Published in final edited form as:

J Neurosci Methods. 2016 May 1; 264: 129–135. doi:10.1016/j.jneumeth.2016.03.012.

The Role of Diversity in Complex ICA Algorithms for fMRI Analysis

Wei Du^a, Yuri Levin-Schwartz^{a,*}, Geng-Shen Fu^a, Sai Ma^a, Vince D. Calhoun^{b,c}, and Tülay Adalı^a

^aDepartment of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA

^bThe Mind Research Network, Albuquerque, NM 87106 USA

^cDepartment of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131 USA

Abstract

Background—The widespread use of data-driven methods, such as independent component analysis (ICA), for the analysis of functional magnetic resonance imaging data (fMRI) has enabled deeper understanding of neural function. However, most popular ICA algorithms for fMRI analysis make several simplifying assumptions, thus ignoring sources of statistical information, types of diversity, and limiting their performance.

New Method—We propose the use of complex entropy rate bound minimization (CERBM) for the analysis of actual fMRI data in its native, complex, domain. Though CERBM achieves enhanced performance through the exploitation of the three types of diversity inherent to complex fMRI data: noncircularity, non-Gaussianity, and sample-to-sample dependence, CERBM produces results that are more variable than simpler methods. This motivates the development of a minimum spanning tree (MST)-based stability analysis that mitigates the variability of CERBM.

Comparison with Existing Methods—In order to validate our method, we compare the performance of CERBM with the popular CInfomax as well as complex entropy bound minimization (CEBM).

Results—We show that by leveraging CERBM and the MST-based stability analysis, we are able to consistently produce components that have a greater number of activated voxels in physically meaningful regions and can more accurately classify patients with schizophrenia than components generated using simpler models.

Conclusions—Our results demonstrate the advantages of using ICA algorithms that can exploit all inherent types of diversity for the analysis of fMRI data when coupled with appropriate stability analyses.

*Corresponding author: ylevins1@umbc.edu (Yuri Levin-Schwartz).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

ICA; complex fMRI; AOD task

1. Introduction

Recent advances in technology have greatly enhanced both the spatial resolution of functional magnetic resonance imaging (fMRI) data and the speed at which it can be collected, thus allowing researchers the unprecedented ability to characterize and classify brain disorders including schizophrenia. Since relatively little is known about the exact processes underlying neural activation, it is important to make as few assumptions as possible when performing an analysis on such data, especially when the analysis is exploratory in nature. This incentivizes the use of data-driven methods, such as independent component analysis (ICA) (McKeown et al., 1998). Despite the popularity of ICA for fMRI analysis, see *e.g.* (McKeown et al., 2003, Rosazza and Minati, 2011, Calhoun and Adalı, 2012, Smith et al., 2013, Lee et al., 2013, Griffanti et al., 2014), most of the popular ICA algorithms used for fMRI analysis, such as Infomax (Bell and Sejnowski, 1995) and FastICA (Hyvärinen, 1999), make several implicit assumptions, which can substantially affect their performance. The first implicit assumption of many popular ICA algorithms for fMRI analysis involves the use of a fixed nonlinearity. The use of a fixed nonlinearity can significantly bias the estimated sources when there is a difference between distribution implied by the nonlinearity and the true source distribution (Adalı et al., 2015). A second assumption involves analyzing the fMRI data in the real domain. Since fMRI data is naturally complex, performing an analysis in the real domain will lead to a loss of information (Hoogenraad et al., 1998, Menon, 2002, Rauscher et al., 2005, Tomasi and Caparelli, 2007, Arja et al., 2010, Yu et al., 2015) stemming from the transformation of the complex signal to the real domain and ignoring the potentially useful property of noncircularity, which fMRI data has been shown to exhibit (Li et al., 2011, Rodriguez et al., 2012, Loesch and Yang, 2013). A final implicit assumption made by the majority of ICA algorithms for fMRI analysis is that there exists no sample-to-sample dependence between adjacent voxels. In fMRI data, activation patterns of blood-oxygenation-level-dependent (BOLD) signals tend to be spatially smooth and clustered, hence, ignoring the potentially informative property of spatial dependence can lead to inferior performance (Du et al., 2014). To avoid such assumptions, the adoption of a more general ICA framework is necessary. For this reason, we propose to use complex entropy rate bound minimization (CERBM) to study fMRI data in its native complex domain and take full advantage of all important statistical properties—types of diversity—inherent to complex fMRI data: dynamic non-Gaussianity, noncircularity, and sample-to-sample dependence (Fu et al., 2015). We compare the performance of the popular complex Infomax (CInfomax) (Calhoun and Adalı, 2002) to that of two algorithms, complex entropy bound minimization (CEBM) (Li and Adalı, 2010), a more recent ICA algorithm that exploits both noncircularity and dynamic non-Gaussianity, but ignores sample-to-sample dependence, as well as CERBM. Note that we use the term dynamic non-Gaussianity to differentiate techniques that use a fixed nonlinearity, such as CInfomax, from those that use a dynamic nonlinearity, such as CEBM and CERBM.

Though able to exploit a greater number of statistical properties due their more general source models, the iterative nature of CEBM and CERBM means that they can converge to local minima, hence producing solutions that are more variable than those of less general methods. This variability motivates the performance of a stability analysis, in order to identify a result that represents a consistent estimate to use as a final output. A widely used approach for resolving this issue in spatial ICA is ICASSO (Himberg and Hyvarinen, 2003), an exploratory visualization method for investigating the relations between estimates. ICASSO runs an ICA algorithm several times and clusters estimated components from all runs based on the absolute value of the correlation between estimates and then selects the centroid of each cluster as the consistent estimate. However, the direct use of centroids can lead to loss of information, since more than one type of component may be grouped into the same cluster, especially when the ICA model order is high, but only one type of component can be selected as the centroid. Additionally, some approaches for selecting centroids with ICASSO can result in the selection of different runs for different components, which breaks the connection with the ICA mixing model. A related and much explored area of research, particularly for multi-subject fMRI data, is a clustering analysis, see *e.g.*, (Sabuncu et al., 2010, Lashkari et al., 2010a,b, Deligianni et al., 2011, Varoquaux et al., 2013, Ng et al., 2013). Generally, these methods focus on matching estimated components across subjects, though it is a simple task to adjust the methods to match components based on runs. However, a natural goal when performing a stability analysis on multi-subject fMRI data is to determine a set of robust *T*-maps for use in further analyses, an issue that clustering methods, in general, ignore. Additionally, the method for performing a stability analysis should be readily applicable to both real-valued as well as complex fMRI data, the latter being a case that is rarely explored in the context of clustering. In order to address such a diverse set of challenges, we discuss a novel minimum spanning tree (MST)-based approach to investigate the stability of ICA algorithms in both the real and complex domain (Du et al., 2014). Our method not only addresses the issue of determining robust *T*-maps in both the complex and real domains, but also provides a metric for interpreting the statistical reliability of estimated components for fMRI data. Note that an MST-based method has been previously used for clustering in the context of an fMRI analysis (Baumgartner et al., 2001), however in this study the MST is to cluster fMRI time courses and not fMRI spatial maps.

The remainder of the paper is organized as follows. In Section 2 we discuss the statistical basis of CERBM, the MST-based stability analysis method, as well as the complex fMRI data used in this study. Section 3 contains the discussion of the simulation results as well as the experimental results using actual complex-fMRI data. We present our conclusions in Section 4.

2. Materials and Methods

2.1. ICA of Complex fMRI and CERBM

The noiseless spatial ICA model for complex fMRI data gathered from a single subject is given by

$$\mathbf{x}(v) = \mathbf{A}\mathbf{s}(v), \quad 1 \leq v \leq V, \quad (1)$$

where the N latent sources or spatial neural activation patterns, $\mathbf{s}(v) \in \mathbb{C}^N$, are strictly linearly mixed by the mixing matrix, $\mathbf{A} \in \mathbb{C}^{N \times N}$, producing observations, $\mathbf{x}(v) \in \mathbb{C}^N$. If we gather the observations together into a single complex matrix, $\mathbf{X} \in \mathbb{C}^{N \times V}$, where the i th row is formed by flattening the i th brain volume of V voxels. By assuming spatial independence on the part of the latent sources, ICA seeks to estimate a demixing matrix, $\mathbf{W}$, such that the estimated components are given by $\hat{\mathbf{S}} = \mathbf{W}\mathbf{X}$.

The model in (1) can be extended in a straightforward manner to K data sets as

$$\mathbf{x}^{[k]}(v) = \mathbf{A}^{[k]} \mathbf{s}^{[k]}(v), \quad 1 \leq v \leq V, \quad 1 \leq k \leq K.$$

Since there are scaling and permutation ambiguities associated with ICA, running ICA individually for each subject and aligning the results is impractical. For this reason and to exploit similarities across subjects, we assume that each subject has N common components but individual mixing matrices. Following two successive steps of principal component analysis, first at the subject level and then at the group level, a single Group ICA (Calhoun et al., 2001), implemented in the Group ICA of fMRI Toolbox (GIFT) (GIFT, 2011), is performed and individual subject maps are back-reconstructed. Clearly, the proper choice of ICA algorithm is vital to the success of Group ICA, thus using an algorithm which can take advantage of multiple types of statistical properties, such as CERBM, which we describe next, is important.

For CERBM, mutual information rate, which leads to the minimization of entropy rate, provides the umbrella under which dynamic non-Gaussianity, noncircularity, and sample-to-sample dependence are jointly exploited to yield an independent decomposition. The mutual information rate of the estimated sources, $\mathcal{J}_r(\hat{\mathbf{s}}_1, \dots, \hat{\mathbf{s}}_N)$, where $\hat{\mathbf{s}}_i$ is the i th estimated component, can be written as

$$\mathcal{J}_r(\hat{\mathbf{s}}_1, \dots, \hat{\mathbf{s}}_N) = \sum_{i=1}^N H_r(\hat{\mathbf{s}}_i) - \log |\det(\mathbf{W}\mathbf{W}^H)| - H_r(\mathbf{x}),$$

where H_r is the entropy rate and $(\cdot)^H$ is the Hermitian operator. Since $H_r(\mathbf{x})$ is constant with respect to $\mathbf{W}$ and $\det(\mathbf{W}\mathbf{W}^H) = |\det(\mathbf{W})|^2$, the cost function for CERBM can be written as (Fu et al., 2015)

$$\mathcal{J}_r(\mathbf{W}) = \sum_{i=1}^N H_r(\hat{\mathbf{s}}_i) - 2 \log |\det(\mathbf{W})|. \quad (2)$$

From (2), it can be seen that CERBM naturally incorporates diversity in the form of sample-to-sample dependence through the use of entropy rate in the formulation. When samples are

independent and identically distributed, entropy rate reduces to simple entropy. In addition, by considering a density model that takes full noncircularity into account (Adalı et al., 2008), CERBM is able to exploit another form of diversity, noncircularity (Fu et al., 2015).

Note that the performance of CERBM is greatly affected by the estimation of $H_f(\hat{s}_j)$. For this reason and due to the difficulty of estimating the entropy rate for a complex random variable, CERBM uses an invertible source model to model the sample-to-sample dependence and bounds the entropy rate with four measuring functions. These four measuring functions each correspond to a different class of distributions: symmetric, asymmetric, unimodal, and bi-modal, with CERBM selecting the measuring function corresponding to the tightest bound on the entropy rate (Fu et al., 2015). This dynamic selection of the appropriate bound enables CERBM to exploit non-Gaussianity dynamically; this is contrasted to CInfomax, which does not exploit non-Gaussianity dynamically due to its use of a fixed hyperbolic tangent nonlinearity Adalı et al. (2004). It should be noted that the Infomax algorithm has been extended in the real domain to fit both super-Gaussian as well as sub-Gaussian sources (Lee et al., 1999), however this is a weaker form of dynamic non-Gaussianity than that possessed by CEBM and CERBM since extended Infomax cannot describe either bi-modal or asymmetric distributions.

2.2. Stability Analysis using MST

The general framework provided by CERBM allows enhanced performance through the exploitation of three types of statistical properties: dynamic non-Gaussianity, noncircularity, and sample-to-sample dependence. However, this enhancement comes at a cost. As a method becomes more general, its stability from one run to another decreases, incentivizing the development of a method to determine a consistent estimate for use in further analyses or to compare the performance of different ICA algorithms. We must stress that the goal of our method is not to determine a “best run,” which can simply be found by selecting the run with the lowest value of the likelihood function, but to select a stable run, which represents the “average” performance of the algorithm and varies little from one set of runs to another. The computation and use of such a stable run, here referred to as a “final run,” enables greater confidence in the reproducibility of results and is described next. Note that in the complex regime considered in this paper, the scaling ambiguity inherent to ICA algorithms includes a phase term. Thus, in order to account for the phase information, rather than normalize the magnitude of the components, the components are normalized using the Mahalanobis distance (Mahalanobis, 1936, Rodriguez et al., 2012) prior to performing the stability analysis. For this reason, we refer to the spatial maps corresponding to components normalized using the Mahalanobis distance as Mahalanobis distance maps, Z_c -maps, to differentiate them from spatial maps corresponding to components normalized using only the magnitude, Z -maps.

The MST-based method depicted in Fig. 1 considers the problem of finding a stable run, first, as a generalized assignment problem (GAP) and then exploits the inherent structure of the GAP using graph theory (Du et al., 2014). In the first stage, the objective is to align the N estimated components from each pairwise combination of runs, such that the assignment cost is minimized. For run k and run l , this is accomplished through the use of a cost matrix,

$\mathbf{C}_{kl}$, whose i j th element is given by $c_{ij} = 1 - \Psi(\hat{s}_i^{[k]}, \hat{s}_j^{[l]})$, $0 \leq \Psi \leq 1$, where $\hat{s}_i^{[k]}$, $\hat{s}_j^{[l]}$, and Ψ , are the i th estimated source from run k , j th estimated source from run l , and an appropriate similarity measure, respectively. In this work we use the Pearson correlation coefficient as the distance measure for two main reasons: first, it is symmetric thus $c_{ij} = c_{ji}$ and second, the correlation coefficient describes statistical properties of the sources and, unlike mutual information, is bounded between 0 and 1, thus c_{ij} is 0 if and only if $\hat{s}_i^{[k]} = \hat{s}_j^{[l]}$. The Hungarian algorithm (Kuhn, 1955) is used to solve the assignment problem, *i.e.*, align the components, for each pair of runs and to obtain the cost of aligning the runs. Then, a fully connected, undirected graph is constructed with nodes representing runs and edges with weights representing the cost of aligning each pair of runs (Du et al., 2014).

The second stage of the MST-based method processes the graph produced in the first stage, aligning the estimated components across all runs. First, an MST (Kruskal, 1956) and corresponding central node, $\hat{\mathbf{S}}^{[m]}$, are found for the graph, allowing all runs to be connected with low, though sub-optimal, cost. The central node provides local information on the run that is most similar to the other runs, since it is only affected by runs that share the greatest degree of similarity and not affected by those that share little similarity with the other runs, which may not be the case for other metrics, such as lowest average cost of alignment. Having selected the central run, the components of the other runs are aligned with the components of the central run. Then, for each corresponding component across K runs, a one-sample t -test across subjects is performed, producing a single T -map for each of the N groups of components. The use of T -maps is well justified for fMRI analysis, since T -maps can be used to investigate the reliability of components across subjects or runs. Finally, the “final run,” $\hat{\mathbf{S}}^{[F]}$, is selected as the run with highest average correlation between its estimated components and the corresponding T -maps. This final step gives us the run whose maps best correspond with the group result, *i.e.*, the run that, unlike the central run, is a good representation of the “average” performance of the algorithm. Note that this step avoids the two major issues associated with ICASSO, since the most consistently estimated components are used to cluster components and a single run is guaranteed to be selected as the “final run,” thus maintaining the relation with the ICA mixing model. Additionally, the MST-based method naturally takes advantage of the expectation that meaningful components will be more consistently estimated than noise components, leading to a higher similarity measure, and thus the cost of aligning meaningful components will be lower than the cost of aligning less consistent components.

2.3. Complex fMRI Data

The complex fMRI data used in this study was gathered from 69 schizophrenia patients and 31 healthy controls. The participants were scanned while listening to three kinds of auditory stimuli: standard (500 Hz occurring with probability 0.80), novel (nonrepeating random digital sounds with probability 0.10), and target (1 kHz with probability 0.10, to which a button press with the right index finger was required). Further details of the AOD paradigm and image acquisition parameters are described in (Kiehl et al., 2005).

2.4. Simulation Setup

For real brain data, the ground truth, *i.e.*, the collection of true latent sources and subject covariations, is not available. Thus, in order to test the performance of CERBM, we generate simulated data similar to the artificial complex fMRI sources used in (Xiong et al., 2008) and examine the performance of CInfomax, CEBM, and CERBM, in terms of intersymbol interference (ISI) (Amari et al., 1996), as the number of samples, V , increases from 100 to 100,000. The 8 sources are corrupted by first order auto-regressive (AR) noise, with an AR coefficient of 0.4, and non-circularity between 0.2 and 0.8. The noise followed was driven by a generalized Gaussian distribution (GGD) with shape parameter, β , between 0.4 and 7. The shape parameter controls the non-Gaussianity of the GGD, with $\beta > 1$ being sub-Gaussian, *i.e.*, distributions with tails that approach zero faster than a Gaussian, and $\beta < 1$ being super-Gaussian, *i.e.*, distributions with tails that approach zero slower than a Gaussian. Thus, generating sources with a range of values for the shape parameter favors algorithms that can incorporate non-Gaussianity dynamically. In addition, added information provided by the sample-to-sample dependence imposed by the AR source model favors algorithms that can exploit this added type of diversity. The plots presented here are the average result of 100 independent simulations, where, for each simulation, every method was used on the same data and each algorithm was initialized at the same random point per simulation.

3. Results and Discussion

3.1. Simulation Results

We observe in Fig. 2 the high level of performance obtained using CERBM for simulated sources. This performance is due to the exploitation of the three types of diversity: non-circularity, dynamic non-Gaussianity, and sample-to-sample dependence. The value of exploiting sample-to-sample dependence can be seen in the performance improvement of CERBM over CEBM, which only exploits non-circularity and dynamic non-Gaussianity. Additionally, the value of exploiting dynamic non-Gaussianity can be seen in the improvement in performance of methods that use this diversity, *i.e.* CEBM and CERBM, over CInfomax, which does not.

3.2. Complex fMRI Results

In order to evaluate the performance of CERBM, Group ICA using CInfomax, CEBM, and CERBM is performed on the complex fMRI data and 30 independent components (ICs) were estimated for each algorithm. The number of components estimated was based on the order selection results achieved using a modified minimum description length (MDL) criterion (Xiong et al., 2008). In order to compare the results using different algorithms, each algorithm was run 10 times and the representative “final run” for each algorithm was determined using the MST-based method described previously.

3.2.1. Component Maps—Task-related components were selected to evaluate the performance of CERBM on the complex fMRI data used in this study. The task-related components were extracted based upon the magnitude of the correlation between IC timecourse and the model timecourse. The ICs were further evaluated visually using prior information about the task as well as the spatial properties of fMRI, such as smooth and focal

regions of activity. Figure 3 shows the Z -maps corresponding to the “final run” for each of the three algorithms under study. Numbered 1 through 10, these ICs include: 1. DMN-1; 2. DMN-2; 3, temporal; 4, motor-temporal; 5, motor; 6, sensorimotor; 7, right frontoparietal; 8, left frontoparietal; 9, medial visual; and 10, lateral visual, respectively. Note that to aid in displaying the results, the maps have been flipped so as to only show activation. These maps show a clear increase in the strength and bilateral nature of activation for CEBM and CERBM over CInfomax in IC1, IC2, IC3, IC5, IC6, IC7, IC9, and IC10. We also note a reduction in the number of unwanted “noise voxels,” corresponding to fragmented or physically meaningless activation, for CEBM and CERBM over CInfomax in IC3, IC4, IC5, and IC10. These results demonstrate, in a qualitative manner, the performance improvement, in terms of the extraction of physically meaningful components, achieved by using CEBM and CERBM over CInfomax.

In order to quantify the performance of CInfomax, CEBM, and CERBM in the results presented in Figure 3, receiver operator characteristic (ROC) curves were generated for all components from binary masks. These masks were composed of joined dilated Brodmann areas (BAs), generated using WFU PickAtlas (Maldjian et al., 2003, 2004), where for components that represented the same functional area but were split between more than one IC during the estimation, such as: IC1 and IC2 for default mode network (DMN), IC5 and IC6 for motor, IC7 and IC8 for frontoparietal, as well as IC9 and IC10 for visual, the corresponding ICs were combined prior to mask formation. Note that for brevity, ROC curves for all subjects and corresponding spatial masks for only the DMN, temporal, and motor components are presented in Figures 4 and 5, respectively. Additionally, note that, ROC curves were generated for the Z_c -maps as well as for the corresponding Z -maps to show the significant improvement achieved when accounting for phase information. The generation of the ROC curves is described below.

At a given value of Z_c — Z for the magnitude map—the number of voxels within the mask, N_{in} , and the number of voxels outside the mask, N_{out} , with values higher than Z_c or Z are found. Then, the probability of detection, P_D , and false alarm, P_{FA} , are calculated using the ratios, N_{true}/N_{in} and N_{false}/N_{out} , respectively, where N_{true} and N_{false} are determined by the masks. The ROC curve is generated by plotting P_D and P_{FA} as the value of Z_c and Z is increased from 0.

From Figure 4, we can see that for all three components CERBM has better performance than CInfomax, though such claims cannot be made about CEBM. We can also see the clear improvement in the performance for each of the algorithms when Z_c -maps are used rather than the Z -maps. This demonstrates the importance of accounting for the phase information and performing an ICA analysis on fMRI data in its native, complex, domain (Li et al., 2011, Rodriguez et al., 2011, 2012).

3.2.2. Classification of Patients—Additionally, we perform classification on selected components to further investigate the estimation of spatial components. In the classification, we apply a linear support vector machine (SVM) (Cortes and Vapnik, 1995), to the Z_c maps of the “final runs” for each algorithm in order to identify schizophrenia patients from controls. Table 1 shows classification results of each selected component, including

accuracy, sensitivity, and specificity, which are calculated by using a leave-one-out approach. Sensitivity measures the proportion of actual patients which are correctly identified in classification and specificity represents the percentage of correctly classified controls. Overall, the results indicate that CERBM tends to estimate better spatial components than CEBM and CInfomax in terms of classification performance. We should note that it may be possible to improve the classification performance through the use of feature selection and other classifiers.

4. Conclusions

Data-driven methods such as ICA have achieved wide-spread popularity for the analysis of fMRI data due to their ability to let the data “speak.” However, the popular ICA algorithms for fMRI analysis, such as Infomax, make several simplifying assumptions that can substantially limit their performance. These include: the analysis of fMRI data in the real domain, the use of a fixed nonlinearity, and the assumption of no sample-to-sample dependence. For this reason, we propose the use of more general ICA algorithms, such as complex entropy bound minimization (CERBM), for the analysis of fMRI data. We have shown that CERBM achieves improved performance of Infomax and CEBM in both simulations and the analysis of actual complex fMRI data. Though able to achieve enhanced performance, CERBM produces results that are more variable than simpler methods, motivating the development of a stability analysis that can identify a result that represents a consistent estimate to use as a final output. To this end, we also discussed a minimum spanning tree (MST)-based method through which we are able to compare the performance of different algorithms on actual fMRI data and mitigate the inherent instability of CERBM. We have shown that by leveraging CERBM and the MST-based stability analysis, we are able to produce components that tend to more accurately classify patients with schizophrenia than those generated using simpler models. These results show the advantage of CERBM when combined with the MST-based stability analysis. The code for CERBM is freely available on the authors’ webpage (mlsp.umbc.edu) and the code for the MST-based stability analysis will be added soon. Currently the code for the MST-based stability analysis is available as part of the open access Group ICA of fMRI Toolbox (GIFT). A final note is that leveraging a more general ICA algorithm in combination with the MST-based stability analysis can be used in real-valued ICA, potentially enhancing both the quality and interpretability of the results of the fMRI analysis.

Acknowledgments

This work was supported by the following grants NIH-NIBIB R01 EB 005846 and NSF-CCF-1117056.

References

- Adali, T.; Kim, T.; Calhoun, V.D. Independent component analysis by complex nonlinearities. 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP); 2004. p. 525-528.
- Adali T, Levin-Schwartz Y, Calhoun V.D. Multi-modal data fusion using source separation: Application to medical imaging. *Proceedings of the IEEE*. 2015; 103(9):1494–1506.
- Adali T, Li H, Novey M, Cardoso J.F. Complex ica using nonlinear functions. *IEEE Transactions on Signal Processing*. Sep; 2008 56(9):4536–4544.

- Amari, S.; Cichocki, A.; Yang, HH. Advances in Neural Information Processing Systems. Vol. 8. MIT Press; Cambridge, MA: 1996. A new learning algorithm for blind signal separation; p. 757-763.
- Arja SK, Feng Z, Chen Z, Caprihan A, Kiehl KA, Adali T, Calhoun VD. Changes in fMRI magnitude data and phase data observed in block-design and event-related tasks. *NeuroImage*. 2010; 49(4): 3149–3160. [PubMed: 19900561]
- Baumgartner R, Somorjai R, Summers R, Richter W. Ranking fMRI time courses by minimum spanning trees: Assessing coactivation in fMRI. *NeuroImage*. 2001; 13(4):734–742. [PubMed: 11305900]
- Bell A, Sejnowski T. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*. 1995; 7:1129–1159. [PubMed: 7584893]
- Calhoun, VD.; Adali, T. Complex Infomax: convergence and approximation of infomax with complex nonlinearities. *Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing*; 2002. p. 307-316.
- Calhoun VD, Adali T. Multisubject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Reviews in Biomedical Engineering*. 2012; 5:60–73. [PubMed: 23231989]
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*. 2001; 14(3):140–151. [PubMed: 11559959]
- Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995; 20(3):273–297.
- Deligianni, F.; Varoquaux, G.; Thirion, B.; Robinson, E.; Sharp, DJ.; Edwards, AD.; Rueckert, D. Information Processing in Medical Imaging. Vol. 6801 of Lecture Notes in Computer Science. Springer; Berlin Heidelberg: 2011. A probabilistic framework to infer brain functional connectivity from anatomical connections; p. 296-307.
- Du, W.; Fu, G-S.; Calhoun, V.; Adali, T. Performance of complex-valued ICA algorithms for fMRI analysis: Importance of taking full diversity into account. *2014 IEEE International Conference on Image Processing (ICIP)*; Oct. 2014 p. 3612-3616.
- Fu GS, Phlypo R, Anderson M, Adali T. Complex independent component analysis using three types of diversity: Non-Gaussianity, nonwhiteness, and noncircularity. *IEEE Transactions on Signal Processing*. 2015; 63(3):794–805.
- GIFT. Group ICA of fMRI Toolbox (GIFT). 2011. <http://mialab.mrn.org/software/gift/index.html>
- Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, Moeller S, Xu J, Yacoub E, Baselli G, Ugurbil K, Miller KL, Smith SM. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*. 2014; 95:232–247. [PubMed: 24657355]
- Himberg, J.; Hyvarinen, A. ICASSO: software for investigating the reliability of ICA estimates by clustering and visualization. *IEEE 13th Workshop on Neural Networks for Signal Processing*; Sep. 2003 p. 259-268.
- Hoogenraad FG, Reichenbach JR, Haacke EM, Lai S, Kuppusamy K, Sprenger M. In vivo measurement of changes in venous blood-oxygenation with high resolution functional MRI at 0.95 tesla by measuring changes in susceptibility and velocity. *Magnetic resonance in medicine*. 1998; 39(1):97–107. [PubMed: 9438443]
- Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*. 1999; 10(3):626–634. [PubMed: 18252563]
- Kiehl KA, Stevens MC, Laurens KR, Pearlson G, Calhoun VD, Liddle PF. An adaptive reflexive processing model of neurocognitive function: supporting evidence from a large scale (n = 100) fMRI study of an auditory oddball task. *NeuroImage*. 2005; 25(3):899–915. [PubMed: 15808990]
- Kruskal, Joseph BJ. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*. 1956; 7(1):48–50.
- Kuhn HW. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*. 1955; 2(1–2):83–97.
- Lashkari D, Sridharan R, Golland P. Categories and functional units: an infinite hierarchical model for brain activations. *Advances in neural information processing systems*. 2010a:1252–1260. [PubMed: 24839377]

- Lashkari D, Vul E, Kanwisher N, Golland P. Discovering structure in the space of fMRI selectivity profiles. *NeuroImage*. 2010b; 50(3):1085–1098. [PubMed: 20053382]
- Lee MH, Smyser CD, Shimony JS. Resting-state fMRI: a review of methods and clinical applications. *American Journal of Neuroradiology*. 2013; 34(10):1866–1872. [PubMed: 22936095]
- Lee TW, Girolami M, Sejnowski TJ. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*. 1999; 11(2): 417–441. [PubMed: 9950738]
- Li H, Correa N, Rodriguez P, Calhoun V, Adalı T. Application of independent component analysis with adaptive density model to complex-valued fMRI data. *IEEE Transactions on Biomedical Engineering*. 2011; 58(10):2794–2803. [PubMed: 21690000]
- Li XL, Adalı T. Complex independent component analysis by entropy bound minimization. *IEEE Transactions on Circuits and Systems I: Regular Papers*. Jul; 2010 57(7):1417–1430.
- Loesch B, Yang B. Cramér-Rao bound for circular and noncircular complex independent component analysis. *IEEE Transactions on Signal Processing*. 2013; 61(2):365–379.
- Mahalanobis, PC. On the generalised distance in statistics. *Proceedings National Institute of Science; India*. Apr. 1936 p. 49-55.
- Maldjian JA, Laurienti PJ, Burdette JH. Precentral gyrus discrepancy in electronic versions of the talairach atlas. *NeuroImage*. 2004; 21(1):450–455. [PubMed: 14741682]
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*. 2003; 19(3):1233–1239. [PubMed: 12880848]
- McKeown MJ, Hansen LK, Sejnowski TJ. Independent component analysis of functional MRI: what is signal and what is noise? *Current opinion in neurobiology*. 2003; 13(5):620–629. [PubMed: 14630228]
- McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI Data by Blind Separation Into Independent Spatial Components. *Human Brain Mapping*. 1998; 6:160–188. [PubMed: 9673671]
- Menon RS. Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magnetic Resonance in Medicine*. 2002; 47(1):1–9. [PubMed: 11754436]
- Ng, B.; Varoquaux, G.; Poline, JB.; Thirion, B. A novel sparse group Gaussian graphical model for functional connectivity estimation. *Proceedings of the 23rd International Conference on Information Processing in Medical Imaging*. IPMI'13; Berlin, Heidelberg: Springer-Verlag; 2013. p. 256-267.
- Rauscher A, Sedlacik J, Barth M, Mentzel HJ, Reichenbach JR. Magnetic susceptibility-weighted MR phase imaging of the human brain. *American journal of neuroradiology*. 2005; 26(4):736–742. [PubMed: 15814914]
- Rodriguez PA, Calhoun VD, Adalı T. De-noising, phase ambiguity correction and visualization techniques for complex-valued ICA of group fMRI data. *Pattern Recogn*. 2012; 45(6):2050–2063.
- Rodriguez PA, Correa NM, Eichele T, Calhoun VD, Adalı T. Quality map thresholding for de-noising of complex-valued fMRI data and its application to ICA of fMRI. *Journal of signal processing systems*. 2011; 65(3):497–508.
- Rosazza C, Minati L. Resting-state brain networks: literature review and clinical applications. *Neurological Sciences*. 2011; 32(5):773–785. [PubMed: 21667095]
- Sabuncu MR, Singer BD, Conroy B, Bryan RE, Ramadge PJ, Haxby JV. Function-based intersubject alignment of human cortical anatomy. *Cerebral Cortex*. 2010; 20(1):130–140. [PubMed: 19420007]
- Smith SM, Vidaurre D, Beckmann CF, Glasser MF, Jenkinson M, Miller KL, Nichols TE, Robinson EC, Salimi-Khorshidi G, Woolrich MW, Barch D, Ugurbil K, Van Essen D. Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*. 2013; 17(12):666–682. [PubMed: 24238796]
- Tomasi DG, Caparelli EC. Macrovascular contribution in activation patterns of working memory. *Journal of Cerebral Blood Flow & Metabolism*. 2007; 27(1):33–42. [PubMed: 16639427]

- Varoquaux, G.; Schwartz, Y.; Pinel, P.; Thirion, B. Information Processing in Medical Imaging. Vol. 7917 of Lecture Notes in Computer Science. Springer; Berlin Heidelberg: 2013. Cohort-level brain mapping: Learning cognitive atoms to single out specialized regions; p. 438-449.
- Xiong, W.; Li, Y-O.; Li, H.; Adalı, T.; Calhoun, VD. On ICA of complex-valued fMRI: Advantages and order selection. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); Mar. 2008 p. 529-532.
- Yu MC, Lin QH, Kuang LD, Gong XF, Cong F, Calhoun VD. ICA of full complex-valued fMRI data using phase information of spatial maps. *Journal of neuroscience methods*. 2015; 249:75–91. [PubMed: 25857613]

Highlights

- Most ICA algorithms used for fMRI analysis make several simplifying assumptions.
- We use CERBM and an MST-based analysis to exploit all information in fMRI data.
- Our method finds more meaningful discriminative components than current methods.
- General ICA algorithms achieve superior performance in the analysis of fMRI data.

Figure 1.

Illustration of the MST-based stability analysis framework. The framework is as follows: (a) multiple independent ICA runs are aligned and (b) formed into an MST (c) all sources from other runs are aligned to the central run of the MST and (d) *T*-maps are generated for each cluster of components (e) the final, most stable, is the run with highest average correlation to the group *T*-maps.

Figure 2.

Performance of CInfomax, CEBM, and CERBM in terms of ISI versus sample size. Note the improved performance of CEBM and CERBM over CInfomax due to dynamic exploitation of non-Gaussianity. Additionally, note the improved performance of CERBM over CEBM due to utilization of the additional type of diversity, sample-to-sample dependence.

Figure 3.

Z_c -maps corresponding to the “final run” for each of the three algorithms under study. These ICs include: 1. DMN-1; 2. DMN-2; 3. temporal; 4. motor-temporal; 5, motor; 6. sensorimotor; 7. right frontoparietal; 8. left frontoparietal; 9. medial visual; and 10. lateral visual. The maps have a Z_c threshold corresponding significance for activation of $p > 0.05$. Note that to aid in displaying the results, all maps have been adjusted so as to only show activation.

Figure 4. ROC curves obtained for the Z_c -maps as well as the Z -maps of the estimated DMN, temporal, and motor components. Note the improvement in performance achieved by all algorithms when phase information is exploited.

Figure 5.

Binary masks for the DMN, temporal, and motor components. The Brodmann areas used to create the masks are: DMN: 7, 10, 23, 31, 39; temporal: 20, 21, 22, 37, 38, 42; and motor: 4, 6, 43 47.

Table 1

Classification results of the three complex ICA algorithms. CERBM has the best performance in terms of classification accuracy for DMN, temporal, and motor components.

Algorithms	Classification accuracy (sensitivity / specificity)		
	DMN	Temporal	Motor
CInfomax	0.85 (0.93/0.68)	0.80 (0.97/0.42)	0.81 (0.93/0.55)
CEBM	0.81 (0.93/0.55)	0.82 (0.97/0.48)	0.83 (0.96/0.55)
CERBM	0.86 (0.96/0.65)	0.82 (0.99/0.45)	0.83 (0.96/0.55)