NeuroImage 56 (2011) 554–569

Contents lists available at ScienceDirect

NeuroImage

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

Effects of repeatability measures on results of fMRI sICA: A study on simulated and real resting-state effects

Jukka J. Remes [a,b,*,1], Tuomo Starck [a], Juha Nikkinen [a], Esa Ollila [c,f], Christian F. Beckmann [d,e], Osmo Tervonen [a], Vesa Kiviniemi [a], Olli Silven [b]

[a] Department of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland
[b] Department of Electrical and Information Engineering, University of Oulu, Oulu, Finland
[c] Department of Mathematical Sciences, University of Oulu, Oulu, Finland
[d] Division of Neuroscience and Mental Health, Imperial College London, London, UK
[e] Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, Oxford, UK
[f] Department of Signal Processing and Acoustics, SMARAD CoE, Aalto University School of Science and Technology, Espoo, Finland

ARTICLE INFO

Article history:
Received 19 November 2009
Revised 24 March 2010
Accepted 30 April 2010
Available online 6 May 2010

Keywords:
fMRI
BOLD
Resting-state
FastICA
ICASSO
Clustering
Repeatability analysis
Bootstrapping

ABSTRACT

Spatial independent components analysis (sICA) has become a widely applied data-driven method for fMRI data, especially for resting-state studies. These sICA approaches are often based on iterative estimation algorithms and there are concerns about accuracy due to noise. Repeatability measures such as ICASSO, RAICAR and ARABICA have been introduced as remedies but information on their effects on estimates is limited. The contribution of this study was to provide more of such information and test if the repeatability analyses are necessary. We compared FastICA-based ordinary and repeatability approaches concerning mixing vector estimates. Comparisons included original FastICA, FSL4 Melodic FastICA and original and modified ICASSO. The effects of bootstrapping and convergence threshold were evaluated. The results show that there is only moderate improvement due to repeatability measures and only in the bootstrapping case. Bootstrapping attenuated power from time courses of resting-state network related ICs at frequencies higher than 0.1 Hz and made subsets of low frequency oscillations more emphasized IC-wise. The convergence threshold did not have a significant role concerning the accuracy of estimates. The performance results suggest that repeatability measures or strict converge criteria might not be needed in sICA analyses of fMRI data. Consequently, the results in existing sICA fMRI literature are probably valid in this sense. A decreased accuracy of original bootstrapping ICASSO was observed and corrected by using centrotype mixing estimates but the results warrant for thorough evaluations of data-driven methods in general. Also, given the fMRI-specific considerations, further development of sICA methods is strongly encouraged.

Introduction

Spatial independent component analysis (sICA) (McKeown et al., 1998) has become one of the main data-driven analysis methods for functional magnetic resonance imaging (fMRI) data (for most current developments c.f. e.g. Calhoun et al., 2009; Demirci et al., 2009; Esposito et al., 2009; Kim et al., 2009a; Kim et al., 2009b; Kokkonen et al., 2009; Meda et al., 2009; Menz et al., 2009; Mohammadi et al., 2009; Rombouts et al., 2009; van de Ven et al., 2009; Wu et al., 2009; Zhang et al., 2009a; Zhang et al., 2009b). Yet, the validity and accuracy of such results are dependent on the ICA approach and parameters used. For example, repeatability analyses accounting for variability due to noise and methodological issues have been proposed (Himberg et al., 2004; Yang et al., 2008; Ylipaavalniemi and Vigario, 2008) but they have not been widely adopted. The issue becomes especially important when perspectives of these studies move from scientific interests to the clinical domain. This study evaluates experimentally the effects of precision and repeatability analyses on resulting independent components (ICs) corresponding to simulated and resting-state fMRI effects. The aim is to provide information for assessing existing studies and for improving future sICA analyses of fMRI data.

In ICA, in general, ICs are computed by projecting original observed data onto a subspace which makes the resulting data as non-Gaussian as possible. The projection directions are presented as unmixing vectors and estimated by the ICA approach used. sICA analyses of fMRI data often use optimization based estimation. One common example is the FastICA algorithm (Hyvarinen, 1999). These optimization processes can lead to varying results depending on starting points and to overly fitted estimates due to noise in the data (Himberg et al., 2004; Yang et al., 2008; Ylipaavalniemi and Vigario, 2008). ICASSO (Himberg et al., 2004),

* Corresponding author.
E-mail address: jukka.remes@iki.fi (J.J. Remes).
URL: http://www.iki.fi/jjremes (J.J. Remes).
[1] Main contributor. Planned the experimental setup for comparisons and carried out all the experiments.

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.04.268

RAICAR (Yang et al., 2008) and ARABICA (Ylipaavalniemi and Soppela, 2009) are frameworks which allow for selecting ICs appearing repeatedly during different runs of the ICA method. You can account for effects of different initializations, overfitting or both. An improved IC can be acquired through selecting an estimate from a single ICA repetition that corresponds to a cluster 'center' ("centrotype") when estimates from all repetitions are pooled and clustered (ICASSO and ARABICA). Alternatively, estimates in a cluster can be averaged (RAICAR). To obtain the clusters, ICASSO uses different flavors of hierarchical clustering, ARABICA applies an approach similar to hierarchical clustering (complete linkage) and RAICAR utilizes a special algorithm (c.f. also section "Limitations and other considerations on methods used in this study" for more details). The original motivation behind these frameworks has been to study the variability of IC estimates. Thus the frameworks serve as useful tools for studying data whether or not the improved IC estimates are used.

Another factor contributing to optimality of the estimates is convergence threshold, i.e. how small a difference between consecutive iterations of estimates is considered relevant. The threshold affects the precision of final overall estimates thus affecting e.g. IC map values and consequently spatial specificity of results.

The effects of different parameterizations of sICA have not been studied thoroughly in the same setting regarding fMRI results. Repeatability measures have been demonstrated in the context of stimulus studies and comparable simulated effects (Himberg et al., 2004; Yang et al., 2008; Ylipaavalniemi and Vigario, 2008) but comparisons on how different repeatability measures (initialization randomization, bootstrapping or both) affect the results are missing.

Moreover, resting-state fMRI data, which has also become a major point of interest in neuroscience (for most current developments c.f. e.g. Filippini et al., 2009; Kiviniemi et al., 2009b; Smith et al., 2009; Zhang et al., 2009a), is being often studied with sICA. Compared to effects induced by stimuli in activation studies, effects in resting-state fMRI data have significantly lower signal-to-noise ratio (SNR). As such, sICA of resting-state data should benefit from more sophisticated signal detection methods but there has not been information on performance differences between single-run sICA approaches and the repeatability measures available for resting-state data.

In this study, we investigated how using repeatability measures based on 1) random initializations and 2) both random initializations and data bootstrapping affect the IC estimates. Comparisons were made on single subject resting-state data with SNR-wise comparable simulated effects added to the imaged data. Quantifications of accuracy were carried out with respect to the simulated effects and the differences of all IC estimates between both repeatability approaches (1 and 2) were computed relative to each other. The results are provided as a function of convergence threshold used in ICA repetitions. In addition, spectral features of IC time courses (estimated mixing vectors) were visualized for ICs corresponding to executive function, visual cortex and posterior and anterior 'default-mode' resting-state networks (RSNs) as a demonstration of how bootstrapping affects overlearning regarding cognitively interesting ICs.

The main goal and the novel contribution of this work were to test if repeatability analyses and strict convergence thresholds are needed in sICA of fMRI. The results have implications for the reliability of existing fMRI sICA studies. In addition, we address the significance of testing analysis methodologies thoroughly prior to their wide deployment. As an example, new results concerning bootstrapping in this context are presented. On the other hand, future prospects in method development are also briefly discussed given the restrictive assumptions in established fMRI sICA methodology.

Materials and methods

The effects of repeatability measures were evaluated in two ways. Firstly, effects on detection of simulated signal sources embedded in real resting-state fMRI data were quantified with respect to original mixing. FastICA implementations were compared with repeatability analyses using multiple ICASSO approaches. Comparisons were done both regarding non-bootstrapping and bootstrapping analyses. ICASSO analyses included estimation of the mixing matrix based on the pseudoinverse of unmixing vector centrotypes (original ICASSO) and based on mixing vector centrotypes (modified ICASSO). Secondly, the similarities of all estimated components were quantified between FastICA and ICASSO analyses in the non-bootstrapping case and between ICASSO analyses in the bootstrapping case.

In addition, ICASSO analyses were compared between non-bootstrapping and bootstrapping approaches. Recognizable components corresponding to functional brain networks were selected. To determine the effects of bootstrapping, their mixing vectors (or IC time courses) were inspected with regard to their power spectra in non-bootstrapping and bootstrapping cases.

The following subchapters describe the undertaken steps in detail. These include 1) data acquisition, 2) pre-processing, 3) simulation of effects, 4) ICA methods, 5) IC sorting, 6) descriptive summary statistics used in inference, and 7) selection of RSN-related ICs.

fMRI data and pre-processing

A single male subject (age 30) with no history of psychiatric or neurological conditions was imaged at 1.5 T with T1-weighting and with BOLD EPI during rest (eyes closed, TR 1.8 s) in Oulu University Hospital. The session was part of a larger fMRI study approved by the ethical committee of Oulu University Hospital and written consent was provided by the subject. (See supplementary section "The details of MRI data acquisition" for more technical details on acquisition.)

The pre-processing included automatic correction for head motion in BOLD data (Jenkinson et al., 2002) followed by brain extraction (Smith, 2002) manually tuned for both the T1 and BOLD data. See supplementary section "The details of MRI data pre-processing" for technical details.

Simulated signal sources

Simulated signal sources were generated to serve as a main comparison point for ICA estimation quality. Sources were defined so that with regard to their signal-to-noise ratio (SNR) they would be comparable to real RSN-related fMRI signal sources. In addition, they were generated in compliance with models used for actual signal sources in inferential steps (mixture modeling) following estimation.

First the pre-processed fMRI data was subjected to analysis with probabilistic ICA (PICA) (Beckmann and Smith, 2004) to determine how much variance of data cognitively interesting ICs explain. FSL4 Melodic was used for PICA with symmetric approach and using skewness ("pow3" parameter for −nl option) to estimate the number of non-Gaussian components in the data. The default dimensionality estimation based on probabilistic PCA (PPCA) (Beckmann and Smith, 2004) was chosen.

The pre-processed data was then spatially smoothed with a 5 mm FWHM Gaussian kernel using the FSL4 fslmaths tool and subjected to another similar PICA analysis. Smoothing was carried out for better visual identification of RSN-related ICs. In this PICA the model order was set to the number of ICs estimated during previous PICA as PPCA overestimates model order in a case of smoothed data. All identifiable RSN-related ICs (e.g. visual cortex, auditory cortex, default-mode network, not shown in this article) attributed from 1% to 2% to the overall variance in the data according to Melodic generated report.

Two simulated signal sources were generated by independently sampling values from a gamma distribution ($k = 2$, theta $= 2$) for two random independent subsets of all brain voxels. Both positive and negative 'activity' were introduced to the sources. Negative values were just negated positive values from the same distribution. This way the generated sources correspond to noise-free effects regarding the mixture model used in the PICA setting (Beckmann and Smith, 2004) for positive and negative effects. Brain voxels were defined by the mask produced in the first Melodic run which was used for dimensionality estimation. The voxel subsets of both sources covered 10% (2939) of all brain voxels. One source contained 25% positive values and 75% negative values, the other one vice versa. All samplings were performed in Matlab using function "random" for gamma distributions and function "randperm" for random selection of voxels.

Other voxels were left to zero values. Natural noise in the real fMRI data, into which simulated values were later mixed, inevitably results in non-zero values in all voxels in the final estimated ICs. This accounts for Gaussian distribution in the mixture model as assumed (Beckmann and Smith, 2004) in the case of mixture modeling ICs related to real fMRI signal sources (IC estimation in "signal + noise" subspace).

We wanted to ensure generality of later results and to minimize effects of simulated sources on real fMRI sources in their joint estimation. Due to this random selection of 'active' voxels for simulated sources was preferred instead of defining coherent spatial structures. Spatial locations of effects can be randomized as sICA approaches compared in this study do not employ spatial priors such as grey matter segmentations.

In line with the previously determined percentage range given by the Melodic report for ICs of the real fMRI data, the simulated sources were then scaled so that one of them corresponded to 1% of whole mixed data variance, simulating that of a very low SNR RSN-related source and the other one to 2% simulating SNR-wise a very high SNR RSN-related source. These calculations could be performed before actual mixing using information from the first Melodic run (output "pcaD").

Finally, the scaled simulated sources were mixed to the pre-processed and non-smoothed fMRI data used as input for the first Melodic step. Mixing vector values for both sources were (pseudo) randomly sampled from uniform distribution (Matlab function "rand"), zero-averaged to correspond to temporal BOLD modulation associated with the source (interpretation given for mixing vectors in fMRI sICA models) and scaled to unit length. Even though temporally fMRI data can be considered autocorrelated, the random sampling of mixing vector values was preferred for generality. The mixing vectors were later used as the reference to which IC estimation was compared in terms of estimated mixing vectors.

The mixing vectors were collected to columns of mixing matrix (A) and signal sources to corresponding rows of signal matrix (S). The data produced by multiplying mixing matrix with simulated sources (A*S) and mapped back to image volumes was added to the pre-processed and non-smoothed fMRI data with the FSL4 fslmaths tool.

ICA analyses

The mixed data containing simulated effects mixed with real fMRI data were analyzed with 6 sICA approaches. All compared methods were based on widely used FastICA (Hyvarinen, 1999) and included original FastICA, a FastICA variant implemented in FSL4 Melodic (v. 3.05) software, original ICASSO (Himberg et al., 2004) and our modification of ICASSO. Both ICASSO methods were applied both with and without bootstrapping of data. Random initialization was used in both cases.

All 6 approaches were used to analyze the data 25 times at convergence threshold ("epsilon") values 0.0005, 0.0001, 0.00005, 0.00001, 0.000005, 0.000001, 0.0000005 and 0.0000001. The thresholds are defined in the same way in software implementations of all approaches. The model order estimation in Melodic (c.f. section "Simulated signal sources") was again first used once to determine the appropriate model order for mixed data and all decompositions were carried out using that model order. Skewness based contrast and symmetric approach were also used again in all analyses. Step size ("mu") was set to 1 and maximum iterations to 10,000. In ICASSO runs 100 repetitions were used. Instances of FastICA were sampled (pseudo)randomly (Matlab function randperm) from non-bootstrapped FastICA repetitions in respective 25 ICASSO runs.

Normal skewness is not as sensitive a contrast as e.g. kurtosis to non-Gaussian features. However, it was chosen as concerning fMRI sICA there has been suggestions that it models brain activity more plausibly than contrasts not accounting for asymmetry in signal source distributions (Suzuki et al., 2002; Stone et al., 2002). Skewness is also the default contrast in FSL Melodic which is a widely used analysis package and consequently the results obtained in this study are in this sense comparable with other results obtained with Melodic. Concerning the plurality of ICA repetitions all ICA computations were for practical reasons restricted to this single contrast only.

Results from all approaches were made comparable to each other in terms of pre-processings by applying FastICA and ICASSO methods on data pre-processed and reduced (PCA) by respective 25 FSL4 Melodic runs (Melodic output "melodic_pca", e.g. with option −Oall). Melodic pre-processings (e.g. variance normalization) follow the PICA model (Beckmann and Smith, 2004) and regarding FSL4 are documented in our previous study (Kiviniemi et al., 2009b). The specific details of group analysis do not apply here.

For original FastICA, Matlab implementation (www.cis.hut.fi/projects/ica/fastica) provided by FastICA authors was used (version 2.5). Also FSL4 Melodic was included in comparisons due to its widespread use and because it uses a FastICA version that differs a little from the original algorithm. Specifically, in FSL4 Melodic (up to v. 3.09), FastICA has been modified such that it does not subtract the mean from the data prior to the ICA step even though otherwise data are prewhitened. The data not having a mean of zero leads to a situation where FastICA update formulas for skewness based optimization contrast do not strictly hold as they have been originally derived with the assumption of prewhitened data. It is important to note, though, that global time course removal from fMRI data prior to sICA is more complicated than purely mathematical.

sICA is basically used to identify independent or sparse spatial volumes which, when back-mapped to anatomy, show the locations of non-Gaussian effects in the data. There is, however, also a domain specific need and goal to perceive mixing vectors (or time courses) related to the ICs as underlying coherent temporal dynamics in those brain areas. As in the case of correlation analyses that are based on the reference time course from a seed voxel or brain area, the IC maps and IC time courses are interpreted to reflect functional connectivity between brain areas. Some effects in the data like those related to RSNs can contribute considerably more to the mean than other effects just because some effects cover larger anatomical areas and as such more spatial samples (voxels) than others. If such a mean is subtracted then effects contributing more to it become attenuated in the resulting data and their projections might not contribute to values of optimization contrasts (e.g. skewness) as much as some other effects. This can then lead to poorer sensitivity in detection of these effects. Also there is potential risk of false positive findings as has been demonstrated with correlation analyses that utilize different mean removal methods as pre-processing steps (Fox et al., 2009; Murphy et al., 2009).

For original ICASSO, Matlab implementation (www.cis.hut.fi/projects/ica/icasso) provided by the ICASSO authors was used (version 1.22). In addition to using original ICASSO, we made a modified version (by enabling a feature already existing in the source code). In the modified version, final mixing vector estimates are based on cluster centrotype mixing vectors. In the original published ICASSO, mixing vectors are determined through the pseudoinverse of the unmixing matrix comprised of unmixing vectors corresponding to the cluster centrotypes. The pseudoinverse ensures that the resulting estimated ICA model is generative (i.e. original data can be reconstructed as a matrix product of mixing and ICs). Our preliminary studies suggested that, regarding the accuracy of mixing matrix values, there was a problem with this approach when applying ICASSO on fMRI data in bootstrapping settings. Such problems can make inferences based on temporal dynamics interpretation of mixing vectors invalid. Estimation by ICASSO software gives an unmixing matrix containing unmixing vectors from different FastICA repetitions and the corresponding pseudoinverse (mixing matrix). The modified version (denoted later by "modified ICASSO") is just a variant which returns the mixing matrix comprising corresponding mixing vectors from the same FastICA repetitions.

The resulting spatial IC maps were post-processed as in the PICA framework and by Melodic software (Beckmann and Smith, 2004). Maps from different runs were converted to statistical scores by dividing IC values by voxel-wise noise standard deviations estimated from Gaussian data left over by the PCA step (by multiplying maps with "Noise_stddev_inv" output from the corresponding Melodic run, produced by option −Oall). The mixture modeling was then applied on the score maps using the mixture modeling option in the FSL4 Melodic software.

Matching of estimated ICs between analyses

IC estimates were matched between all sICA runs in order for them to be comparable between approaches. This enabled computation of statistics considering 25 repetitions of each analysis. Additionally, ICs corresponding to the simulated effects were identified by comparing mixing vectors of all ICs with mixing originally defined for the simulated effects. All matchings were implemented in a Matlab environment.

The matching was done with the assumption of one-to-one correspondence of components between two sets of estimates. Due to this, some components end up with a lower matching IC from another set than what would be available in unrestricted matching. In the case of multiple-to-one matches, comparisons would have been practically impossible due to yielding branches to the execution path of this study. Possibly countless different permutations of possible matches would have had to have been investigated. Due to the great number of sICA analyses performed in this study also heuristic expert selection of matching ICs had to be ruled out for both practical considerations (time available) and for the possibility of bias introduced by subjective assessment.

The matching was based on computing inner product values between estimated mixing vectors, and in the case of simulated effects, between estimated and original mixing vectors. (See supplementary section "The details of component matching" for more technical details.) The following cases were carried out using the procedure:

- ICs from FastICA and Melodic runs were matched with modified ICASSO in the non-bootstrapped case.
- ICs from original ICASSO runs were matched with modified ICASSO separately in bootstrapping and non-bootstrapping cases.
- ICs from modified ICASSO in the bootstrapping case were matched with ICs from modified ICASSO in the non-bootstrapping case to compare the effects of bootstrapping on mixing vectors of IC related to real RSNs (c.f. section "Inspection of selected RSN-related ICs regarding bootstrapping").

Statistics

Statistics were defined to compare how well the ICs that corresponded to simulated effects match with the original mixing, and also to compare, concerning all ICs, how similar were the results that different approaches produce. As in the case of IC matching, the statistics were based on absolute values of inner products 1) between the original mixing vectors and corresponding estimates and 2) between all estimated mixing vectors from a pair of approaches. All values vary between zero and one as all vectors were scaled to unit length. All steps were carried out in a Matlab environment.

Concerning simulated effects, inner products between their original mixing and the mixing vectors of ICs related to the effects were computed separately for each repetition at each convergence threshold level and in each sICA approach. Since the original data has temporally zero mean and the original mixings were also defined so, the inner product values are very close to (if not equivalent to) temporal (Pearson's product-moment) correlation coefficients between the mixing vectors. Sample means of inner product values/correlations in the 25 repetitions were then computed at each convergence threshold level and in each sICA approach along with 99% level confidence intervals for the actual means.

After matching and the corresponding sorting, inner products were computed for the mixing vectors of the same ICs between modified ICASSO and other sICA approaches in each repetition and at each convergence threshold level. As in the matching procedure (c.f. section "Matching of estimated ICs between analyses") original FastICA and Melodic were compared in this way with modified ICASSO in the non-bootstrapping case and original ICASSO with modified ICASSO in both bootstrapped and non-bootstrapped cases. Also, inner products were computed for respective repetitions of modified ICASSO between bootstrapped and non-bootstrapped cases. As in the case of simulated effects, sample means and 99% level confidence intervals were computed at each convergence threshold level for inner products reflecting the similarity of repetitions of different approaches.

As we had our previous indications of problems in the use of the pseudoinverse, we also computed condition numbers for the unmixing matrices produced in ICASSO approaches. The condition number provided relative measures between non-bootstrapping and bootstrapping cases for how applicable matrix (pseudo)inversion is as an estimation procedure for mixing matrices in the context of ICASSO. The condition number was computed for an unmixing matrix as the product of a) 2-norm (largest singular value, Matlab norm-function) of the unmixing matrix estimated by ICASSO and b) 2-norm of the corresponding mixing matrix calculated through the pseudoinverse (Matlab pinv-function) by the original ICASSO approach.

Inspection of selected RSN-related ICs regarding bootstrapping

In addition to the quantification of similarity between different sICA approaches, RSN-related ICs were inspected. Corresponding mixing vectors were analyzed with respect to their frequency content to study how bootstrapping affects them.

As an illustrative example of spatial localization of studied components, thresholded score maps overlaid on the subject anatomy were created for the ICs. The maps made were based on (modified version) ICASSO ICs from a non-bootstrapped run at a convergence threshold of 0.0000001. For a more anatomically identifiable visualization of ICs and since no spatial smoothing was applied to the data in the pre-processing phases, thresholding masks for score maps were first smoothed using spatially smoothing probability maps produced by mixture modeling prior to thresholding them at alternative hypothesis test (Beckmann and Smith, 2004) level $P = 0.5$. (See supplementary section "The details of making the spatial maps" for more technical details.)

FSL4 MISCVIS tools ("overlay" and "slicer") were used to produce sagittal, coronal and axial slice images of thresholded score maps overlaid on subject anatomy at locations reported in our previous study (Kiviniemi et al., 2009b). Resulting overlay slice images were inspected visually and ICs clearly identifiable relating to real RSNs were selected for further study.

Time courses (i.e. mixing vectors) corresponding to the selected ICs were analyzed with respect to their power spectral content to study the effects of bootstrapping on IC estimation. Power spectral densities (PSDs) were estimated for the time courses from the 25 repetitions of ICASSO (modified version) in both bootstrapped and non-bootstrapped cases and at all 8 convergence threshold levels. The PSDs were estimated according to our previous methodology (Kiviniemi et al., 2009a). However, the window width used in this study was 256 points and the number of windows was 23, yielding better frequency resolution and still adequately robust estimation of power at individual frequency bins. The 25 PSDs from matched ICs were averaged for the bootstrapped and non-bootstrapped cases separately for all convergence thresholds and plotted for visual comparison. Matlab software was used for all the steps.

Assessing the effects of model order

However a model order is chosen – empirically or using automated methods (such as PPCA used in this study) – it is possible that the model order has an effect on similarities between the compared ICA methods. To test what kind of effects the choice of model order might have on the results, a coarse battery of control experiments were performed. ICA analyses (c.f. section "ICA analyses") were carried out with three other model orders in addition to the one acquired through PPCA. Each case was followed by the component matching (c.f. section "Matching of estimated ICs between analyses") and the computation of the statistics (c.f. section "Statistics").

The model orders chosen for these tests included automatically estimated model order (PPCA) minus 10, automatically estimated model order (PPCA) plus 10, and as a third model order 48. The first two were set to simulate choosing a model order off from the number estimated automatically for the data used in this study. In the last case we followed an experimental rule of thumb suggested by Greicius et al. (Greicius et al., 2007) according to which you use a number of components equal approximately to one sixth of the time points in the fMRI data. This number was then increased by two because of the simulated sources added to the original fMRI data.

A spatial matching and visual assessment was used to identify same or most similar RSN-related ICs between model orders. Thresholded score maps produced by Melodic mixture modeling (c.f. section "ICA analyses") were correlated with each other using a FSL4 fslcc-tool. Best matches were confirmed by viewing changes in Melodic reports between the cases of automatically estimated model order and additional model orders. The maps for each model order were chosen as described in section "Inspection of selected RSN-related ICs regarding bootstrapping".

Results

On simulated effects

Dimensionality estimation by Melodic on original fMRI data yielded a model order of 48. Adding two simulated effects to the real fMRI data increased the estimated model order to 51. Skewness' of the added data components were −1.8297 for the low variance source and 2.2707 for the high variance source. Fig. 1 shows the histograms of voxel values for non-zero voxels and Fig. 2 demonstrates the spatial spread of randomized voxel locations for both sources at a single slice. In addition to 48, the control experiments were carried out at sICA model orders 41 and 61 due to 51 components being estimated by PPCA (c.f. section "Assessing the effects of model order).

Similarity between simulated effects and results of different sICA approaches

Fig. 3 shows how different sICA approaches performed when estimating mixing vectors (i.e. time courses) of ICs relating to simulated effects. Solid lines and markers are mean values over 25 repetitions and dotted lines represent 99% corresponding confidence intervals for the actual distribution mean. For both low and high variance simulated signal sources there is no difference between non-bootstrapping ICASSOs and original FastICA. Melodic FastICA, however, has slightly poorer performance. In the bootstrapping case the pseudoinverse utilizing original ICASSO performs in a considerably poorer manner compared to modified ICASSO, which uses centrotype ICs for determination of the final mixing vectors.

In all approaches, excluding the original ICASSO in the bootstrapping case, convergence threshold does not have any significant effect on results. Modified ICASSO using bootstrapping gives best results compared to all other approaches. Considering simulated effects with high variance, all differences (between Melodic and other approaches, between ICASSOs in the bootstrapped case and between modified ICASSOs in the non-bootstrapped and bootstrapped cases) are clearly smaller. Confidence intervals are very narrow considering all results except for the bootstrapped original ICASSO.

Histograms of condition numbers for unmixing matrices estimated in the ICASSO approaches can be seen in Fig. 4. The results show that condition numbers in the bootstrapping case are orders of magnitude larger than in the case of no bootstrapping. These results correspond with differences in Fig. 3 concerning the performance of original ICASSO between non-bootstrapping and bootstrapping. As condition numbers are large, they correlate with large numerical errors in pseudoinversion based mixing estimates and thus with performance differences in Fig. 3 concerning original ICASSO. Also, a more pronounced presence of outliers (some excluded from the histogram

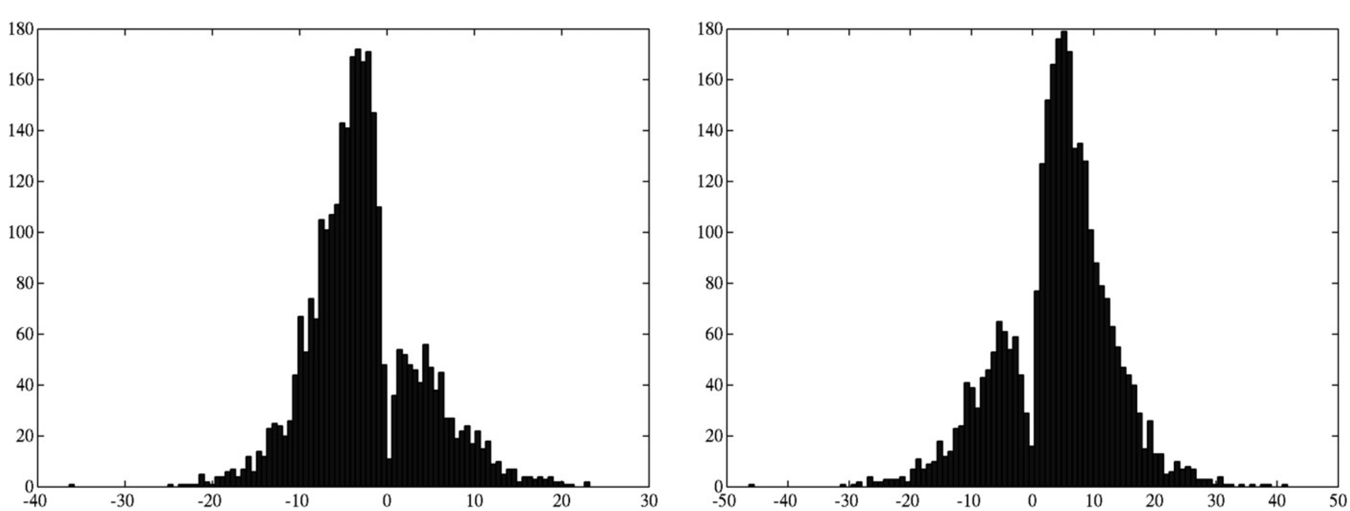

Fig. 1. Histograms of simulated effects values (left: the artificial signal source contributing to 1% of all variance, right: the signal source contributing to 2%).

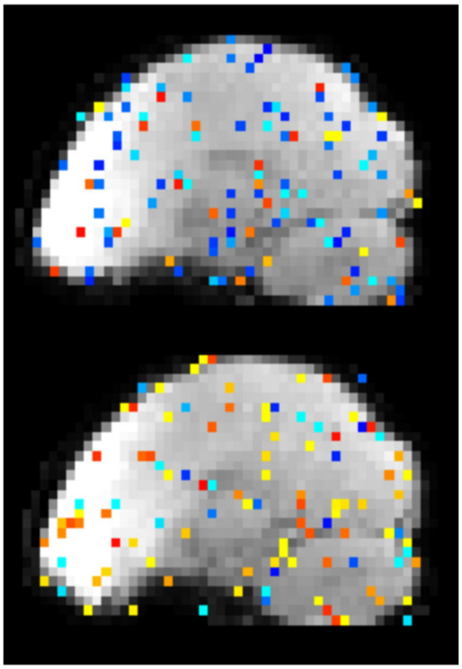

Fig. 2. Example sagittal slice showing the artificial signal sources (top: low variance source, bottom: high variance source) overlaid on the mean image of the BOLD data analyzed.

due to scale issues) in the bootstrapping case correlate with wide confidence intervals in Fig. 3 in the bootstrapping case.

Effects of bootstrapping on temporal characteristics of IC estimates

Identified RSN-related components included executive function related IC, medial and lateral visual cortex IC, and posterior and anterior "default-mode network" ICs (later EXEC, VIS, DMNP and DMNA, respectively). The thresholded spatial maps of EXEC, VIS, DMNP and DMNA can be seen top-down in Fig. 5. According to MNI coordinates and FSL4 atlases (Harvard-Oxford and Juelich) the EXEC component was dominated by activity in the frontal pole, VIS included large parts from V1 and V2 and some part of V5, DMNA consisted mainly of the frontal pole, paracingulate gyrus, superior frontal gyrus and anterior cingulate gyrus, and DMNP included mainly the precuneous cortex and posterior division of the cingulate gyrus with also some part of the anterior division, angular gyrus and right auditory cortex.

Mean power spectra of 25 IC time course estimates of EXEC, VIS, DMNP and DMNA components can be seen Figs 6, S1, S2 and 7 respectively. The power is displayed as percentage relative to all power in the spectra. The extra axis in the figures gives the spectra as a function of the FastICA convergence threshold level. For EXEC and VIS, using bootstrapping moderately reduces power on frequencies higher than 0.1 Hz, for DMNP and especially DMNA the reduction is more apparent. In all cases, the subsets of low frequencies (below 0.1 Hz) become more distinctive as bootstrapping is deployed. When using bootstrapping the convergence threshold has some effect on power frequency-wise. This effect is more profound concerning DMNP and DMNA results.

Variability of all ICs between different sICA approaches

The similarity of Melodic FastICA and original FastICA estimates of all 51 IC mixing vectors to modified ICASSO estimates in the non-bootstrapped case can be seen in Fig. 8. Solid lines show the sample mean and dotted lines give the 99% confidence intervals as in the case of results concerning simulated sources. Melodic estimates are less alike than estimates of original FastICA but most of them are still very close to the reference. Correspondingly, the similarity of original ICASSO to modified ICASSO in non-bootstrapping and bootstrapped cases can be seen in Fig. 9. In the bootstrapping case a systematic difference is apparent even though the difference is very much dependent on component.

Fig. 10 shows how much all or only RSN-related IC estimates differ between non-bootstrapping and bootstrapping cases. The 4 RSNs top-down are EXEC, VIS, DMNP and DMNA. The RSN-related IC estimates are all over 0.7 but both mean and expected variability regarding widths of confidence intervals are very much IC dependent. The IC estimates that are less sensitive to bootstrapping have also smaller confidence intervals. The mutual order of similarity values regarding RSN-related ICs corresponds with the amount of corresponding differences in PSD results in Figs 6, S1, S2 and 7.

As in the case of simulated effects, convergence threshold does not have a dominant effect on results in any approach.

The effects of model order

Concerning simulated sources there is no difference between original FastICA and modified ICASSO, and Melodic FastICA gives a little lower accuracy compared to them. The accuracy of original ICASSO in bootstrapping is considerably lower compared to all other results. Modified ICASSO with bootstrapping is the best option most of the time. Excluding original ICASSO, the convergence threshold has no significant effect on results. The SNR affects the similarity values and their differences between the compared methods.

Concerning all ICs, Melodic FastICA is less similar to ICASSO reference than original FastICA. The difference between original and modified ICASSO are much more pronounced in the bootstrapping case. As a whole, the differences in similarity due to bootstrapping in the case of modified ICASSO remain similar and IC-specific regarding RSN-related ICs. In most comparisons the similarity does not change dominantly as a function of converge threshold.

Figs S3, S4 and S5 show results similar to Fig 3 for model orders 41, 48 and 61 respectively. Similarly concerning all ICs, Figs S6, S7 and S8 correspond to Fig 8. Figs S9, S10 and S11 correspond to Fig 9, while Figs S12, S13 and S14 correspond to Fig 10.

The differences to results obtained with model order 51 are as follows. Concerning the low SNR simulated source, the lower model orders (41 and 48, Figs S3 and S4) yield decreased similarity values concerning non-bootstrapping in the case of all methods except Melodic FastICA. Melodic FastICA results, on the other hand, improve. The accuracy of original ICASSO estimates both decreases and becomes more variable in the non-bootstrapping setting when using high convergence threshold (>0.00005). Regarding low SNR simulated source and the lowest model order (41), the accuracy also decreases throughout convergence thresholds compared to original FastICA and modified ICASSO and remains mostly at the level of Melodic FastICA. When using high model order (61) the differences in accuracy between best non-bootstrapped and bootstrapped results disappear. Also, the accuracy of all methods increases excluding original ICASSO in the bootstrapping case.

Concerning all ICs there is a general tendency towards similar estimates in the direction of increasing model order when FastICAs are compared with ICASSO reference (Figs S6, S7, 8 and S8). In the non-bootstrapping case there is a decreased similarity for a few ICs when model order is decreased (41, Fig S9) or increased (61, Fig S11) much from the automatically estimated one. As with simulated sources, the similarity values drop at high convergence threshold values (>0.00005) and especially when lower model orders (41 and 48, Figs S9 and S10) are used. The similarity of RSN-related ICs between bootstrapping and non-bootstrapping cases vary differently IC-wise as a function of model order (Figs S12, S13, 10 and S14). As a whole, most similar results between bootstrapping and non-bootstrapping cases are obtained at the automatically estimated model order. The effect of bootstrapping on EXEC IC decreases with increasing model order. VIS IC becomes less similar between the cases if the model order is increased or decreased

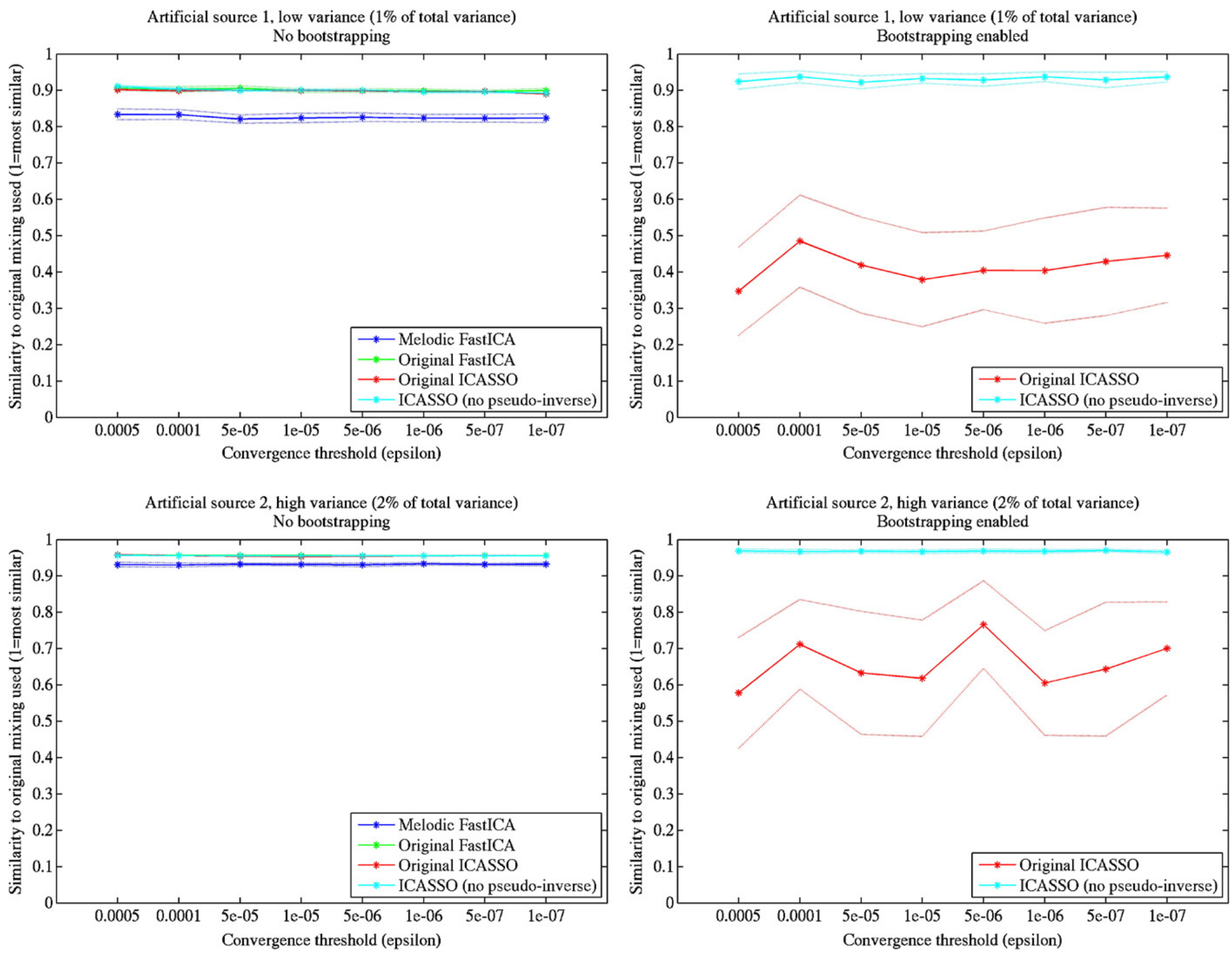

Fig. 3. Similarity of estimates of different approaches to original mixing. Solid lines and markers are mean values over 25 repetitions and dotted lines are confidence intervals for the actual mean of similarity value distributions.

from the automatically estimated order. DMNP IC remains approximately at the same level. The results of DMNA IC change notably when a big change (41/61) is applied and the changes are more pronounced in the case of decreasing model order.

Discussion

The main objective has been to investigate how the use of different repeatability measures affects the estimates and if repeatability measures are necessary in fMRI sICA. We have compared different FastICA-based IC estimation approaches regarding detection accuracy of the simulated effects comparable to non-Gaussian components in real resting-state fMRI data. We have also shown how temporal dynamics (time series/mixing vectors) of RSN-related IC estimates are affected by bootstrapping and how alike different approaches are for all effects in resting-state fMRI data. Comparisons also show the role of the convergence threshold used in FastICA-based sICA of resting-state data. The effect of model order was assessed with control experiments covering several scenarios.

Similarity between simulated effects and results of different sICA approaches

Concerning simulated effects there were no differences in performance between original FastICA and original FastICA -based repeatability approaches (ICASSOs) if bootstrapping was not used. The slightly improved estimates by modified ICASSO in the case of bootstrapping were to be expected as bootstrapping in the presence of noise (estimation of fMRI ICs in signal + noise subspace (Beckmann and Smith, 2004)) should decrease the overfit of estimates to the noise (Ylipaavalniemi and Vigario, 2008). Therefore bootstrapping should be used in order for repeatability measures to have any effect. However, the performance improvement was only moderate and also more profound in a case of low variance signal source, which is also to be expected considering SNR. In conclusion, the use of repeatability measures might not be as warranted as one might think in a case of resting-state fMRI data. fMRI data acquired in the presence of a cognitive stimulus or other challenge can have much higher SNR effects than resting-state data, so this principle should be applicable also in wider sense in sICA analysis of fMRI data.

The lack of differences between FastICA and repeatability analysis results therefore give further evidence for validity of results in plenty of existing sICA studies of both stimulus and resting-state fMRI data which have not used repeatability measures. However, since properties of data can vary, possibly being considerably worse than in the case of this study, there is no harm in using repeatability measures to be on the safe side if reliable tools are available. Also, a similar lack of effect concerning the use of strict convergence thresholds can be observed and a similar conclusion made on validity of existing studies.

Original ICASSO was shown not to work properly in the bootstrapping case with regard to estimated mixing vectors. As both

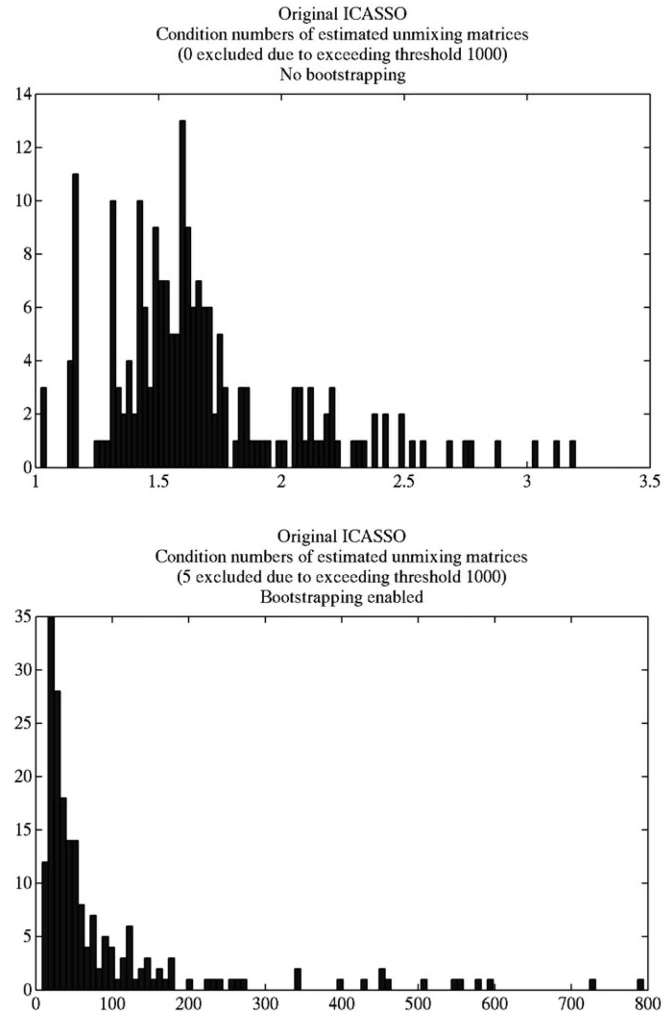

Fig. 4. Condition numbers of unmixing matrices for assessing the effects of bootstrapping on the use of the pseudoinverse in the original ICASSO approach.

original and modified ICASSO use IC estimate cluster centrotypes for final determination of unmixing vectors and because IC values are computed as projections along these unmixing vectors, there is no difference in spatial maps acquired from the two methods. However, since time courses (mixing vectors) of ICs are poor estimates for bootstrapped original ICASSO runs, any neuroscientific inference based on such temporal dynamics is highly unreliable. The confidence intervals for the mean of different runs were also larger than in the case of any other approach, which predicts that possible outcomes of time course estimates from bootstrapped original ICASSO runs can randomly vary greatly depending on the analysis session. When going back to the point of using repeatability methods for possibly greater reliability of sICA results, caution is advised. Our results warrant a proper evaluation of used repeatability measures if accurate mixing vector estimates are preferred. As an example, we wish to highlight that in the case of fMRI data the original ICASSO has not been tested before with as much detail as in this study. Our results complement those of Himberg et al. (Himberg et al., 2004).

The difference between original and modified ICASSO approaches is that the original method uses the pseudoinverse of the unmixing matrix to compute mixing vectors, whereas the modified version determines mixing vectors through clustered IC estimate centrotypes (in the same way both original and modified ICASSO determine unmixing vectors). The pseudoinverse assures that the ICA model remains generative, which almost always is not the case with the modified version because centrotype-based subspaces are not necessarily orthogonal and may present overlapping portions of original data variance. As was seen from the condition numbers of unmixing vectors in Fig. 4, IC centrotype unmixing vectors corresponding to different bootstraps form an ill-conditioned unmixing matrix and the (pseudo)inverse of such a matrix is susceptible to large numerical errors leading to distorted mixing vectors (showing in results in Fig. 3).

Melodic FastICA was found to perform in a slightly poorer manner compared to the other approaches (excluding original ICASSO in bootstrapping case). This is likely due to the data not being mean zero prior to actual IC estimation step as it is the main difference between the FSL4 Melodic FastICA and the original FastICA approaches. The implementations do use different libraries (C++ and Matlab respectively) which could potentially affect the results but similar results were obtained when the mean was left to the data when running repetitions of original FastICA (results not shown here).

As pointed out in section "ICA analyses", there are arguments both against and for removal of the mean from fMRI data prior to sICA. However, non-Gaussian projections can be searched for by using many kinds of optimization contrasts and it is not known that any single contrast available in FastICA or any other ICA framework would lead to ICs that are in some sense spatially optimal presentations of the effects in the data. It is likely that all effects in the fMRI data do not even follow the same assumptions distribution-wise; some might be skewed, others symmetric. This very probably makes any set of ICs made with a single choice of contrast a crude approximation. The simulated effects in this study were explicitly made such that true skewness based contrast should yield the best results. In that sense the performance difference of Melodic FastICA is understandable concerning the simulated effects.

The mean subtraction problem possibly does not manifest as clearly in temporal ICA (such as ICA of electroencephalography (EEG) data) because temporal components might not have as uneven representation over the whole sample as spatial components can have. As temporal ICA in general is a more common use of ICA, this could have had led to the current situation where this problem has not yet been thoroughly investigated.

As a summary concerning simulation results it can be stated that:

1. The repeatability measures or strict convergence thresholds do not necessarily improve sICA results of fMRI data
2. Consequently fMRI sICA results reported in existing literature are probably valid enough, from ICA points of view, even if the repeatability measures or strict convergence thresholds have not been used
3. Some performance increase can be obtained when bootstrapping repeatability measures are used
4. Original ICASSO fails to accurately estimate IC temporal dynamics in the bootstrapping case due to the use of the pseudoinverse
5. The modified ICASSO approach using centrotype-based mixing estimates is proposed instead
6. Any methods, new and old, should be thoroughly tested in the data domain context (e.g. fMRI in general or resting-state fMRI) in which they are intended to be used

Variability of all ICs between different sICA approaches and the effects of bootstrapping

Regarding the comparison of all ICs between different sICA approaches, the results showed that Melodic FastICA estimates in general deviate from ICASSO more than the original FastICA estimates do. The results concerning systematic differences in achieved skewness due to mean subtraction in the case of simulated effects could thus be generalized also to the estimation of ICs related to real effects in resting-state fMRI data. However, most of the 51 components in both original FastICA and Melodic FastICA are the same as those obtained through repeatability measures when bootstrapping is not used.

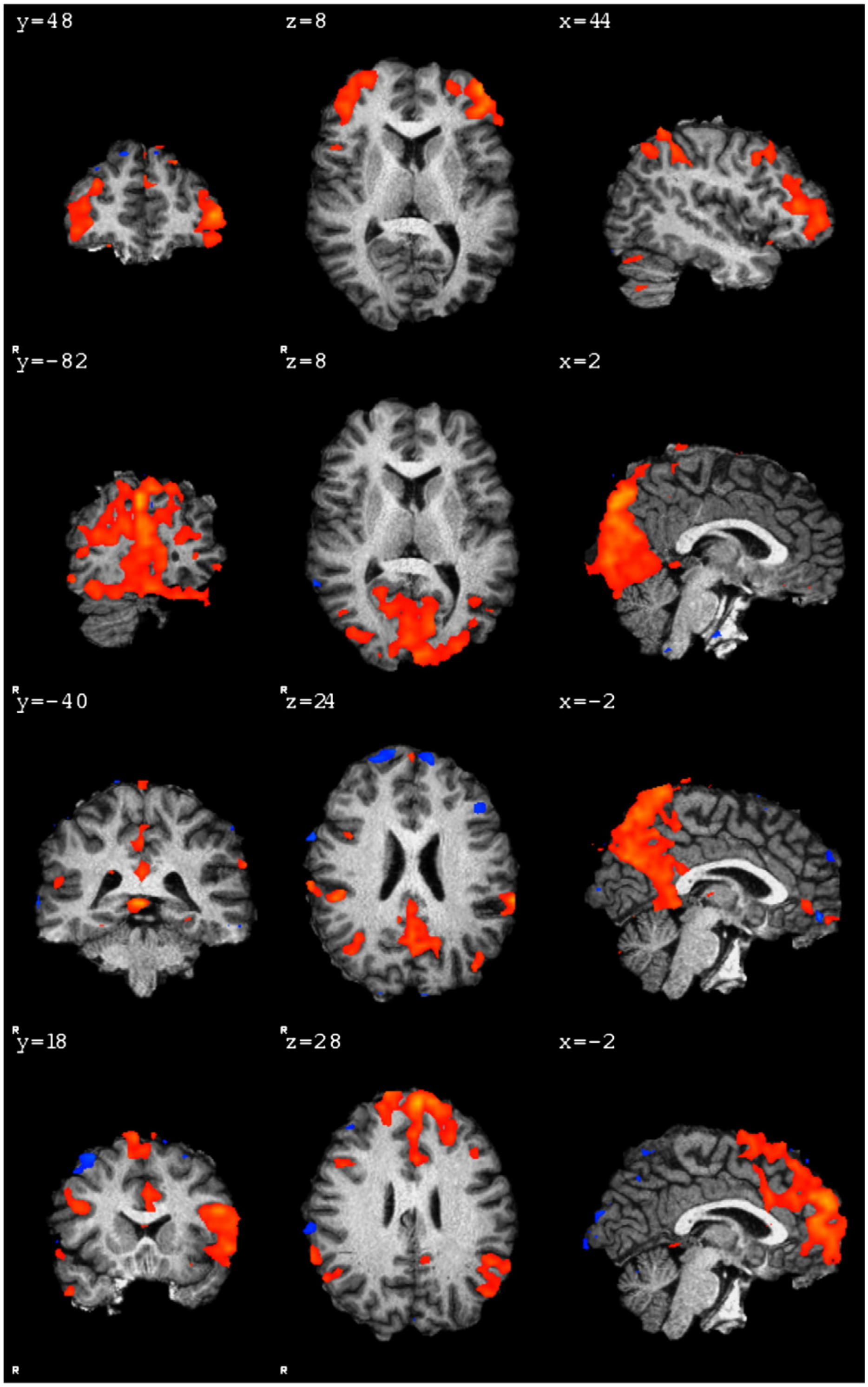

Fig. 5. Identified and selected RSN-related IC maps with statistically significant brain activation (red and blue) overlaid on subject anatomy. Top-down: EXEC, VIS, DMNP and DMNA. Coordinates are millimetres in MNI standard space.

On the other hand, bootstrapping makes ICs in general, as well as ICs related to RSNs, different from the non-bootstrapping case (Fig. 10). These findings combined with results concerning simulated effects further suggest that repeatability measures should be used with bootstrapping enabled in order for them to have an effect. Also, results comparing original ICASSO in non-bootstrapping and bootstrapping cases (Fig. 9) show differences similar to those with simulated effects which means that the use of the pseudoinverse in the computation of mixing vectors fails in general with resting-state fMRI data and not just in simulations.

The RSN-related components, EXEC, VIS, DMNP and DMNA, detected in this study correspond anatomically with components computed in previous studies (e.g. Kiviniemi et al., 2009b, anatomical overlaps). The mean PSDs of their time series also match the criteria for RSN-related effects: main power at frequencies below 0.1 Hz (Raichle and Mintun, 2006). As such, EXEC, VIS, DMNP and DMNA

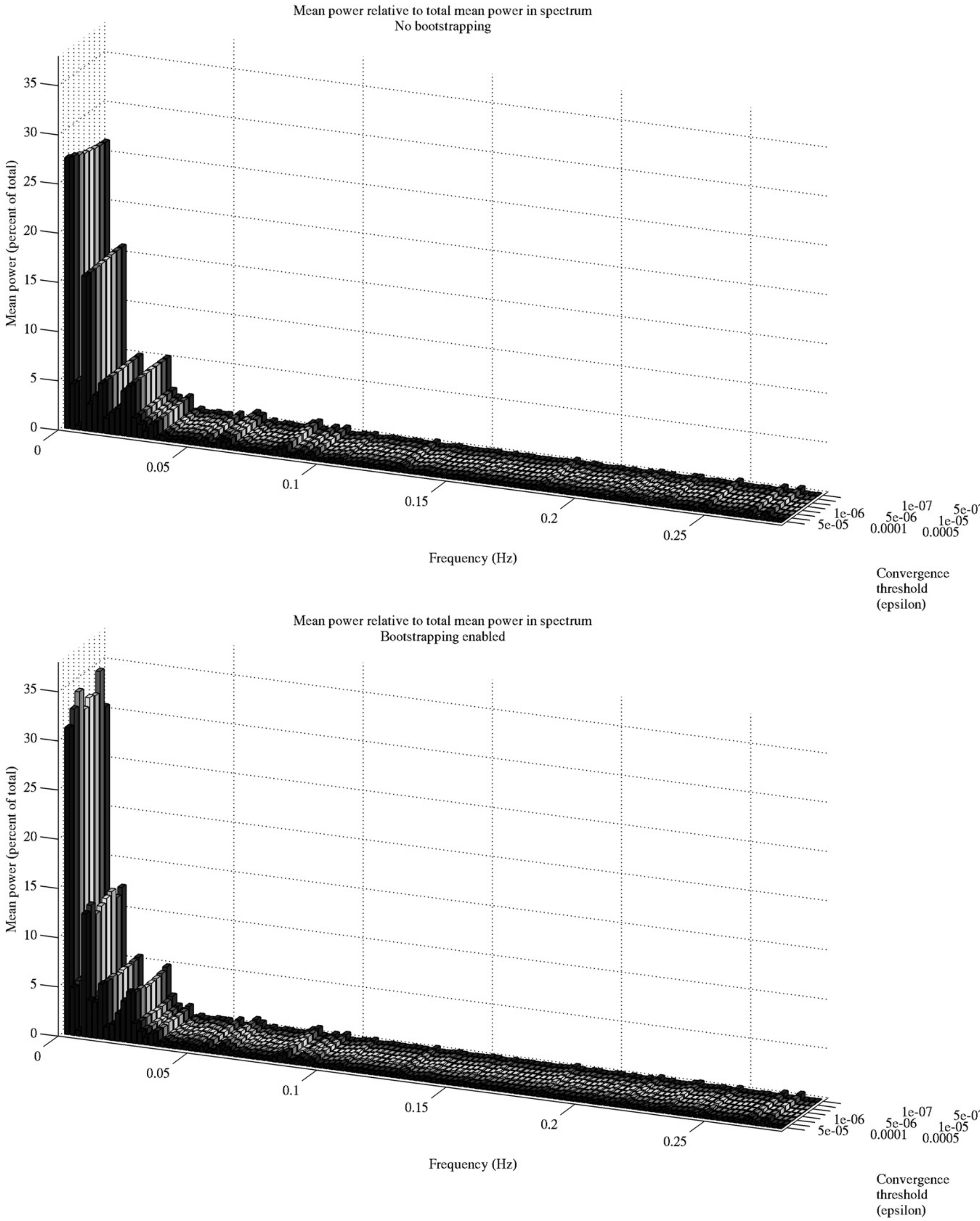

Fig. 6. Mean PSD over 25 repetitions as a function of convergence threshold for EXEC IC.

can be perceived to reflect brain activity and serve as examples of RSNs. These example maps were computed from ICASSO IC estimates using data pre-conditioning and post-processing (mixture modeling of IC value distributions) from the PICA framework. For future reference we denote this combined PICA + ICASSO approach as "PICASSO".

The effect of bootstrapping on the selected ICs was found to be two-fold. Firstly, power was attenuated on frequencies over 0.1 Hz

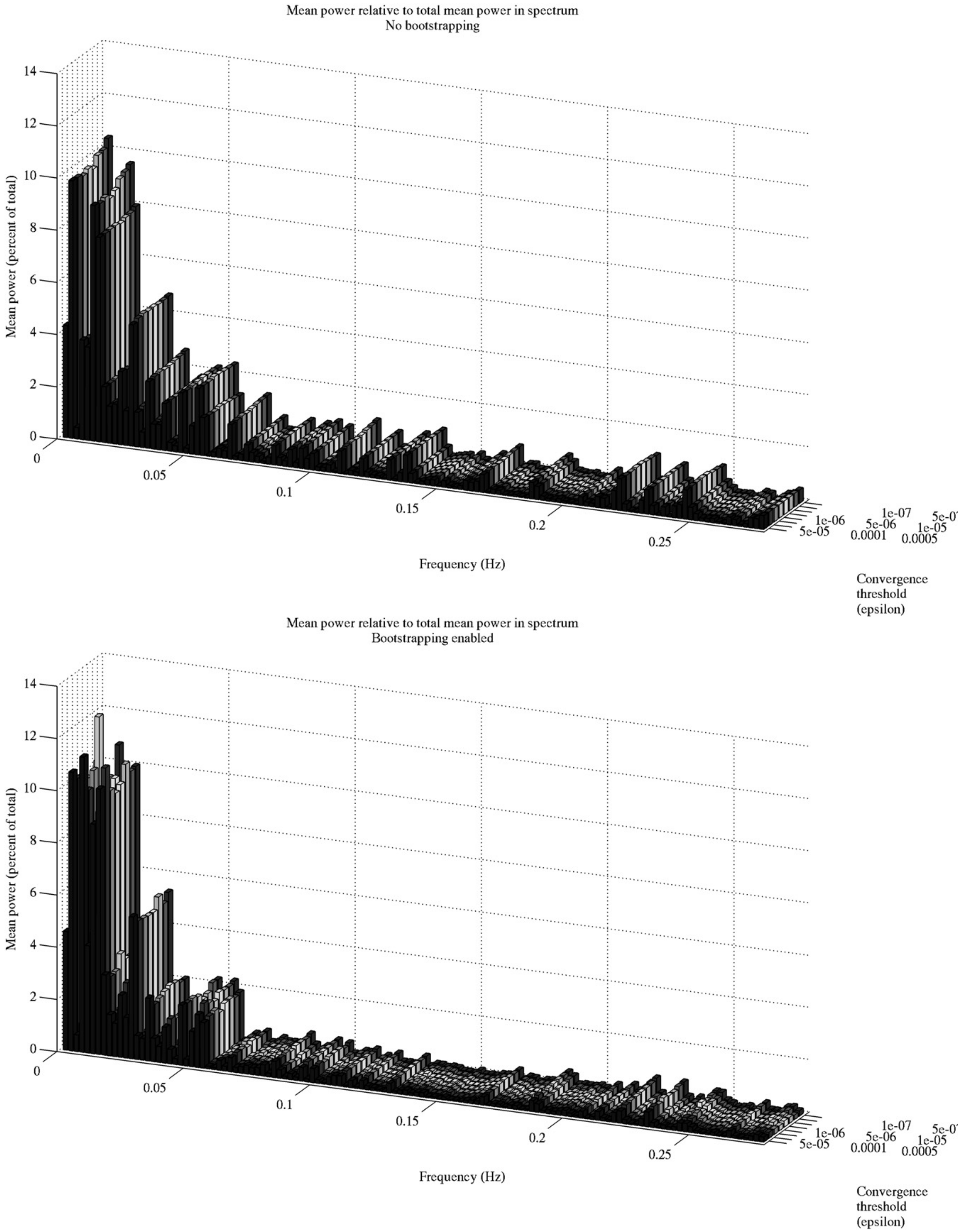

Fig. 7. Mean PSD over 25 repetitions as a function of convergence threshold for DMNA IC.

that are in general viewed as uncharacteristic for RSN-related effects. Secondly power at frequencies below 0.1 Hz became more differentiated and focused on fewer frequency bins which could be interpreted as RSN-wise characteristic frequencies becoming more evident. In this perspective bootstrapping may well improve estimates of RSN-related temporal dynamics and consecutively make neuroscientific inference based on them more reliable. The improvement was very dependent on IC. The amount of change

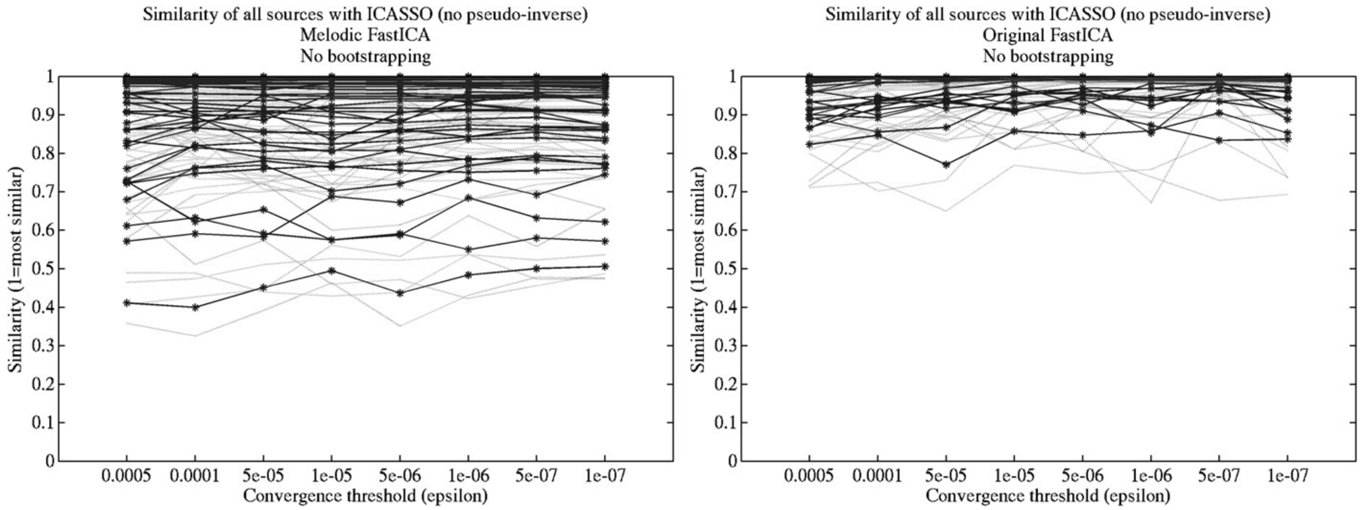

Fig. 8. Similarity between the reference (matching modified ICASSO estimates) and the estimates produced by Melodic FastICA (left) and original FastICA (right). Solid lines and markers are mean values over 25 repetitions and dotted lines are confidence intervals for the actual mean of the similarity value distribution.

naturally corresponded with how much modified ICASSO runs matched between non-bootstrapping and bootstrapping approaches (Fig. 10). In conclusion, the advantage gained from using bootstrapping depends highly on the effects of interest. Again, we see no harm in using some repeatability measures given that the method gives reliable estimates.

Finally, the convergence thresholds do not seem to have any significant role in FastICA-based approaches compared here. Regarding simulated effects, there are only differences within results of bootstrapping original ICASSO which performs poorly in any case. Also, in comparison of all ICs between different sICA approaches, mean similarities and corresponding confidence intervals of single ICs are quite constant as a function of the convergence threshold. In addition to previous conclusions about the little need for repeatability analysis in sICA of fMRI data, these findings also give further merit to the accuracy of existing sICA fMRI studies. In the existing studies high value convergence thresholds may be in place as default parameters of ICA implementations are used. For example, Melodic and original FastICA packages and consequently also ICASSOs based on original FastICA use a default convergence threshold of 0.0005.

Most systematic variability of similarity values for many ICs as a function of convergence threshold can be seen in the bootstrapped case (Fig. 10). By using RSN-related ICs one can qualitatively assess (Figs. 6, S1, S2 and 7) that this variability is coupled with the degree of effect bootstrapping has on the IC. A possible explanation is just that, in the repetitions of bootstrapping modified ICASSO runs at certain convergence threshold level, an outlier estimate (with regard to estimates at other convergence threshold levels) may occur repeatedly over all 25 runs. Such an occurrence would make mean similarity values, confidence intervals and mean power at frequency bins a little different from other convergence threshold levels. In conclusion, the use of bootstrapping can have this kind of trade-off in parameter-wise stability of estimation and several converge threshold levels should possibly be explored in bootstrap analyses for optimal results.

As a summary concerning all data components it can be stated that

1. Bootstrapping gives different results also in the estimation of real effects from fMRI data
2. The use of the pseudoinverse (original ICASSO) gives deviant results also in the estimation of real effects from fMRI data
3. Bootstrapping makes temporal dynamics of resting-state activity related ICs less noisy and more distinctive
4. Detections of the networks benefit from bootstrapping very variably

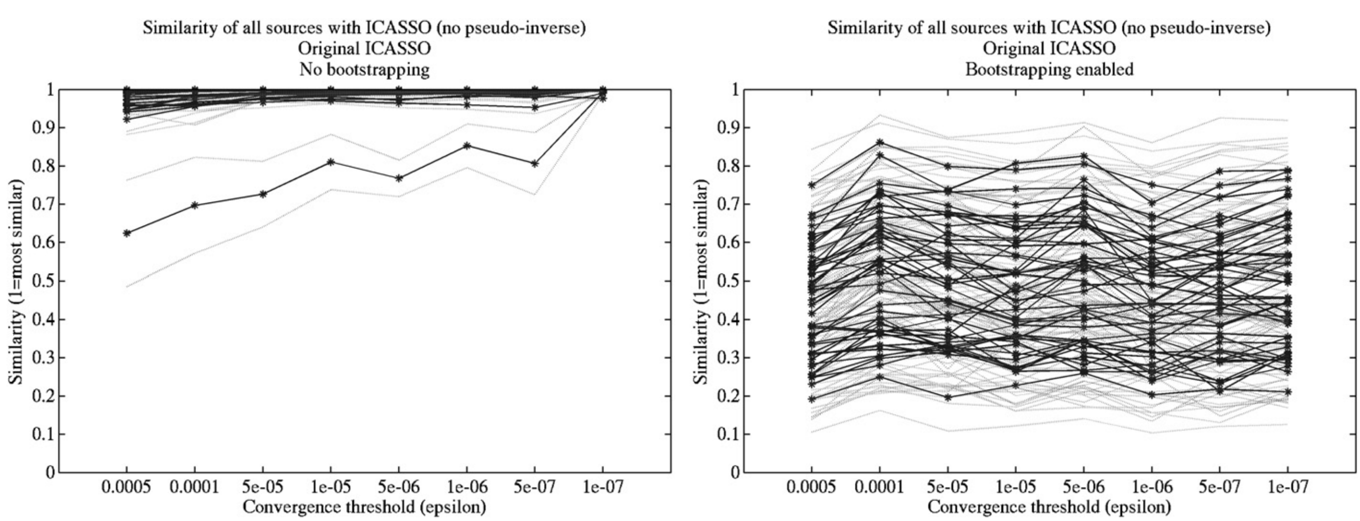

Fig. 9. The similarity of original ICASSO produced estimates to the reference (matching modified ICASSO estimates) in non-bootstrapping (left) and bootstrapping (right) cases. Solid lines and markers are mean values over 25 repetitions and dotted lines are confidence intervals for the actual mean of the similarity value distribution.

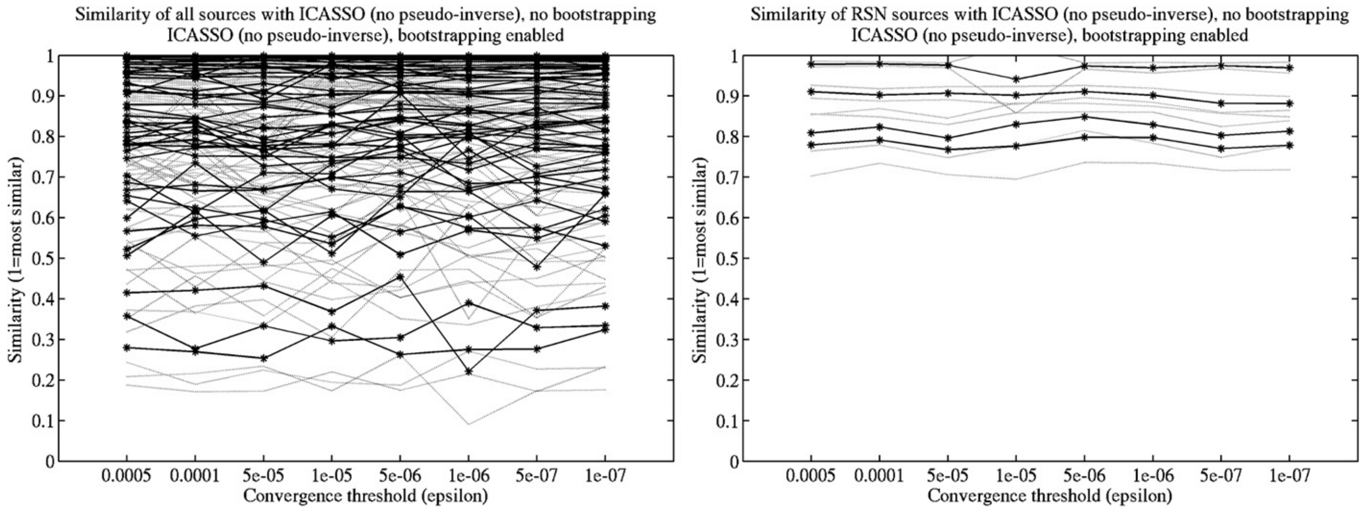

Fig. 10. The similarity of modified ICASSO produced estimates between non-bootstrapping and bootstrapping cases for all non-Gaussian components of the data (left) and the RSN-related ICs (right, top-down: EXEC, VIS, DMNP and DMNA). Solid lines and markers are mean values over 25 repetitions and dotted lines show confidence intervals for the actual mean of the similarity value distribution.

5. Bootstrapping should only be used with properly validated methods
6. When using bootstrapping the convergence threshold has effect regarding resting-state activity estimates and multiple thresholds may have to be explored for optimum results

The effects of model order

Overall the changes in results due to changes in model order were only moderate or small. The moderate amount of change makes the conclusions in sections "Similarity between simulated effects and results of different sICA approaches" and "Variability of all ICs between different sICA approaches and the effects of bootstrapping" largely applicable to different sICA model order selection scenarios. The main addition to section "Similarity between simulated effects and results of different sICA approaches". Is that, in addition to bootstrapping, also increasing the model order can make estimates slightly improved and even remove the differences due to the use of bootstrapping.

On the other hand, the differences in results concerning RSN-related ICs suggest that for real fMRI effects, the situation can be very variable. Thus bootstrapping cannot be dismissed as a contributing and possibly important factor. DMNP IC estimates remained at the same level with respect to bootstrapping whereas e.g. VIS IC became more variant away from automatically estimated order 51. One explanation for these differences could be that splitting (and merging) of components occur depending on model order used. For example previous results (Abou-Elseoud et al., 2010) suggest that VIS is probably more prone to splitting than DMNP also within this study.

EXEC IC was found to become more invariant to bootstrapping with increasing model order. This could be in line with improving accuracy in the case of simulated sources where differences between non-bootstrapping and bootstrapping cases disappear. Assuming that correspondence, the results could further suggest that (among EXEC, VIS, DMNP and DMNA estimated in this study) EXEC is the data component which most truly acts as one would expect from an actual independent component. That, on the other hand, raises a question about the nature of 1-dimensional IC estimates computed from resting-state fMRI data. The discrepancy that we observed between RSN-related ICs could be related to the existence of multidimensional independent subspaces (Cardoso, 1998; Ylipaavalniemi and Vigario, 2008) in the data instead of all IC estimates representing 1-dimensional signal sources that are actually independent. The RSN-related ICs detected in this study might be just forced 1D-projections of such subspaces with EXEC more likely representing an actual 1-dimensional signal source compared to other RSN-related ICs.

The issue of IC-specific considerations in choosing model order is provided here merely as a proof of concept without aims at population level inferences. Anatomical analyses of ICs involved here are beyond the scope of this study and would require group data for accurate generalizations.

Concerning convergence thresholds, the use of high values was clearly found to affect the results when original ICASSO is used at model orders lower than the automatically estimated order. This should be considered when using original ICASSO in such settings.

Future prospects for fMRI sICA method development

When increased reliability of ICA estimates is preferred, repeatability measures are not the only option. Another, interesting possibility is to use improved versions of FastICA. Zarzoso and Comon (Zarzoso and Comon, 2008) have pointed out that the basic FastICA algorithm is suboptimal and can be made to evade local optima (and also made faster) with relatively simple changes to its optimization routine.

On the other hand, the ICA model itself could be too restrictive for fMRI. sICA results of fMRI data may be basically neuroscientifically convenient and informative sparse representations of the data rather than actual statistically independent signal sources (Ylipaavalniemi and Vigario, 2008). Given these considerations it is a good question whether strict ICA model assumptions should be applied in any case or if less restrictive projection methods should be applied instead. A recent study (Daubechies et al., 2009) has shown both through simulations and stimulus fMRI experiments that sparseness (instead of statistical independence) is very likely the explaining factor behind meaningful fMRI sICA decompositions. Furthermore, regarding the possibility of multidimensional independent subspaces instead of 1-dimensional ICs (c.f. Discussion section "The effects of model order"), interpretations made on corresponding 1-dimensional estimates (sparse or otherwise) may well give an incomplete picture of underlying brain activity. Thus fusion of interrelated 1D-estimates into independent subspaces (Cardoso, 1998) that are more representative of effects in the data, is possibly also an important model-related consideration, especially when high model orders are used and splitting of ICs is more pronounced (Abou-Elseoud et al., 2010).

The mean subtraction problem (c.f. section "ICA analyses" and Discussion section "Similarity between simulated effects and results of different sICA approaches" for discussion) is yet another point about whether it is really necessary to follow the normal ICA model in pursuit of interesting non-Gaussian projections of the fMRI data. In this sense the variant of FastICA used in Melodic is a trade-off between data domain specific concerns about results representing fMRI effects of interest and the accuracy of mathematical formulations to achieve as skewed distribution of IC values as possible. One way for developing further the detection of interesting non-Gaussian projections of fMRI data could be the use of more sophisticated optimization contrasts. For example, already in 2004 (Karvanen and Theis, 2004) Pearson ICA (Karvanen and Koivunen, 2002) has been applied to fMRI data. In Pearson ICA the optimization contrasts are fitted to the data in an iterative manner jointly with maximization of non-Gaussianity, enabling effect specific modeling of non-Gaussianity in IC estimation. This could be combined with the aforementioned optimization related improvements reported for FastICA (Zarzoso and Comon, 2008).

Also, in the work of Ylipaavalniemi and Vigario (Ylipaavalniemi and Vigario, 2008) the repeatability framework philosophy has been developed towards detecting interesting subspaces not necessarily mutually orthogonal as in the strict ICA model. This feature enabling exploratory analysis of fMRI data might be the advantage of repeatability approaches even if they are not needed for stability.

Limitations and other considerations on methods used in this study

The steps in this study were based on existing theoretical formulations and results in order to minimize the effects of less established methodology on comparisons. FastICA was selected as a basis for ICA as it is in general a widely used ICA method, including sICA for fMRI data. Also, Daubechies (Daubechies et al., 2009) has recently reported that if sensitivity to actual statistical independence (instead of just sparseness) is preferred, then FastICA is more proper choice compared to Infomax, which is also widely used in fMRI sICA. The PICA framework was used as it offers fMRI-specific conditioning of data in the pre-processing step and justifiable post-processing (mixture modeling and alternative hypothesis testing) for inferential steps considering produced anatomical maps of significant effects. Consequently, a mixture model was assumed to be an adequate model of resting-state fMRI components for creation of related simulated effects. In addition, the use of Melodic for PICA provided practical means to make all comparisons comparable due to unified pre-processing (including PCA) before actual IC estimation.

Melodic was first used to estimate model order with PPCA. Later addition of two simulated sources increased the model order from 48 to 51, which seems reasonably accurate for such an estimation and 51 was used as model order in all later estimations. Simulated effects were created according to the PICA mixture model (Fig. 1) and scaled to 1% and 2% of overall variance as suggested by another Melodic analysis for the upper and lower boundaries of variance when actual RSN-related resting-state fMRI effects in the same data were inspected. Spatial locations of simulated effects were randomized within brain voxels for generality even though in fMRI the effects often have a spatial cluster structure and as such estimation of all related IC values is affected by similar noise properties within the spatial clusters. In addition, the original mixing, used to add the simulated effects to the data, was also randomly generated for generality and to enable more reliable estimation of ICs related to real effects without simulated effects interfering considerably. As a contrast in the work by Yang (Yang et al., 2008) RAICAR produced estimates have been compared with original FastICA using simulated effects which have both temporal and statistical structure. It is debatable if either approach is more valid than the other.

Overall, regarding our creation parameters for the simulated effects, our results should describe estimation differences between different sICA approaches equally well for all RSN-related ICs. Enabling joint estimation further simplified analysis setups as no separate studies on simulated effects and real effects were required. This made the evaluations computationally more feasible in terms of disk space and processing time, especially regarding the additional control experiments concerning model order. On the other hand, we cannot account for possible RSN specific estimation factors that might be revealed if temporally and spatially structured simulated effects simulating some specific RSNs were used. Still, all the effects presenting in fMRI data are not yet known explicitly. Thus it would be very challenging to make simulations that are spatially representative of all component possibilities anatomy-wise. A large number of permutations with different subsets of brain regions would be required to enable good generalizations. Such complexity would make the evaluations ensuing the simulations computationally intractable.

We have also not studied spatial properties of IC estimates. The analysis has been limited to mixing vectors for two reasons. Firstly, mixing and unmixing estimates are commonly used for comparing estimates in general ICA literature, cf. e.g. Amari, Douglas and Ollila (Amari et al., 1996; Douglas, 2007; Ollila, 2009). The spatial maps are just projections of original data into subspaces spanned by unmixing vectors. Secondly, as spatial study would introduce additional methodological steps including mixture modeling of all IC estimates in every run we think that current study extent using mixing vectors is adequate in showing differences in estimation due to use of repeatability measures. Another apparent limit of our study is that we have made analyses using only single subject fMRI data. This was because making the number of repetitions and data for a similar group analysis would require considerably more time and data storage resources. The significance of using repeatability measures to population level mean results is thus not available, but we feel that the current results give an adequate description of how they perform at the subject-level. Namely the main results concerning simulated effects are independent from chosen resting-state fMRI data set and the data set specific results only support main findings derived for simulated effects.

Concerning contrast function we have limited this study to skewness for reasons outlined in section "ICA analyses". Because of this our finding do not directly address the same issues when using e.g. kurtosis which is also a widely used contrast. Nevertheless, since both contrasts still optimize for non-Gaussianity, we think that relative differences between sICA approaches in this study are descriptive also in a general sense. Also, as we pointed out in Discussion section "Similarity between simulated effects and results of different sICA approaches", it is debatable if any single contrast can ever give optimum results.

Our result statistics and evaluations thereof were based on the premise that good enough matches have been found between ICs from different ICA sessions. In the case of simulated sources, the second best matches had considerably lower inner product values compared to the best matches. Considering the high mean similarity values between ICs and quite narrow confidence intervals of all ICs in almost all comparisons, it is rather improbable that matching would work insufficiently. There could be some mismatches between ICs if some sessions have not converged to the same set of components but such situations would show in the results only as lower mean similarity values and broader confidence intervals.

Regarding the effects of model order, only a coarse battery of control experiments were carried out for practical reasons. However, the two model orders (41 and 48) lower than the automatically estimated number of components demonstrates the effects of model order for cases of low order choices made in previous studies. 41 (or 39, c.f. section "Assessing the effects of model order") is in the upper range of model orders used in several prior resting-state fMRI sICA studies (Abou-Elseoud et al., 2010). 48 corresponds to an experimental rule of thumb suggested in the literature (Greicius et al., 2007), c.f. section "Assessing the effects of model order". On the other hand, the case of higher model order (61) demonstrates results that are clearly different from the lower orders. Moreover, even though a linear relationship between the changes in results and a difference from the automatically estimated number of components may be unlikely, 41 and 61 can be considered to represent larger changes of model order from such a dimensionality estimate whereas 48 represents the changes in results with respect to only little deviation from an automatically estimated number of components.

Concerning the generality framework-wise, the results of this study concerning ICASSOs should be to a large extent applicable to other similar repeatability approaches, such as ARABICA (Ylipaavalniemi and Soppela, 2009) and RAICAR (Yang et al., 2008). ARABICA is very similar to ICASSO but allows for more freedom in selecting interesting subspaces through centrotypes of adjustable clusters whereas in ICASSO some fixed model order is assumed and in default settings the same number of clusters are formed even if there are, for example, outliers distorting the clusters. In RAICAR, averaging over a cluster instead of cluster centrotypes is used for computation of final estimates. This is done selectively to exclude cluster member estimates that are too variable. This might have its benefits as well as weaknesses compared to ICASSO and ARABICA but that is not within the scope of this study. We think it is likely that the results of this study would have been evidently a little different due to this factor if RAICAR instead of ICASSO had been used. For practical reasons concerning study extent, such comparisons could be the topic of a future study.

Finally, an important point is that we have studied the effects of repeatability just from the view point of improving IC estimation. Whatever the results in that regard, the frameworks can still be useful tools for measuring variability of estimates in different sICA setups as performed in the original studies (Himberg et al., 2004; Ylipaavalniemi and Vigario, 2008; Yang et al., 2008).

Conclusions

In this study we have shown that, in contrast to prior understanding, repeatability measures do not necessarily improve sICA results of fMRI data, or more specifically in the case of resting-state fMRI. Also, the accuracy of results did not vary significantly as a function of convergence threshold used in FastICA approaches. These findings give new evidence to support the validity of existing sICA fMRI studies and neuroscientific inferences based on their results. For improved results, bootstrapping should be employed in repeatability analyses but the observed problems with bootstrapping warrant for additional testing of the repeatability measures with different kinds of fMRI data. For ICASSO the bootstrapping problem was corrected by using cluster centrotype mixing vectors instead of the pseudoinversion of a matrix containing centrotype unmixing vectors.

The effect of bootstrapping on estimated mixing was evaluated further for ICs attributed to brain activity in executive function, the visual cortex and default-mode resting-state networks. Power was attenuated at frequencies over 0.1 Hz perceived as uncharacteristic to RSN-related effects and low frequencies (below 0.1 Hz) became component-wise more specific. The improvement due to bootstrapping, however, was very IC specific with executive network related time course estimates benefiting very little and anterior default-mode network related estimates benefited the most. Furthermore, the model order should be considered when applying bootstrapping. Our results suggest that, as a whole, RSN-related ICs, that are most invariant with respect to bootstrapping, are obtained at model orders based on automatic dimensionality estimation although the changes in results due to changes in model order were only moderate or small. Low model order also decreases estimation accuracy in the case of original ICASSO when loose convergence threshold is used concurrently.

To our knowledge this is first study to compare several FastICA-based approaches (conventional and repeatability measures) in analysis of resting-state fMRI data and also in general one of few studies to make in-depth, quantitative comparisons between different data-driven methods in the resting-state fMRI domain. We think that more such comparisons should be done in each subdomain of fMRI (e.g. stimulus studies, resting-state studies, physiological challenge based studies of BOLD signals) to learn more about how and what different data-driven analysis tools should be used for optimal results and what are the dangers of misuse on neuroscientific inference based on those results. Also, e.g. the issue with the mean subtraction and our results concerning the mean subtraction raise the question whether the conventional ICA model should be fundamentally developed further for the specific needs in fMRI, even given the valuable results obtained so far. Many improvement possibilities for conventional ICA-based multivariate analyses exist in the literature. We suggest that interesting non-Gaussian projections should be in future based on less restrictive models than conventional ICA or related estimation approaches.

Acknowledgments

This study was supported by Academy of Finland Grant # 111711, Finnish Medical Foundation and Finnish Neurological Association grants. We thank Nick Hayward, MSc from the A. I. Virtanen Institute of Molecular Sciences, University of Kuopio, Finland, for linguistic editing of the manuscript.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.04.268.

References

Abou-Elseoud, A., Starck, T., Remes, J., Nikkinen, J., Tervonen, O., Kiviniemi, V., Jan 2010. The effect of model order selection in group pica. Hum. Brain Mapp. (Electronic publication ahead of print) URL http://dx.doi.org/10.1002/hbm.20929.

Amari, S.-I., Cichocki, A., Yang, H., 1996. A new learning algorithm for blind source separation. Adv. in Neural Inf. Process. Syst. 8, 757–763.

Beckmann, C.F., Smith, S.M., Feb 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23 (2), 137–152 URL http://dx.doi.org/10.1109/TMI.2003.822821.

Calhoun, V.D., Liu, J., Adali, T., Mar 2009. A review of group ica for fmri data and ica for joint inference of imaging, genetic, and erp data. Neuroimage 45 (1 Suppl), S163–S172 URL http://dx.doi.org/10.1016/j.neuroimage.2008.10.057.

Cardoso, J.-F., 1998. Multidimensional independent component analysis. Proc. Int. Workshop on Higher-Order Stat, In. pp. 111–120.

Daubechies, I., Roussos, E., Takerkart, S., Benharrosh, M., Golden, C., D'Ardenne, K., Richter, W., Cohen, J.D., Haxby, J., Jun 2009. Independent component analysis for brain fmri does not select for independence. Proc Natl Acad Sci U S A 106 (26), 10415–10422 URL http://dx.doi.org/10.1073/pnas.0903525106.

Demirci, O., Stevens, M.C., Andreasen, N.C., Michael, A., Liu, J., White, T., Pearlson, G.D., Clark, V.P., Calhoun, V.D., Jun 2009. Investigation of relationships between fmri brain networks in the spectral domain using ica and granger causality reveals distinct differences between schizophrenia patients and healthy controls. Neuroimage 46 (2), 419–431 URL http://dx.doi.org/10.1016/j.neuroimage.2009.02.014.

Douglas, S.C., 2007. Fixed-point algorithms for the blind separation of arbitrary complex-valued non-gaussian signal mixtures. EURASIP J. Appl. Signal Process. 2007 (1), 83.

Esposito, F., Aragri, A., Latorre, V., Popolizio, T., Scarabino, T., Cirillo, S., Marciano, E., Tedeschi, G., Salle, F.D., Mar 2009. Does the default-mode functional connectivity of the brain correlate with working-memory performances? Arch. Ital. Biol. 147 (1–2), 11–20.

Filippini, N., MacIntosh, B.J., Hough, M.G., Goodwin, G.M., Frisoni, G.B., Smith, S.M., Matthews, P.M., Beckmann, C.F., Mackay, C.E., Apr 2009. Distinct patterns of brain activity in young carriers of the apoe-epsilon4 allele. Proc Natl Acad Sci U S A 106 (17), 7209–7214 URL http://dx.doi.org/10.1073/pnas.0811879106.

Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E., Jun 2009. The global signal and observed anticorrelated resting state brain networks. J. Neurophysiol. 101 (6), 3270–3283 URL http://dx.doi.org/10.1152/jn.90777.2008.

Greicius, M.D., Flores, B.H., Menon, V., Glover, G.H., Solvason, H.B., Kenna, H., Reiss, A.L., Schatzberg, A.F., Sep 2007. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol. Psychiatry 62 (5), 429–437 URL http://dx.doi.org/10.1016/j.biopsych.2006.09.020.

Himberg, J., Hyvärinen, A., Esposito, F., Jul 2004. Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage 22 (3), 1214–1222 URL http://dx.doi.org/10.1016/j.neuroimage.2004.03.027.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. 10 (3), 626–634 URL http://dx.doi.org/10.1109/72.761722.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., Oct 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841.

Karvanen, J., Koivunen, V., 2002. Blind separation methods based on pearson system and its extensions. Signal Process. 82 (4), 663–673.

Karvanen, J., Theis, F., 2004. Spatial ICA of fMRI data in time windows. Proc. MaxEnt 2004. Vol. 735 of AIP conference proceedings. Garching, Germany, pp. 312–319. URL http://www.biologie.uni-regensburg.de/Biophysik/Theis/publications/karvanen04windowfmri_MaxEnt04.pdf.

Kim, D.I., Mathalon, D.H., Ford, J.M., Mannell, M., Turner, J.A., Brown, G.G., Belger, A., Gollub, R., Lauriello, J., Wible, C., O'Leary, D., Lim, K., Toga, A., Potkin, S.G., Birn, F., Calhoun, V.D., Jan 2009b. Auditory oddball deficits in schizophrenia: an independent component analysis of the fmri multisite function birn study. Schizophr. Bull. 35 (1), 67–81 URL http://dx.doi.org/10.1093/schbul/sbn133.

Kim, D.I., Manoach, D.S., Mathalon, D.H., Turner, J.A., Mannell, M., Brown, G.G., Ford, J.M., Gollub, R.L., White, T., Wible, C., Belger, A., Bockholt, H.J., Clark, V.P., Lauriello, J., O'Leary, D., Mueller, B.A., Lim, K.O., Andreasen, N., Potkin, S.G., Calhoun, V.D., May 2009a. Dysregulation of working memory and default-mode networks in schizophrenia using independent component analysis, an fbirn and mcic study. Hum. Brain Mapp. 30 (11), 3795–3811 URL http://dx.doi.org/10.1002/hbm.20807.

Kiviniemi, V., Remes, J., Starck, T., Nikkinen, J., Haapea, M., Silven, O., Tervonen, O., 2009a. Mapping transient hyperventilation induced alterations with estimates of the multi-scale dynamics of bold signal. Front Neuroinformatics 3, 18 URL http://dx.doi.org/10.3389/neuro.11.018.2009.

Kiviniemi, V., Starck, T., Remes, J., Long, X., Nikkinen, J., Haapea, M., Veijola, J., Moilanen, I., Isohanni, M., Zang, Y.-F., Tervonen, O., Jun 2009b. Functional segmentation of the brain cortex using high model order group pica. Hum. Brain Mapp. 30 (12), 3865–3886 URL http://dx.doi.org/10.1002/hbm.20813.

Kokkonen, S.-M., Nikkinen, J., Remes, J., Kantola, J., Starck, T., Haapea, M., Tuominen, J., Tervonen, O., Kiviniemi, V., Jul 2009. Preoperative localization of the sensorimotor area using independent component analysis of resting-state fmri. Magn. Reson. Imaging 27 (6), 733–740 URL http://dx.doi.org/10.1016/j.mri.2008.11.002.

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 (3), 160–188.

Meda, S.A., Calhoun, V.D., Astur, R.S., Turner, B.M., Ruopp, K., Pearlson, G.D., Apr 2009. Alcohol dose effects on brain circuits during simulated driving: an fmri study. Hum. Brain Mapp. 30 (4), 1257–1270 URL http://dx.doi.org/10.1002/hbm.20591.

Menz, M.M., McNamara, A., Klemen, J., Binkofski, F., Apr 2009. Dissociating networks of imitation. Hum. Brain Mapp. 30 (10), 3339–3350 URL http://dx.doi.org/10.1002/hbm.20756.

Mohammadi, B., Kollewe, K., Samii, A., Krampfl, K., Dengler, R., Münte, T.F., May 2009. Changes of resting state brain networks in amyotrophic lateral sclerosis. Exp. Neurol. 217 (1), 147–153 URL http://dx.doi.org/10.1016/j.expneurol.2009.01.025.

Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A., Feb 2009. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44 (3), 893–905 URL http://dx.doi.org/10.1016/j.neuroimage.2008.09.036.

Ollila, E., Aug. 2009. On the robustness of the deflation-based fastica estimator. Proc. IEEE/SP 15th Workshop on Statistical Signal Processing SSP '09, pp. 673–676.

Raichle, M.E., Mintun, M.A., 2006. Brain work and brain imaging. Annu. Rev. Neurosci. 29, 449–476.

Rombouts, S.A.R.B., Damoiseaux, J.S., Goekoop, R., Barkhof, F., Scheltens, P., Smith, S.M., Beckmann, C.F., Jan 2009. Model-free group analysis shows altered bold fmri networks in dementia. Hum. Brain Mapp. 30 (1), 256–266 URL http://dx.doi.org/10.1002/hbm.20505.

Smith, S.M., Nov 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155 URL http://dx.doi.org/10.1002/hbm.10062.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., Aug 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106 (31), 13040–13045 URL http://dx.doi.org/10.1073/pnas.0905267106.

Stone, J.V., Porrill, J., Porter, N.R., Wilkinson, I.D., Feb 2002. Spatiotemporal independent component analysis of event-related fmri data using skewed probability density functions. Neuroimage 15 (2), 407–421 URL http://dx.doi.org/10.1006/nimg.2001.0986.

Suzuki, K., Kiryu, T., Nakada, T., Jan 2002. Fast and precise independent component analysis for high field fmri time series tailored using prior information on spatiotemporal structure. Hum. Brain Mapp. 15 (1), 54–66.

van de Ven, V., Esposito, F., Christoffels, I.K., Oct 2009. Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ica study. Neuroimage 47 (4), 1982–1991 URL http://dx.doi.org/10.1016/j.neuroimage.2009.05.057.

Wu, X., Lu, J., Chen, K., Long, Z., Wang, X., Shu, H., Li, K., Liu, Y., Yao, L., May 2009. Multiple neural networks supporting a semantic task: an fmri study using independent component analysis. Neuroimage 45 (4), 1347–1358 URL http://dx.doi.org/10.1016/j.neuroimage.2008.12.050.

Yang, Z., LaConte, S., Weng, X., Hu, X., Jun 2008. Ranking and averaging independent component analysis by reproducibility (raicar). Hum. Brain Mapp. 29 (6), 711–725 URL http://dx.doi.org/10.1002/hbm.20432.

Ylipaavalniemi, J., Soppela, J., 2009. Arabica: robust ica in a pipeline. ICA '09: Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation. Springer-Verlag, Berlin, Heidelberg, pp. 379–386.

Ylipaavalniemi, J., Vigario, R., Jan 2008. Analyzing consistency of independent components: an fmri illustration. Neuroimage 39 (1), 169–180 URL http://dx.doi.org/10.1016/j.neuroimage.2007.08.027.

Zarzoso, V., Comon, P., Aug. 20–25, 2008. Robust independent component analysis for blind source separation and extraction with application in electrocardiography. Proc. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2008, pp. 3344–3347.

Zhang, Z., Lu, G., Zhong, Y., Tan, Q., Yang, Z., Liao, W., Chen, Z., Shi, J., Liu, Y., Jul 2009b. Impaired attention network in temporal lobe epilepsy: a resting fmri study. Neurosci. Lett. (3), 97–101 URL http://dx.doi.org/10.1016/j.neulet.2009.04.040.

Zhang, Z., Lu, G., Zhong, Y., Tan, Q., Liao, W., Chen, Z., Shi, J., Liu, Y., Jun 2009a. Impaired perceptual networks in temporal lobe epilepsy revealed by resting fmri. J Neuro URL http://dx.doi.org/10.1007/s00415-009-5187-2.