{U}_{i,q}$ and $\boldsymbol{\Lambda}_{i,q}$ contain the first q eigenvectors and eigenvalues based on the singular value decomposition of $\boldsymbol{Y}_i = [\boldsymbol{y}_i(1), \ldots, \boldsymbol{y}_i(V)]$, the error variance $\tilde{\sigma}^2_{i,q}$ represents the variability in $\boldsymbol{Y}_i$ that is not accounted by the q extracted ICs, and $\tilde{\sigma}^2_{i,q}$ is estimated by the average of the $T - q$ smallest eigenvalues that are not included in $\boldsymbol{\Lambda}_{i,q}$. The number of ICs, i.e., q, can be determined using the Laplace approximation method (Minka, 2000).

After the dimension reduction and whitening, the first-level model can then be rewritten on the reduced and sphered space,

$$\tilde{\boldsymbol{y}}_i(v)=\tilde{\boldsymbol{A}}_i\boldsymbol{\xi}_i(v)+\tilde{\varepsilon}^{(1)}_i(v), \quad (16)$$

where $\tilde{\mathbf{A}}_i = \boldsymbol{H}_i\boldsymbol{A}_i$ and $\tilde{\varepsilon}^{(1)}_i(v)=\boldsymbol{H}_i\boldsymbol{\varepsilon}^{(1)}_i(v)$ with $\boldsymbol{H}_i=(\boldsymbol{\Lambda}_{i,q}-\tilde{\sigma}^2_{i,q}\boldsymbol{I}_q)^{-1/2}\boldsymbol{U}'_{i,q}$ denoting the transformation matrix. The transformed noise term in model (16) still follows a multivariate Gaussian distribution, i.e., $\tilde{\varepsilon}^{(1)}_i(v)\sim MVN(\mathbf{0},\tilde{\boldsymbol{\Sigma}}^{(1)}_v)$ with $\tilde{\boldsymbol{\Sigma}}^{(1)}_v=\boldsymbol{H}_i\boldsymbol{\Sigma}^{(1)}_v\boldsymbol{H}'_i$. The dimension reduction and whitening is equivalent to multiplying the original group fMRI data with a linear transformation matrix. Therefore, the estimation of the hierarchical group PICA model can be performed using the proposed EM algorithm with slight modifications. Specifically, we note that the dimension reduction is performed on the temporal domain while the spatial domain is not affected. Consequently, the modeling, estimation and thresholding of the spatial maps for the ICs remain the same after replacing $\boldsymbol{y}$ with $\tilde{\boldsymbol{y}}$.

3. Simulation Studies

We conducted simulation studies to evaluate the performance of the proposed hierarchical group PICA model. We simulated fMRI data from five underlying source signals, i.e., $q = 5$,

and considered three sample sizes with the number of subjects of $N = 5$, 15, and 40. For each source, we generated a 3D spatial map consisting of 4 slices where each slice had 40×50 voxels, resulting in a total of $V = 8000$ for each IC (Figure 1). The simulated 3D spatial maps were devised to mimic the sparse nature of activations in fMRI images. The temporal response for each of the source signals was a time series of length $T = 200$ that was estimated from real fMRI data and hence represented realistic temporal dynamics of fMRI source signals.

The population-level spatial maps were generated as images with Gaussian random variability of a standard deviation of 0.5 with linearly added source signals. The subject-specific spatial maps of the source signals were generated by adding subject-specific Gaussian noises to the population-level spatial maps. We considered three levels of between-subject heterogeneity where the standard deviation of the subject-specific effects was 20%, 40% and 67% of the intensity of the source signals, representing low, medium and high between-subject variability. To simulate subject-specific time courses for the source signals, we considered temporal dynamics in two major types of fMRI studies: task-related and resting-state fMRI. For task-related fMRI signals, the subject-specific time courses were generated as combinations of population-level time courses and individualized Gaussian noises. For resting-state fMRI signals, we generated subject-specific time courses that had similar frequency features but different phase patterns (Guo, 2011) (Figure 1).

After simulating the spatial maps and time courses for the source signals, Gaussian background noise with a standard deviation of 2 was added to the source signals to simulate the observed fMRI data. The generated fMRI data has a signal-noise-ratio (SNR) level of 0.43 representing a typical signal level in fMRI studies. Based on the simulated data, we compared the performance of the proposed hierarchical group PICA model to that of the GIFT and tensor PICA. Following previous work (Beckmann and Smith, 2005; Guo and Pagnoni, 2008; Guo, 2011), we evaluated the performance of each method by calculating the correlation between the true and estimated signals in both the temporal and spatial domains. Since ICA recovery is permutation invariant, each estimated IC was matched with the original source with which it had the highest spatial correlation.

We observe similar simulation results for task-related fMRI data and resting-state fMRI data. Therefore, we only present results for task-related fMRI in the paper (Table 1) and provide the results for resting-state fMRI in Web Table 1. The proposed EM algorithm for the hierarchical PICA model converged in all the simulation runs. The results show the proposed hierarchical PICA model provided more accurate estimates for the subject-specific spatial maps and time courses than the two existing methods under all simulation cases, especially for scenarios where the between-subject variability was high or the sample size was small. Specifically, when between-subject variability increased, the accuracy of the GIFT and tensor PICA methods decreased considerably for small to moderate sample sizes while the estimates based on the hierarchical PICA remained highly accurate even for the sample size as small as 5 subjects. The proposed hierarchical PICA also showed higher accuracy for estimating population-level spatial maps. In particular, when there was high level of between-subject variability and a small sample size, the correlations between the true and estimated population-level maps were below 0.5 for the GIFT and tensor PICA,

while the correlation for the hierarchical PICA was about 0.73 (Table 1) which was almost 60% higher than the other two methods. Figure 2 plots the detected spatial activation regions based on the three methods for this simulation scenario. The hierarchical PICA showed much better performance in correctly detecting the true distributed patterns for each IC. In comparison, the average activation maps based on the GIFT and tensor PICA were noisier and also demonstrated some mismatches of ICs due to the low correlations between the true and estimated spatial IC maps (Figure 2).

4. An fMRI data example

We applied the proposed method to an fMRI study on the neurobiological correlates of Zen meditation (Guo and Pagnoni, 2008). Zen meditation practices have recently received considerable attention in the scientific community for their potential contribution to the phenomenological and epistemological aspects of cognitive sciences. An fMRI study was carried out to compare neural activity patterns between zen meditators and controls. We adapted a simple lexical decision paradigm employing semantic and nonsemantic stimuli: 50 words and 50 phonologically and orthographically matched nonword items were presented visually on a screen in pseudo-random temporal order and subjects were asked to respond whether the displayed item was “a real English word” via an MRI-compatible button-box with their left hand (index finger = “yes”, middle finger = “no”). Subjects were instructed to use the awareness of their breathing throughout the session as a reference point to monitor and counteract attentional lapses. The experimental task can be thus conceived as having a dual-layer structure: an ongoing meditative baseline condition and a phasic perturbation of this baseline by semantic and nonsemantic stimuli.

In our study, twelve Zen meditators with more than 3 years of daily practice (MEDT) were recruited along with twelve control subjects (CTRL) who never practiced meditation. The groups were matched for gender, age, and education level. All participants were native speakers of English and right-handed, except one meditator who was ambidextrous. A series of 520 functional MRI images were acquired during the experiment with a 3.0 Tesla Siemens Magnetom Trio scanner. Each of these fMRI images contained 53×63×46 voxels. The images were corrected for slice acquisition timing differences and subject movements, registered and spatially normalized to the MNI standard brain space.

We performed the data dimension reduction and whitening procedure on the Zen meditation data prior to performing ICA. Based on Laplace approximation (Minka, 2000), 14 ICs were appropriate. Therefore, each subject’s data was projected on the subspace spanned by the first 14 eigenvectors from the probabilistic PCA and whitened by the variance estimated from the residual eigenvalues of the projection. The processed multi-subject fMRI data from the 24 subjects in the control and meditation groups were then decomposed with the proposed hierarchical group PICA method. The ICs were estimated using the proposed EM algorithm implemented by in-house MATLAB programs developed by the authors and the computation time was about 5 hours.

Among the extracted components, we identified three components of particular interest on the basis of the putative neural systems involved in meditation and the execution of the

experimental task. Population-level spatial maps showing activated brain region in each component are presented in Figure 3 where voxels with an estimated conditional probability of activation exceeding 0.95 are labeled active. The first network includes the supplementary motor area (SMA), the hand region of the right sensorimotor cortex contralateral to the (left) hand that was pressing the button box, and the visual cortex, and is clearly functionally related to the phasic performance of the lexical decision task. The second network is a fronto-parietal system including the bilateral intraparietal sulcus and the supplementary eye fields, which is consistent with the general architecture of attentional function (Corbetta and Shulman, 2002). The third network represents what has come to be known as “default mode network” (DMN) (Raichle et al., 2001), a set of regions including the posterior cingulate, the medial prefrontal cortex, the lateral parietal cortex and the hippocampi, which is characterized by high metabolism at rest and decreased activity during a variety of demanding tasks.

The proposed hierarchical group PICA model also provides model-based estimation for subject-specific IC maps for each of the 24 subjects in the two groups. To threshold subject-specific spatial maps, we modeled the spatial intensity values in the estimated subject-specific IC maps with a mixture of Gaussian distributions and then calculated the posterior probability of activation based on the fitted mixture distribution. The activated brain regions in the subject-specific IC maps were then identified as voxels with posterior probability of activation exceeding 0.95. For illustration, Figure 4 presents the thresholded subject-specific spatial maps for the default mode network for each subject in the control and meditation groups. The thresholded subject-specific maps for the task-related network and attentional network are presented as Web Figures 1 and 2. These figures provide valuable information in terms of subject-specific distributed patterns of the functional networks. To most effectively demonstrate the between-subject similarity and variability in the functional networks, we provide summary maps for the subject-specific thresholded IC maps (Figure 5). For each brain location, Figure 5 plots the proportion of subjects in each of the two groups (CTL and MEDT) who showed activation in that location in their subject-specific maps. For the task-related network, Figure 5 shows that a large area of the visual cortex is consistently present in the task-related network for all subjects in both the control and meditation groups. For the attentional network, the bilateral intraparietal sulcus especially the left intraparietal sulcus is present in most subjects. For the default mode network, posterior cingulate cortex (PCC) and medial prefrontal cortex is the most consistently represented regions in the network across subjects. Our finding is consistent with previous research which showed that these two regions, particularly PCC, play a central role in the default mode network (Kalcher et al., 2012).

In addition to subject-specific spatial maps, the proposed hierarchical group PICA model allows us to estimate subject-specific time courses associated with the functional networks, which are presented in Web Figures 3, 4 and 5. To summarize between-subject variability in the temporal domain, we calculated the standard deviation of subject-specific temporal responses at each time point and then averaged these standard deviations across time (Web Table 2). Following Guo (2011), we distinguished between the task active state (task-on) when word or nonword stimuli were presented and the resting state (task-off). Results in Web Table 2 indicate that for the attentional network and the default mode network, the

meditators demonstrated lower between-subject variability in temporal responses as compared to the controls. This may be mainly attributed to the fact that the meditators have more regularized neural activity in these two networks, since Zen meditation can help enhance the capacity to voluntarily moderate neural processing in attentional control and baseline brain functioning (Guo and Pagnoni, 2008; Guo, 2011; Pagnoni et al., 2008). Web Table 2 also shows there is higher between-subject variability in the task-off periods than in the task-on periods. This result provides evidence to support findings in exploratory analysis that brain functional networks tend to have more homogenous temporal response patterns across subjects when they are engaged in experimental tasks.

5. Discussion

Compared to classic noise-free ICA, PICA models are theoretically and computationally more challenging to estimate due to the presence of noise terms. We develop a maximum likelihood estimation approach via the EM algorithm to estimate the hierarchical PICA model. Our EM algorithm provides explicit E-step and M-step and hence does not require iterative steps or numerical methods to derive the conditional expectation and to update parameter estimates. However, one issue with the EM algorithm is that the computational load rises quickly with the increase of the number of ICs. In our previous work (Guo, 2011), we proposed a variational approximation EM algorithm that significantly speeded up the estimation of temporal-concatenation group PICA models. Following a similar approach, one can derive a variational approximation EM to reduce the computation time for the proposed hierarchical group PICA model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the editor, the associate editor and the two referees for valuable comments and suggestions. We also thank Dr. Giuseppe Pagnoni for providing the Zen meditation data and interpretations of the results. This work was supported by NIH grants UL1TR000454 (NCATS) and 1 U18 NS082143-01 (NINDS).

References

Beckmann, C.; Mackay, C.; Filippini, N.; Smith, S. Proc HBM. San Francisco: 2009. Group comparison of resting-state fmri using multi-subject ica and dual regression.

Beckmann C, Smith S. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Transactions on Medical Imaging. 2004; 23:137–152. [PubMed: 14964560]

Beckmann C, Smith S. Tensorial extensions of independent component analysis for multisubject fMRI analysis. NeuroImage. 2005; 25:294–311. [PubMed: 15734364]

Bell A, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. Neural Computation. 1995; 7:1129–1159. [PubMed: 7584893]

Biswal B, Ulmer J. Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. Journal of Computer Assisted Tomography. 1999; 23:265–271. [PubMed: 10096335]

Bullmore E, Brammer M, Williams S, Rabe-Hesketh S, Janot N, David A, Mellers J, Howard R, Sham P. Statistical methods of estimation and inference for functional mr image analysis. Magnetic Resonance Medicine. 1996; 35:261–277.

Calhoun V, Adali T, Pearlson G, Pekar J. A method for making group inferences from functional MRI data using independent component analysis. Human Brain Mapping. 2001; 14:140–151. [PubMed: 11559959]

Corbetta M, Shulman G. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience. 2002; 3:201–215.

Delyon B, Lavielle V, Moulines E. Convergence of a stochastic approximation version of the em algorithm. Annals of Statistics. 1999; 27:94–128.

Guo Y. A general probabilistic model for group independent component analysis and its estimation methods. Biometrics. 2011; 67:1532–1542. [PubMed: 21517789]

Guo Y, Pagnoni G. A unified framework for group independent component analysis for multi-subject fMRI data. NeuroImage. 2008; 42:1078–1093. [PubMed: 18650105]

Hyvärinen A. Fast and robust fixed-point algorithm for independent component analysis. IEEE Trans. on Neural Networks. 1999; 10:626–634.

Hyvärinen, A.; Karhunen, J.; Oja, E. Independent Component Analysis. New York: Wiley; 2001.

Kalcher K, Huf W, Boubela R, Filzmoser P, Pezawas L, Biswal B, Kasper S, Moser E, Windischberger C. Fully exploratory network indepdnent component analysis of the 1000 functional connectomes database. Frontiers in Human Neuroscience. 2012; 6:301–311. [PubMed: 23133413]

McKeown M, Makeig S, Brown G, Jung T, Kindermann S, Bell A, Sejnowski T. Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping. 1998; 6:160–188. [PubMed: 9673671]

Meunier D, Achard S, Morcom A, Bullmore E. Age-related changes in modular organization of human brain functional networks. NeuroImage. 2008; 44:715–723. [PubMed: 19027073]

Minka, T. Technical Report 514. MIT Media Lab; 2000. Automatic choice of dimensionality for pca.

Pagnoni G, Cekic M, Guo Y. “thinking about not-thinking”: neural correlates of conceptual processing during zen meditation. PLoS ONE. 2008; 3:e3083. [PubMed: 18769538]

Raichle M, MacLeod A, Snyder A, Powers W, Gusnard D, Shulman G. A default mode of brain function. Proc. National Academy of Science USA. 2001; 98:676–682.

Xu, L.; C, C.; Yang, H.; Amari, S. Proc. of ICNN’97. Houston, TX, USA: 1997. Maximum equalization by entropy maximization and mixture of cumulative distribution functions; p. 1821-1826.

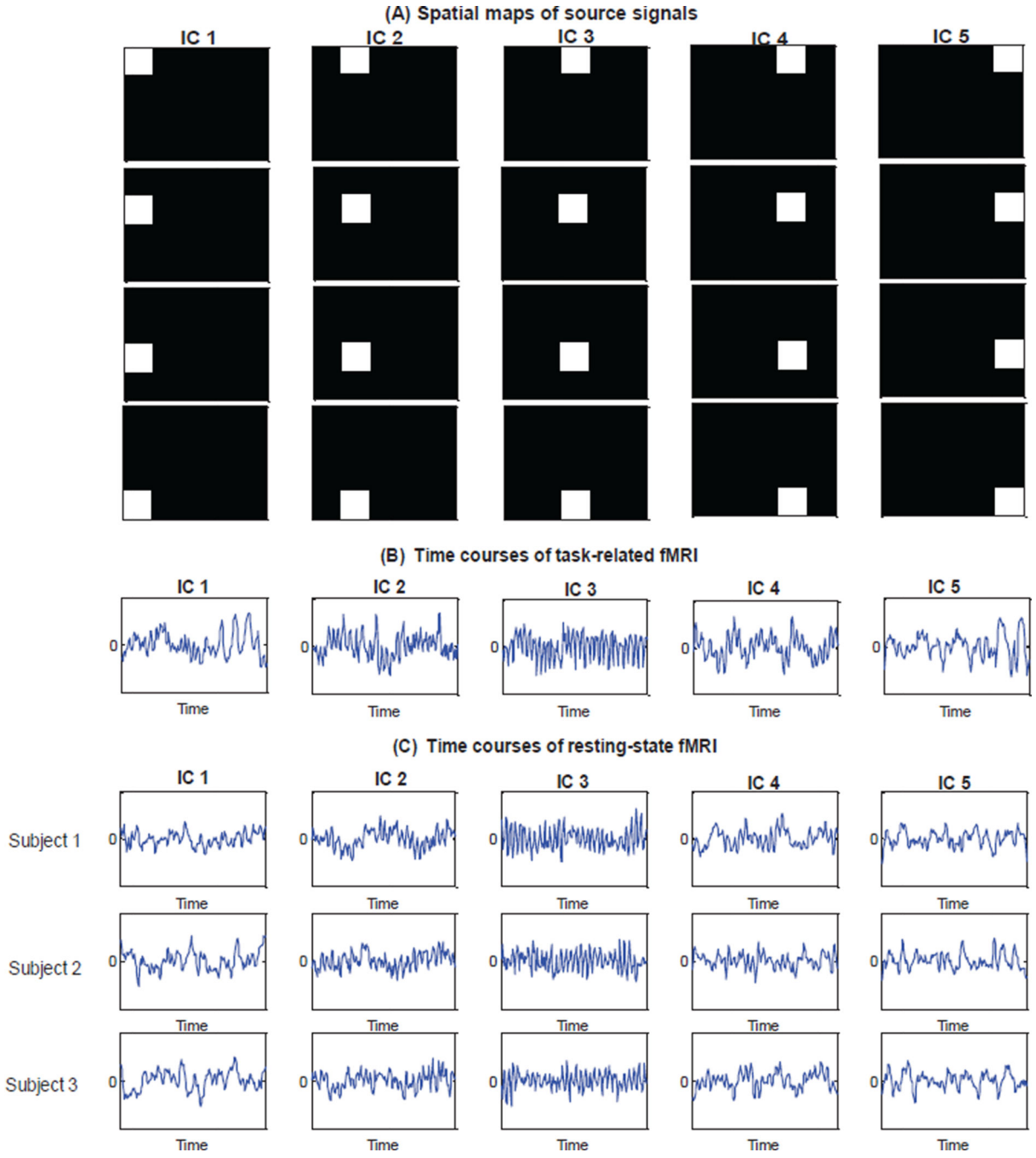

Figure 1. True spatial maps and time courses for generating source signals in the simulation study

The true spatial maps and time courses for the generated source signals in the simulation study. Panel (A) presents the spatial maps of the activated regions for the five independent components (ICs) where the activated locations are colored in white. Panel (B) plots the time courses associated with the five ICs for generating the task-related fMRI data. Panel (C) illustrates the subject-specific time courses for three subjects for the resting-state fMRI data which have similar frequency features but different phase patterns across subjects.

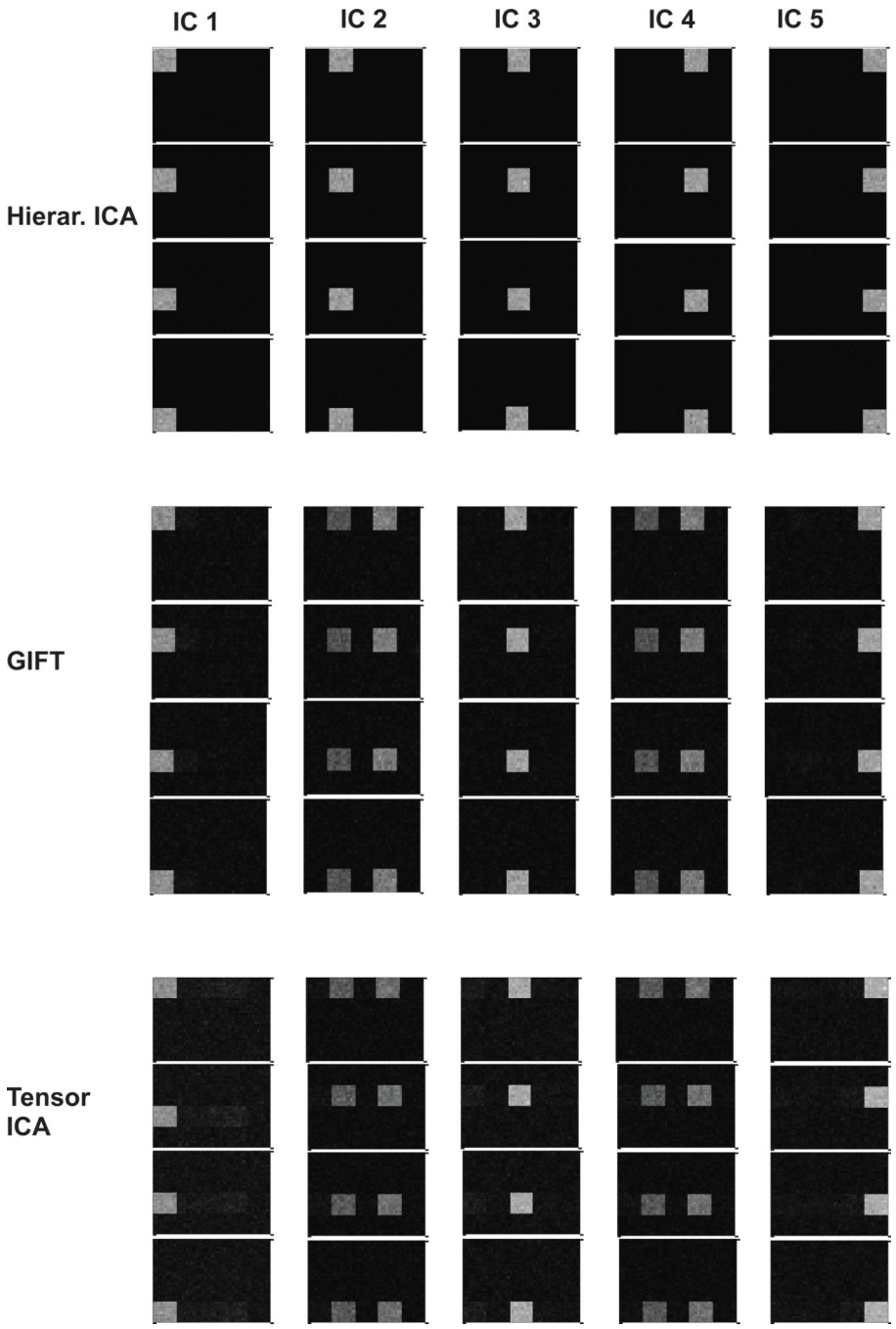

Figure 2. The average detected population-level activation maps in the simulation study

The average detected population-level activation maps for the five independent components in the simulation study when there is high between-subject variability and small sample size of $N = 5$. The average spatial maps are colored using white (1) to black (0) gray scale. The hierarchical group PICA model in the first panel detects the true spatial activated regions much better than the GIFT and tensor PICA in this case. For GIFT and tensor PICA, the average detected activation maps are noisier and also demonstrate some mismatches of ICs due to the low correlations between the true and estimated spatial IC maps.

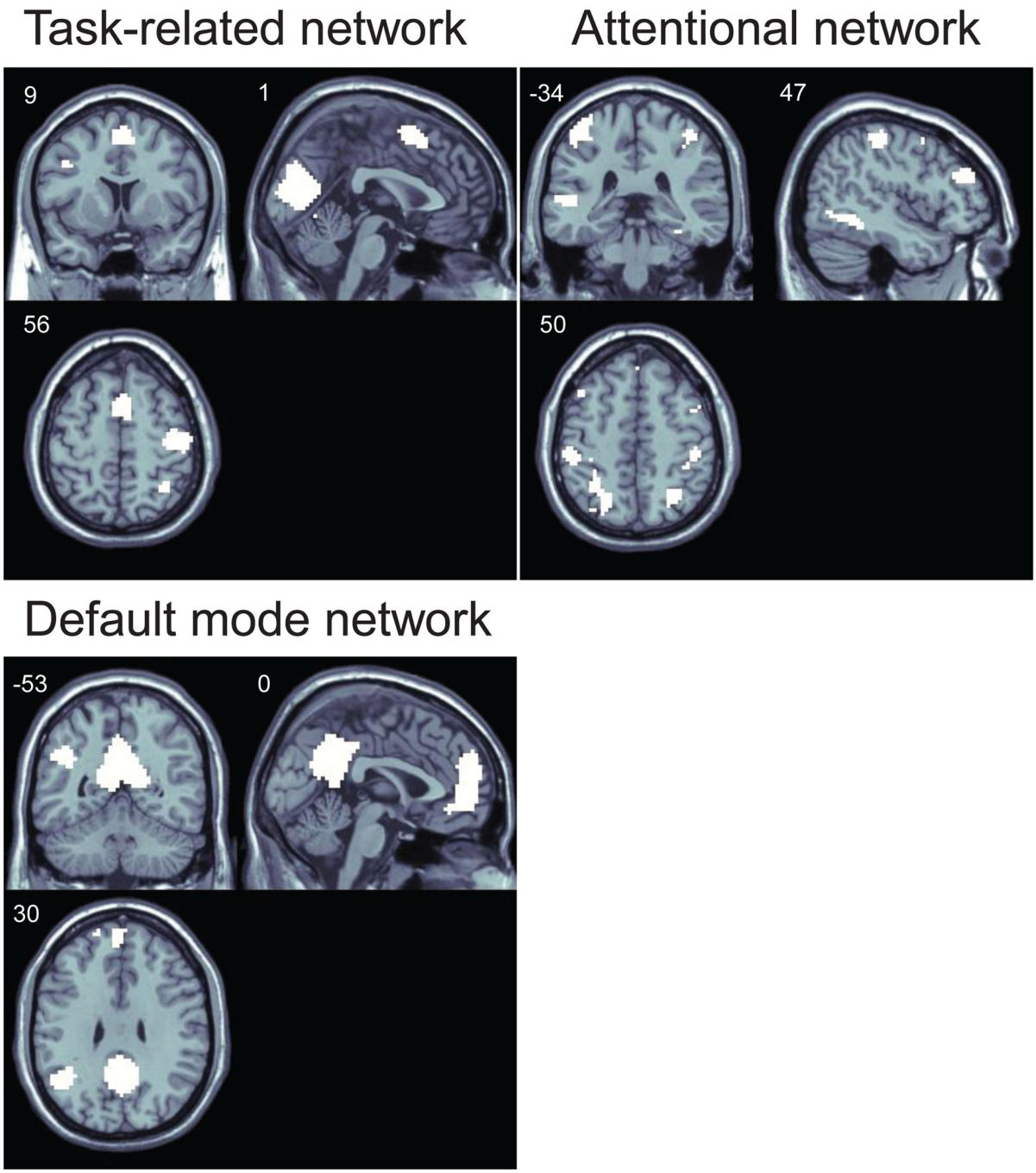

Figure 3. Estimated population spatial maps for three networks

Thresholded population-level activation maps for three selected independent components from the Zen meditation fMRI data based on all 24 subjects in the control and meditation groups. Voxels with a conditional probability of activation exceeding 0.95 are labeled active and colored in white in the maps. For each component, images are shown from coronal, sagittal and axial views.

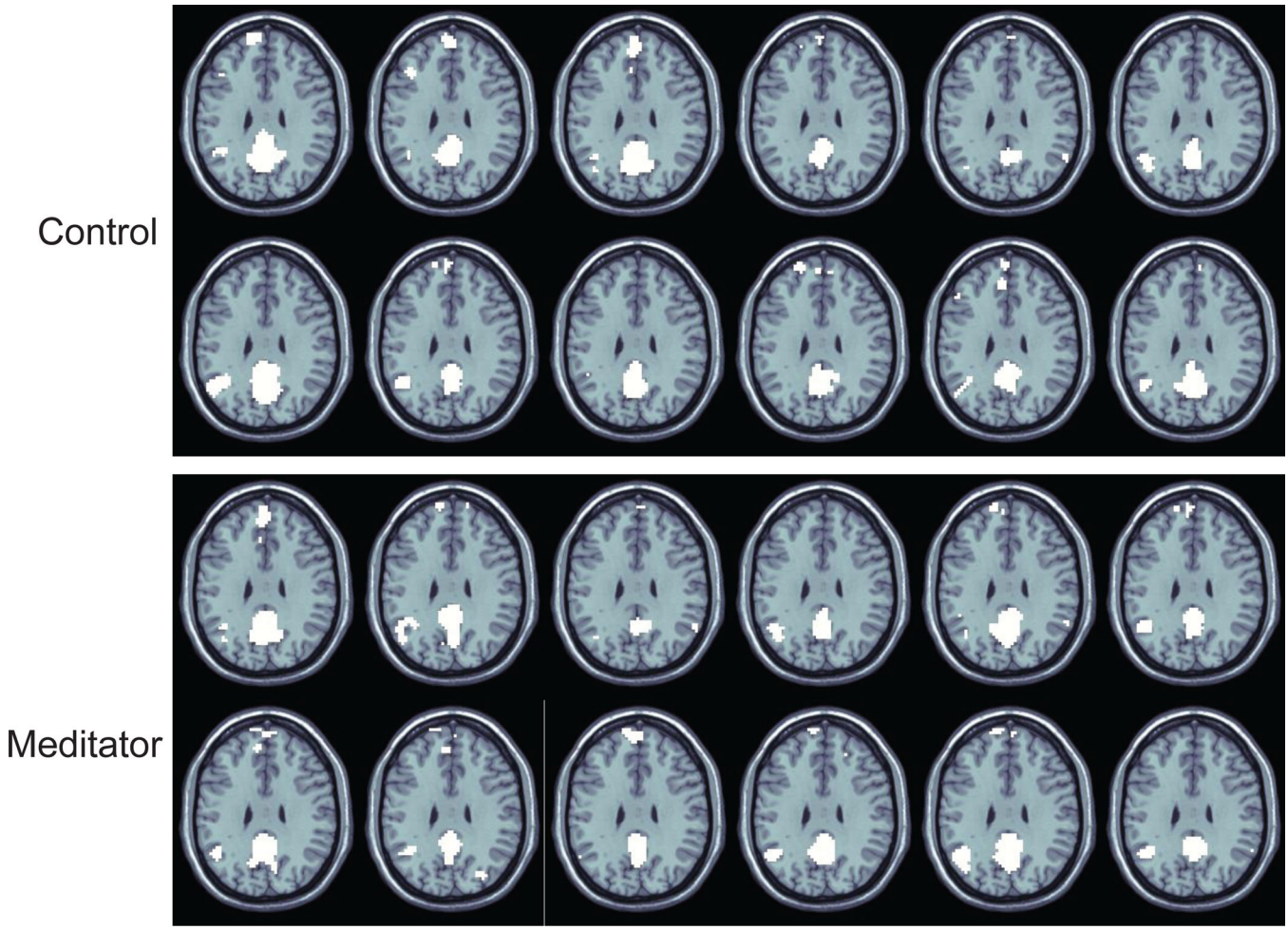

Figure 4. Thresholded subject-specific spatial maps for the default mode network

Thresholded subject-specific activation maps for the default mode network from the Zen meditation fMRI data. Voxels with a conditional probability of activation exceeding 0.95 are labeled active and colored in white in the maps. The axial slices are 30 mm from the anterior commissure. The upper panel presents the thresholded subject-specific maps for the 12 subjects in the Control group and the lower panel presents the thresholded subject-specific maps for the 12 subjects in the Meditation group.

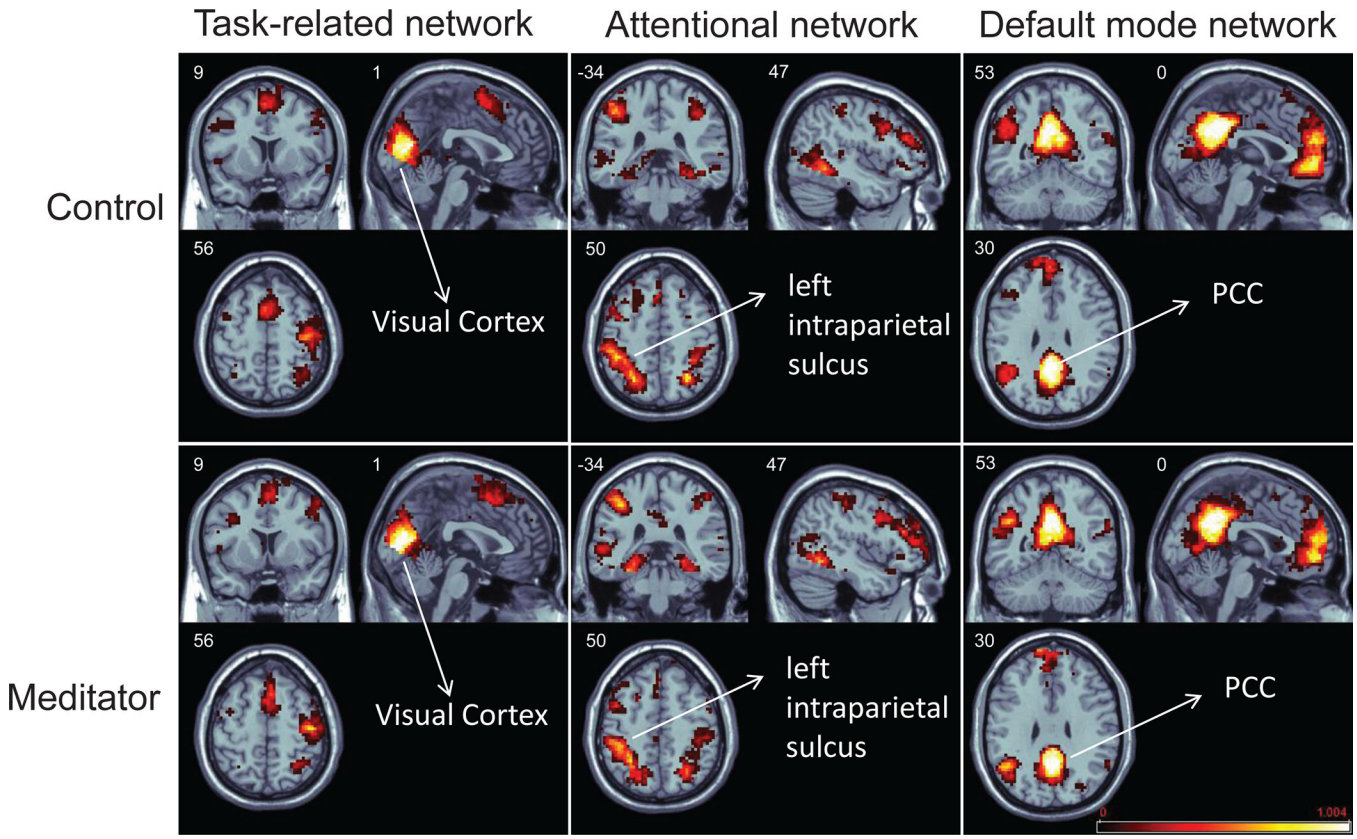

Figure 5. Across-subject consistency in activation patterns for brain functional networks

The between-subject consistency in activation patterns for three selected independent components from the Zen meditation fMRI data. The upper panel presents the between-subject consistency among the control subjects and the lower panel represents the between-subject consistency among the meditators. The figure plots the average of thresholded subject-specific activation maps, which represents the proportion of subjects in each of the two groups (Control and Meditator) that showed activation in their thresholded maps. For each functional network, images are shown from coronal, sagittal and axial views.

Table 1

Simulation results for task-related fMRI data based on 200 runs. fMRI data were generated from 5 underlying source signals. We compare the proposed hierarchical group PICA model (Hierar. ICA) against two existing group ICA methods: the GIFT and Tensor PICA. Values presented are the mean and standard deviation of correlations between the true and estimated: Subject-specific spatial maps, Population-level spatial maps and Subject-specific time courses, across the 200 simulation runs.

Btw-subj Var	**Subject-specific spatial maps Mean(SD)**			**Population-level spatial maps Mean(SD)**		
	GIFT	**Tensor ICA**	**Hierar. ICA**	**GIFT**	**Tensor ICA**	**Hierar. ICA**
Low						
N=5	0.976(0.013)	0.976(0.012)	0.988(0.001)	0.940(0.014)	0.940(0.013)	0.954(0.003)
N=15	0.977(0.010)	0.977(0.010)	0.988(0.001)	0.970(0.011)	0.970(0.011)	0.980(0.002)
N=40	0.983(0.008)	0.984(0.007)	0.989(0.001)	0.976(0.008)	0.977(0.008)	0.983(0.002)
Medium						
N=5	0.952(0.036)	0.949(0.043)	0.986(0.004)	0.811(0.019)	0.805(0.015)	0.878(0.007)
N=15	0.962(0.014)	0.963(0.016)	0.984(0.003)	0.912(0.015)	0.911(0.017)	0.939(0.004)
N=40	0.983(0.004)	0.983(0.005)	0.988(0.001)	0.929(0.008)	0.929(0.009)	0.945(0.003)
High						
N=5	0.721(0.057)	0.734(0.061)	0.975(0.007)	0.489(0.077)	0.477(0.081)	0.728(0.031)
N=15	0.854(0.041)	0.857(0.042)	0.975(0.005)	0.744(0.036)	0.732(0.016)	0.870(0.010)
N=40	0.961(0.031)	0.961(0.033)	0.977(0.003)	0.823(0.019)	0.823(0.035)	0.883(0.005)

Btw-subj Var.	Subject-specific time courses Mean(SD)		
	GIFT	Tensor ICA	Hierar. ICA
Low			
N=5	0.979(0.013)	0.979(0.013)	0.990(0.001)
N=15	0.980(0.010)	0.977(0.010)	0.988(0.001)
N=40	0.990(0.007)	0.991(0.008)	0.995(0.001)
Medium			
N=5	0.956(0.035)	0.951(0.042)	0.988(0.004)
N=15	0.964(0.013)	0.959(0.015)	0.986(0.003)

Btw-subj Var	Subject-specific spatial maps Mean(SD)			Population-level spatial maps Mean(SD)		
	GIFT	Tensor ICA	Hierar. ICA	GIFT	Tensor ICA	Hierar. ICA
N=40	0.987(0.004)	0.985(0.005)	0.991(0.001)			
High						
N=5	0.800(0.056)	0.798(0.064)	0.977(0.007)			
N=15	0.885(0.041)	0.875(0.043)	0.975(0.005)			
N=40	0.963(0.033)	0.963(0.033)	0.979(0.003)			

The proposed EM algorithm for the Hierarchical ICA model converged in all 200 simulation runs in each simulation scenario.