A Method to Compare the Discriminatory Power of Data-driven Methods: Application to ICA and IVA

Yuri Levin-Schwartz[1], Vince D. Calhoun[2,3], and Tülay Adalı[1]

[1]Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250
[2]The Mind Research Network, Albuquerque, NM 87106
[3]Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131

Keywords: fMRI, ICA, Schizophrenia

Abstract

Functional magnetic resonance imaging (fMRI) facilitates the study of neural function and how it is disrupted by psychiatric disorders, such as schizophrenia. Since relatively little is known about the form of neural activation *a priori*, it is important to minimize the assumptions placed on the data. This notion has motivated the development of many data-driven, factorization-based methods for the analysis of fMRI data, such as independent component analysis (ICA) and its multiset extension, independent vector analysis (IVA). The increasing number of these methods motivates their comparison. However, such an investigation is difficult to perform on real fMRI data, since the ground truth is unknown and comparing factors across different techniques is both tedious and imprecise. For this reason, different methods are usually compared using synthetic data, which is generally quite different in nature from real fMRI data. In this paper, we present a novel method, global difference maps (GDMs), to compare the results of different fMRI analysis techniques on real fMRI data, quantify their relative performances, and highlight the differences between the decompositions visually. We apply this method to fMRI data from 109 patients with schizophrenia and 138 healthy controls during the performance of three tasks. Through this application of GDMs, we find that IVA can determine regions that are more discriminatory between patients and controls than ICA, though IVA is less effective at emphasizing regions found in only a subset of the tasks. These results demonstrate that GDMs are an effective way to compare the performances of different factorization-based methods as well as regression-based analyses.

1 Introduction

Due to its high spatial resolution and non-invasive nature, functional magnetic resonance imaging (fMRI) data has become one of the most popular means of understanding normal neural function as well as how it is disrupted by disorders, such as schizophrenia (Logothetis, 2008). The data processing strategies for fMRI data can be roughly grouped into two schemes: hypothesis-driven and data-driven (X. Zhao et al., 2004). Hypothesis-driven methods exploit actual or supposed *a priori* knowledge about brain activity and, generally, study neurological relationships across a few regions or with respect to specific stimuli. Data-driven methods, on the other hand, offer a less targeted and more holistic approach, often by decomposing the observed data into a set of factors. Such techniques include: principal component analysis (PCA) (Thirion & Faugeras, 2003; Zhong et al., 2009), independent component analysis (ICA) (Mckeown et al., 1998; Calhoun, Adalı, Pearlson, & Pekar, 2001; Stone, Porrill, Porter, & Wilkinson, 2002; Wang, 2011; Adalı, Levin-Schwartz, & Calhoun, 2015), dictionary learning/sparse coding (DL) (Abraham, Dohmatob, Thirion, Samaras, & Varoquaux, 2013; Abolghasemi, Ferdowsi, & Sanei, 2015; S. Zhao et al., 2015), non-negative matrix factorization (NMF) (Lohmann, Volz, & Ullsperger, 2007; Ferdowsi, Abolghasemi, & Sanei, 2010), tensor-based methods (Davidson, Gilpin, Carmichael, & Walker, 2013; Kuang et al., 2013), and independent vector analysis (IVA) (J.-H. Lee, Lee, Jolesz, & Yoo, 2008). The performance of each of these factor models depends on the validity of their modeling assumptions for the dataset being analyzed and, thus, motivates a comparison of different factor models on the same dataset. However, it is difficult to compare the performance of different factor models on real data, since the ground truth is not known and each method typically produces multiple factors.

In order to avoid this issue, many papers that compare different factorization techniques focus on their performance on simulated data, see *e.g.*, (X. Zhao et al., 2004; Degras & Lindquist, 2014; Kuang et al., 2015; Engberg, Andersen, Mørup, & Madsen, 2016). However, these artificial datasets are usually simple when compared with real fMRI data (Welvaert & Rosseel, 2014; Eklund & Nichols, 2017). When comparing the performance of different factor methods on real fMRI data, most papers align factors from different methods and then rely on a visual comparison, see *e.g.*, (Sui, Adalı, Pearlson, Clark, & Calhoun, 2009; Erhardt et al., 2011; K. Lee, Tak, & Ye, 2011; Moeller, LeVan, & Gotman, 2011; Schöpf, Windischberger, Kasess, Lanzenberger, & Moser, 2010). However, aligning even a subset of the total number of factors from multiple techniques can be time consuming, due to the potentially large number of factors from each method. Additionally, each method exploits different properties of the signal (Lahat, Adalı, & Jutten, 2015) and such comparisons are inherently subjective, since they rely on visual interpretation. Another metric for the comparison of different fac-

torization methods is reproducibility or generalizability, *i.e.*, their ability to produce similar factors across different subjects and sessions (Strother et al., 2002; Afshin-Pour, Hossein-Zadeh, Strother, & Soltanian-Zadeh, 2012; Rasmussen, Abrahamsen, Madsen, & Hansen, 2012; Roels, Bossier, Loeys, & Moerkerke, 2015). However, focusing solely on reproducibility ignores how informative the extracted factors are for a given task. In the case where multiple groups, such as those affected by a psychiatric disease and those who are healthy are analyzed, the ability of a factor to differentiate between the groups can be used to determine the performance of different factorization methods, see *e.g.*, (Sui et al., 2010; Ramezani, Marble, Trang, Johnsrude, & Abolmaesumi, 2015; Levin-Schwartz, Calhoun, & Adalı, 2017). The use of this criterion as a measure of performance is well motivated and exploits the knowledge that there should be some brain-related differences between the groups. However, the technique does not solve the fundamental problem of alignment, since often there are multiple discriminatory factors for each method.

In this paper, we present a novel model comparison technique, global difference maps (GDMs), and demonstrate how they can be used to visually highlight the differences between factorization methods and quantify the discriminative or relational power of a dataset within a decomposition. We apply this technique to highlight the differences between individual analyses, using ICA, and a joint analysis, using IVA, of three fMRI tasks: an auditory oddball (AOD) task, a Sternberg item recognition paradigm (SIRP) task, and a sensorimotor (SM) task. Through this application, we show how GDMs can be an effective way to compare the performances of different factorization-based methods. Results show that IVA can determine regions that are more discriminatory between patients and controls than ICA, however, this improvement in discriminatory power comes at the cost of not emphasizing some of the regions found using ICA in a subset of the tasks.

2 Materials and Methods

2.1 Feature Extraction

Since the timing of the stimuli in each task is different, it is difficult to jointly analyze multi-task fMRI data. Rather, for each subject, a simple linear regression is run on the data from each voxel using the statistical parametric mapping toolbox (SPM) (SPM5, 2011), where the regressors are created by convolving the hemodynamic response function (HRF) in SPM with the desired predictors for each task. The resulting regression coefficient maps are used as features for each subject and task. This reduction, using a lower-dimensional though still multivariate representation of the data, facilitates exploration of the associations across the features from multiple tasks, see

e.g., (Calhoun et al., 2006; Ramezani et al., 2014), as well as enabling the identification of intrinsic functional networks (Calhoun & Allen, 2013).

2.2 FMRI Tasks and Features

The datasets used in this study are from the Mind Research Network Clinical Imaging Consortium Collection (Gollub et al., 2013) (publicly available at http://coins.mrn.org). These datasets were obtained from 247 subjects, 109 patients with schizophrenia and 138 healthy controls. Next, we introduce the tasks as well as the associated multivariate features analyzed in this study.

2.2.1 Auditory Oddball Task

This auditory task required the subjects to listen to three different types of auditory stimuli in a pseudo-random order: standard (1 kHz tones occurring with a probability of 0.82), novel (complex sounds occurring with a probability of 0.09), and target (1.2 kHz tones occurring with a probability of 0.09, to which a right thumb button press was required) (Kiehl & Liddle, 2001; Gollub et al., 2013). Each run of the task consisted of 90 stimuli, each with a 200 ms duration and a randomly changing interstimulus interval of between 550 and 2,050 ms. The order of novel and target stimuli was changed between runs, thus ensuring that the responses were not dependent on the stimulus order (Gollub et al., 2013). For this task, the regressor was created by modeling both the target and standard stimuli as delta functions and convolving this sequence of delta functions with the default SPM HRF in addition to their temporal derivatives (Michael et al., 2009). Subject averaged contrast images between the target versus the standard tones were used as the feature for this task.

2.2.2 Sternberg Item Recognition Paradigm Task

This visual task required the subjects to remember a set of 1, 3, or 5 randomly chosen numbers between 0 and 9. The task paradigm consisted of: a 1.5 second *learn* condition, a blank screen for 0.5 seconds, a 6 second *encode* condition, where the whole sequence of digits was presented together, and a 38 second *probe* condition, where the subject was shown a sequence of integers and had to indicate, with a right thumb button press, whether or not it was a member of the memorized set (Gollub et al., 2013). Each probe digit was presented for up to 1.1 seconds in a pseudo-random fashion within a 2.7 second interval (Gollub et al., 2013). A total of 84 probes, consisting of 42 targets and 42 foils was obtained per scan and the prompt-encode-probe conditions were run twice for each set size in a pseudo-random order (Gollub et al., 2013). For this task, the regressor was created by convolving the probe response for the three digit set with the default SPM HRF

(Michael et al., 2009). This was done for both runs of the probe response and then the average map was computed and used as the feature for this task.

2.2.3 Sensory Motor Task

This auditory task required the subjects to listen to a sequence of auditory stimuli that consisted of 16 different tones each lasting 200 ms and ranging in frequency from 236 Hz to 1,318 Hz with a 500 ms inter-stimulus interval (Gollub et al., 2013). The first tone presented was at the lowest pitch and each subsequent tone was higher in pitch than the previous one until the highest tone was reached, then the order of the tones was reversed (Gollub et al., 2013). Each tonal change required a right thumb button press. A run consisted of 15 increase-and-decrease blocks, alternated with 15 fixation blocks, with each block lasting 16 seconds in duration (Gollub et al., 2013). For this task, the regressor was created by convolving the entire increase-and-decrease block with the default SPM HRF (Michael et al., 2009). This was done for both runs for each subject and then the average map was used as the feature for this task.

2.3 ICA

For an fMRI feature dataset from M subjects, $\mathbf{X} \in \mathbb{R}^{M \times V}$, where the mth row of $\mathbf{X}$ is formed by flattening the feature of V voxels from the mth subject, the noiseless ICA model can be written as

$$\mathbf{X} = \mathbf{AS}, \tag{1}$$

where the C spatially independent latent sources or neural activation patterns, $\mathbf{S} \in \mathbb{R}^{C \times V}$, are linearly mixed by the mixing matrix, $\mathbf{A} \in \mathbb{R}^{M \times C}$. Given this model, ICA seeks to determine a demixing matrix, $\mathbf{W}$, such that the estimated components, the rows of $\hat{\mathbf{S}} = \mathbf{WX}$, are as independent as possible. Since we seek to maximize independence between the estimated components, $\hat{\mathbf{s}}_1, ..., \hat{\mathbf{s}}_C$, estimation of the demixing matrix can be accomplished through the minimization of the mutual information among the components, written as

$$\mathcal{I}_{\text{ICA}}(\mathbf{W}) = \sum_{i=1}^{C} H\left(\hat{\mathbf{s}}_i\right) - \log|\det(\mathbf{W})| - H(\mathbf{X}), \tag{2}$$

where $H(\cdot)$ is the entropy. Note that, since the data from each subject within a task has been reduced to a feature, the columns of the estimated mixing matrix, $\hat{\mathbf{A}}$, provide the weight of each component across the subjects. This means that the ith column of the estimated mixing matrix, $\hat{\mathbf{a}}_i$, represents the relative weights of the ith source estimate, $\hat{\mathbf{s}}_i$. Therefore, to

look for components that have different expressions—on average—between patients and controls a two-sample t-test can be performed on columns of $\hat{\mathbf{A}}$, where one group is represented by the values corresponding to the patients with schizophrenia and the other by the values corresponding to the healthy controls (Calhoun & Adalı, 2009). When a component has weights that are statistically significantly different between the patients and the controls, the component can be seen as describing the functional differences between patients and controls, and is referred to as a possible biomarker of disease. The t-statistics derived from these tests can serve as the basis for the construction of the GDMs discussed further in Section 2.5.

Since fMRI data is of high dimensionality and quite noisy, *i.e.*, is in a space consisting of signal as well as noise, dimension-reduction using PCA is a crucial preprocessing step in order to avoid over-fitting in subsequent analyses. Determining the appropriate order—the size of the signal subspace—for this PCA step using real fMRI data is an active area of research, see *e.g*, (Cordes & Nandy, 2006; Y.-O. Li, Adalı, & Calhoun, 2007; Xie, Cao, Weng, & Jin, 2009; Chen et al., 2010; Yourganov et al., 2011; Hui et al., 2013). Despite its fairly extensive study, the majority of order estimation techniques are appropriate for single subject or single task analyses, which limits their applicability to the joint analysis of multiple tasks, multiset fusion. The reason for this is that these techniques ignore the complementary information between the datasets, which is the main motivation for performing a joint analysis on such data. In this work, we use the order estimated by a modification of the procedure in (Levin-Schwartz, Song, Schreier, Calhoun, & Adalı, 2016). The reason for this is because, to the best of our knowledge, it is the only method that has shown desirable performance for the task of order estimation in both the sample-poor and sample-rich regimes inherent to the fusion of medical imaging data. The technique, named PCA and canonical correlation analysis (PCA-CCA), estimates the size of the signal subspace shared by two datasets, *i.e.*, the common order, through a sequence of binary hypothesis tests. The process begins by assuming that the common order is 0 and that the null hypothesis is that the current common order is appropriate. If the null hypothesis is rejected, the common order is increased by 1 and the test is repeated, until the null hypothesis cannot be rejected or the common dimension is equal to half the number of subjects, $M/2$. In this application, we estimate the order for each pairwise combination of tasks and then select the highest estimated order to enable the retention of the most complementary information across the datasets.

We should note that although there are many ICA algorithms, in this work, we used the ICA by entropy bound minimization (EBM) algorithm, due to the fact that it has shown superior performance in both simulated and real neurological data when compared with the popular Infomax algorithm (Bell & Sejnowski, 1995), see *e.g.*, (X.-L. Li & Adalı, 2010; Adalı, Anderson, & Fu, 2014; Adalı et al., 2015). This improved performance derives from

the fact EBM does not assume a fixed form for the distribution of the latent sources and instead attempts to upper bound their entropy through the use of measuring functions (X.-L. Li & Adalı, 2010). The use of these measuring functions enables the description of a wide variety of distributions, including those that are: unimodal, bimodal, symmetric, and skewed (X.-L. Li & Adalı, 2010), thus, generally, improving the estimation of all sources within the mixture.

2.4 IVA

FMRI data from subjects performing multiple tasks has been increasingly gathered during the same study, since each task is expected to provide related information regarding neural function and how it is impacted by neurological diseases, such as schizophrenia, see *e.g.*, (Calhoun et al., 2006; Fitzgerald et al., 2008; Michael et al., 2009; Mijovic et al., 2012; Ramezani et al., 2015). In order to address this scenario, we extend the model in (1) to K datasets, or tasks, as

$$\mathbf{X}^{[k]} = \mathbf{A}^{[k]}\mathbf{S}^{[k]}, \quad 1 \leq k \leq K. \tag{3}$$

Due to the scaling and permutation ambiguities inherent to ICA, running a separate ICA individually for each task and aligning the results is both impractical and suboptimal, since it would ignore the complementary information that each task provides (Lahat et al., 2015). This motivates the performance of a joint analysis, such as using IVA, a recent multiset extension of ICA, which exploits similarities across datasets to achieve a successful decomposition (J.-H. Lee et al., 2008).

The mutual information cost function for IVA can be written as,

$$\mathcal{I}_{\text{IVA}}(\mathbf{W}^{[k]}) = \sum_{c=1}^{C} H\left(\hat{\mathbf{S}}_c\right) - \sum_{k=1}^{K} \log\left|\det\left(\mathbf{W}^{[k]}\right)\right| - \sum_{k=1}^{K} H(\mathbf{X}^{[k]}), \tag{4}$$

where $\hat{\mathbf{S}}_c = \left[\hat{\mathbf{s}}_c^{[1]}, ..., \hat{\mathbf{s}}_c^{[K]}\right]^{\mathrm{T}} \in \mathbb{R}^{K \times V}$ is the cth source component matrix (SCM), formed by concatenating the cth estimated component from each of the datasets. Note that the difference between (2) and (4) is that we are now minimizing the mutual information between SCMs and not components. This cost function reduces to the sum of separate ICAs performed on each dataset individually if the sources in each of the datasets are mutually independent of each other (Adalı et al., 2014). In this work, IVA via a multivariate Gaussian and Laplacian algorithm (Ma, Calhoun, Phlypo, & Adalı, 2014) is used, since it has proven to be an effective IVA algorithm for analysis of fMRI data, see *e.g.*, (Ma et al., 2014; Gopal et al., 2016).

2.5 GDMs

Before describing how we construct GDMs, we should note that, since an important focus of this study is differentiating patients with schizophrenia from healthy controls, we focus on how GDMs can highlight the brain regions that differentiate between two groups. However, as we describe at the end of this section, GDMs can also be constructed such that they reflect the regions that are correlated with certain behavioral variables of interest, highlighting associations between psychological tests and neural function.

Due to the scaling ambiguity of ICA, prior to the construction of GDMs, Z-scores are independently computed from the individual biomarkers using the mean and standard deviation computed using all voxels. Let N be subset of the C components estimated using separate ICAs or IVA for each dataset with weights that are statistically significantly different, at $p < 0.05$, between patients and controls. Denoting these components as $\hat{\mathbf{s}}_n^{[k]}, 1 \leq n \leq N$, we construct the GDM for that method and dataset, $\hat{\mathbf{s}}_{\text{GDM}}^{[k]}$, as follows

$$\hat{\mathbf{s}}_{\text{GDM}}^{[k]} = \sum_{n=1}^{N} \frac{t_n}{\sum_{m=1}^{N} t_m} \hat{\mathbf{s}}_n^{[k]}, \quad 1 \leq k \leq K, \tag{5}$$

where t_n is the t-statistic, calculated as described in Section 2.3, for the nth component that is positive or made to be positive by multiplying the corresponding subject covariation and component by -1. Note that due to the sign ambiguity associated with ICA and IVA, multiplying both the component as well as the subject covariation by -1 will produce the same solution (Hyvärinen, Karhunen, & Oja, 2001). The physical interpretation of such a transformation is that, for example, rather than having patients have increased activation in a certain region over controls, controls have increased deactivation over patients. The GDM can be seen as a summary map that describes only the regions that activate/deactivate significantly differently between patients and controls for a given decomposition and dataset within that decomposition. Each biomarker is scaled by the value of its corresponding t-statistic, so has more weight if the component is better able to differentiate between patients and controls. Note that though we describe the construction of GDMs for factors extracted using ICA and IVA, they could also be constructed using the results of any factor analysis, such as PCA, NMF, or DL. Additionally, though beyond the scope of this work, GDMs can also be defined based upon regression models, which are useful when there are less clearly defined groups. Though we should note that using general linear models where the regressor is the disease state of the subjects are, generally, less robust than factor models in this case (Calhoun, Adalı, Stevens, Kiehl, & Pekar, 2005).

We can quantify the discriminative power of a GDM, and thus, indirectly, the whole decomposition, by generating component weights in a nearly identical manner to the GDM spatial maps and performing a two-sample t-test

on the resulting weights. The corresponding component weights for the GDM are given by

$$\hat{\mathbf{a}}_{\text{GDM}}^{[k]} = \sum_{n=1}^{N} \frac{t_n}{\sum_{m=1}^{N} t_m} \hat{\mathbf{a}}_n^{[k]}, \quad 1 \leq k \leq K, \tag{6}$$

where $\hat{\mathbf{a}}_n^{[k]}$ is the subject covariation, obtained using ICA or IVA, corresponding to $\hat{\mathbf{s}}_n^{[k]}$ and, just as for $\hat{\mathbf{s}}_n^{[k]}$, is multiplied by -1 if the sign of t_n is negative. This construction is expected to result in component loadings that show greater discriminatory power than the original $\hat{\mathbf{a}}_n^{[k]}$, since $\hat{\mathbf{a}}_{\text{GDM}}^{[k]}$ is constructed from only those component loadings that are statistically significant. Additionally, the $\hat{\mathbf{a}}_n^{[k]}$ can be modeled as unit step functions with weights corresponding to their disease status and corrupted with Gaussian noise, where higher values of t_n imply lower variances for the noise. Therefore, the summation of multiple $\hat{\mathbf{a}}_n^{[k]}$, weighted as in (6), will tend to result in a lower value for the variance of the noise and, therefore, a more discriminative subject covariation. Additionally, note that the t-tests performed on the columns of the mixing matrices may be seen as a feature extraction step, thus corrections for multiple comparisons should only account for the number of GDMs calculated.

A fundamental goal of neuroimaging studies is to understand the underlying neural basis for psychological measurements (Shen et al., 2014; Wan et al., 2014; Meng et al., 2017). GDMs lend themselves to this goal in a straightforward manner and through their application can facilitate the discovery of which neuroimaging datasets most effectively explain the results of psychometric tests. This extension can be done by using the t-statistics derived from the correlations of the subject weights with the behavioral variables of interest, rather than those derived using the two-sample t-test. This means that the functional form of (5) and (6) would not change, instead t_n would be defined based on the sample correlation and $\hat{\mathbf{a}}_n^{[k]}$ would be selected based on the significance of their correlation to the test score of interest. In this study, the behavioral variables were derived from the measurement and treatment research to improve cognition in schizophrenia (MATRICS) consensus cognitive battery (MCCB), which is widely recognized as a valuable tool that provides a comprehensive evaluation of cognitive function of schizophrenia within clinical trials (Nuechterlein et al., 2008; August, Kiwanuka, McMahon, & Gold, 2012; Sui et al., 2015).

3 Results and Discussion

3.1 Detection of Group-level Differences

In order to investigate the ability of a GDM to summarize the discriminatory power of a decomposition, we estimate 25 components, corresponding to

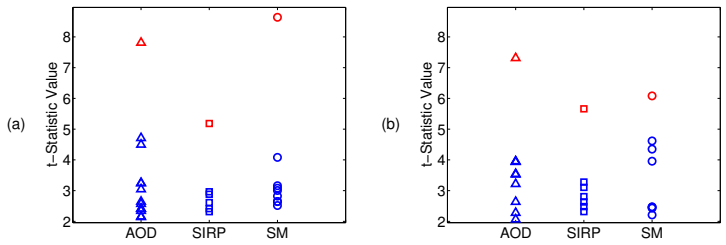

Figure 1: Comparison of the t-statistics from the GDMs compared with the t-statistics of the original biomarkers found using (a) separate ICAs and (b) IVA for each task. The t-statistics of the GMDs are shown in red, while the original biomarkers are shown in blue.

the maximum order that we find by applying PCA-CCA to the pairwise combinations of the three tasks. We display the t-statistics of the subject weights that were found to be statistically significant for each dataset and the corresponding GDMs for the separate ICAs and IVA in Figure 1.

From Figure 1, we see that the t-statistics from the GDMs are an upper-bound for the t-statistics from their constituent biomarkers. This seems to imply that the t-statistics from the GDMs may summarize the discriminative power of a decomposition of a given dataset and can therefore be used to compare the performance of different analysis techniques on the same data. We also note from Figure 1(a) that the AOD and SM datasets seem to provide more discriminatory information than the SIRP dataset, since the significance of the t-statistic the GDM from the SIRP dataset is much lower than those of the AOD and SM datasets. In Figure 1(b), we see that when the datasets are combined in a joint analysis using IVA, the significance of the SIRP dataset increases, while the significance levels of the AOD and SM datasets decrease. This seems to indicate that through a joint analysis of the multi-task data, datasets that are not as discriminative by themselves, like the SIRP, can be made more discriminative if exploiting related information across different tasks. However, we do note a drop in significance for the two datasets that are more significant individually, namely the AOD and the SM.

We next investigate whether the trends that we note using the t-statistics of the GDMs are accurate and that no spurious relations are introduced. This is done by comparing the p-value corresponding to the t-statistic from the GDM with the p-value obtained using Hotelling's T^2-test (Hotelling, 1947). Hotelling's T^2-test, the multivariate counterpart to the two-sample t-test assess the validity of the hypothesis that a two multivariate distributions have the same mean vectors against the alternative hypothesis that they do not. For this application, we evaluate whether the vector containing the subject loadings of all of the biomarkers from the controls is different from the vector containing the corresponding subject loadings from the patients. The results of this comparison are shown in Table 1.

	AOD (ICA)	AOD (IVA)	SIRP (ICA)	SIRP (IVA)	SM (ICA)	SM (IVA)
GDM	1.65×10^{-13}	3.74×10^{-12}	4.54×10^{-7}	4.20×10^{-8}	7.52×10^{-16}	4.53×10^{-9}
T^2-test	8.19×10^{-10}	2.48×10^{-9}	3.95×10^{-5}	3.06×10^{-6}	1.15×10^{-14}	1.00×10^{-6}

Table 1: Significance in terms of p-values for each combination of dataset and analysis method measured using a 2-sample t-test run on the subject loadings of the GDM as well as Hotelling's T^2-test. Note that the trends of the significance of each combination of dataset and analysis method is consistent across the two tests.

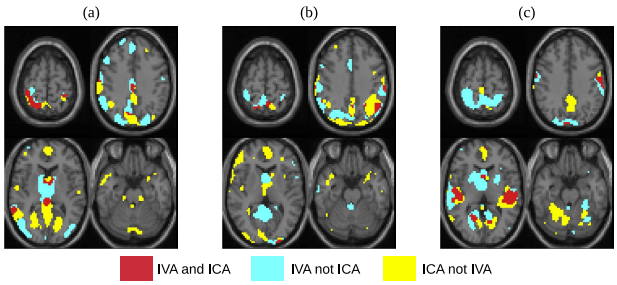

Figure 2: GDMs generated from the (a) AOD, (b) SIRP, and (c) SM tasks. For each plot, only those voxels that had an absolute Z-score above 1.96 are displayed. The red shows regions that are common to both the GDM extracted from the IVA results as well as the GDM extracted from the separate ICAs. The cyan shows the regions that are in the GDM extracted from IVA results but not in the GDM extracted from the separate ICAs. The yellow shows the regions that are in the GDM extracted from separate ICAs but not in the GDM extracted from the IVA results.

Based on the results shown in Table 1, we can see that the relative significance for the t-statistics of the GDMs and the Hotelling's T^2-test is the same. This seems to imply that the GDMs are effectively summarizing the discriminative power of a dataset within a decomposition and not introducing any artificial relations. The reason for this is due to the fact the Hotelling's T^2-test is the classical measure of the total difference between the subject covariations associated with the patients and those associated with the controls. However, it is important to note that the advantage of summarizing the discriminatory power through the use of GDMs rather than through the use of Hotelling's T^2-test is that the GDMs, *i.e.*, the spatial maps, themselves can be related. This enables the identification of regions that differentiate patients and controls in different decompositions and a comparison of these regions. In order to explore the differences between individual ICAs and IVA in terms of the regions that are found to differentiate between patients and controls, we compare the GDMs derived from each of the methods for each of the three tasks after thresholding the maps at $|Z| > 1.96$. Using such a technique, we can visualize how much joint discriminative information each region possesses across tasks. We display the results in Figure 2.

Figure 2 indicates that the discriminatory regions found using IVA but not when using separate ICAs—shown in cyan—correspond, mainly, to the

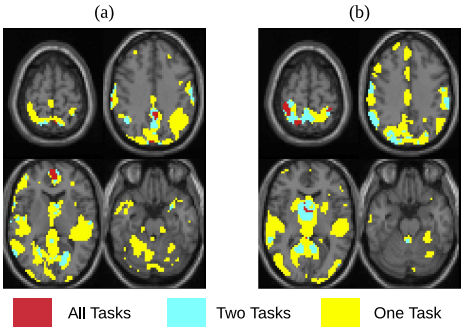

Figure 3: GDMs generated from (a) indivudal ICAs and (b) IVA. For each plot, only those voxels that have an absolute Z-score above 1.96 are displayed. The red shows regions that are common to all three tasks, the number of these voxels is 252 for separate ICAs and 286 for IVA. The cyan shows the regions that two of the three tasks, the number of these voxels is 1,420 for separate ICAs and 2,123 for IVA. The yellow shows the regions that are only activated in one of the three tasks, the number of these voxels is 9,469 for separate ICAs and 8,376 for IVA.

sensorimotor region of the brain. Such a result seems reasonable, since this region is expected to be consistently activated across the three tasks and have been shown to differentiate patients with schizophrenia from healthy controls in these tasks, see *e.g.*, (Liu et al., 2009; Sui et al., 2010; Deserno, Sterzer, Wüstenberg, Heinz, & Schlagenhauf, 2012). We also see regions, such as some of the auditory activation in the AOD and SM datasets as well as the visual activation in the SIRP, which are found using separate ICAs, but not when performing a joint analysis using IVA—shown in yellow. This highlights the ability of a joint analysis using IVA to emphasize similarities across datasets and also highlights the need to compare the results of a joint analysis to those of single dataset analyses, since some of the regions associated with a subset of the tasks may not be as strongly emphasized using IVA (Svensén, Kruggel, & Benali, 2002). Additionally, we can investigate the extent to which IVA emphasizes similar regions across all tasks when compared with individual ICAs by looking at the regions in the GDMs that are found to activate at $|Z| > 1.96$ for all tasks. Such an analysis is important, since a fundamental reason for analyzing multi-task fMRI data is to determine the common information provided by the tasks, which may be useful for identifying dysfunctional regions associated with brain disorders (Plis et al., 2014). The results of this comparison are presented in Figure 3.

From Figure 3, we can see that the number of voxels that are common to two or more tasks is higher when analyzing the data using IVA than when separate ICAs are performed on each of the datasets. Such a result seems reasonable, since IVA is able to exploit the joint information across tasks, whereas individual ICAs cannot. This again highlights the improved power of IVA over separate ICAs when there are similarities across datasets.

	AOD (ICA)	AOD (IVA)	SIRP (ICA)	SIRP (IVA)	SM (ICA)	SM (IVA)
t-statistic	7.81	7.31	5.18	5.66	8.64	6.08
Speed of Processing	4.25×10^{-7}	2.50×10^{-5}	5.91×10^{-3}	7.86×10^{-6}	1.02×10^{-5}	9.14×10^{-4}
Verbal Learning	1.06×10^{-6}	6.45×10^{-6}	2.63×10^{-4}	3.03×10^{-6}	4.70×10^{-9}	1.73×10^{-7}
Verbal Working Memory	7.04×10^{-4}	1.23×10^{-3}	3.03×10^{-3}	1.95×10^{-5}	1.40×10^{-4}	5.19×10^{-4}
Nonverbal Working Memory	5.54×10^{-6}	7.56×10^{-5}	3.58×10^{-3}	7.99×10^{-5}	6.73×10^{-8}	7.10×10^{-5}

Table 2: Estimated t-statistics derived from the subject loadings of the GDMs as well as significance of the correlation between the subject loadings for the GDMs and MCCB scores. No correction for multiple comparisons was performed, however we highlight the entries that are not significant after a Bonferroni correction—$p > 2.02 \times 10^{-3}$ for a p-value of 0.05—in red. Note that the only significance values that do not survive the Bonferroni correction, and the lowest t-statistic value, are those corresponding to the ICA decomposition of the SIRP dataset.

3.2 Association with Psychological Measures

In order to investigate the ability of a GDM to summarize the association of components within a decomposition with a behavioral variable of interest, we display the t-statistics corresponding to the Pearson correlation between the subject weights and MCCB scores for the GDMs as well as its constituent components and display the results in Figure 4.

From Figure 4, we see that, just as they may be seen to summarize the discriminatory power of a decomposition, GDMs can be used to summarize the associative power between components within a decomposition and behavioral variables of interest. Using this fact, we explore the use GDMs to assess the value of performing a joint analysis on multi-task fMRI using IVA rather than performing individual analyses using ICA. We present these results in Table 2.

From Table 2, we see that the GDM found using ICA on the SIRP dataset does not correlate significantly, after a Bonferroni correction, with the majority of behavioral scores. This means that the results of a unimodal factorization of this dataset cannot be strongly associated with the measured clinical outcomes. However, the GDM created from the IVA decomposition of that dataset does correlate significantly, even after a conservative correction for multiple comparisons. This trend is similar to the one observed in Figure 1, *i.e.*, that the factors extracted from datasets, which may not, inherently, be strongly associated with behavioral variables of interest, using a joint analysis can have a much stronger association with the behavioral variables. However, this may come at the cost of reducing the associative power of datasets that are strongly associated with that variable by themselves, in this case the AOD and SIRP datasets. This motivates estimating the amount of joint information that exists between fMRI tasks prior to performing an analysis, see *e.g.*, (Löfstedt & Trygg, 2011; Alter, Brown, & Botstein, 2003; van Deun, Smilde, Thorrez, Kiers, & van Mechelen, 2013; Levin-Schwartz et al., 2016), as well as comparing the results of a joint analysis with those from individual analyses (Svensén et al., 2002). It is

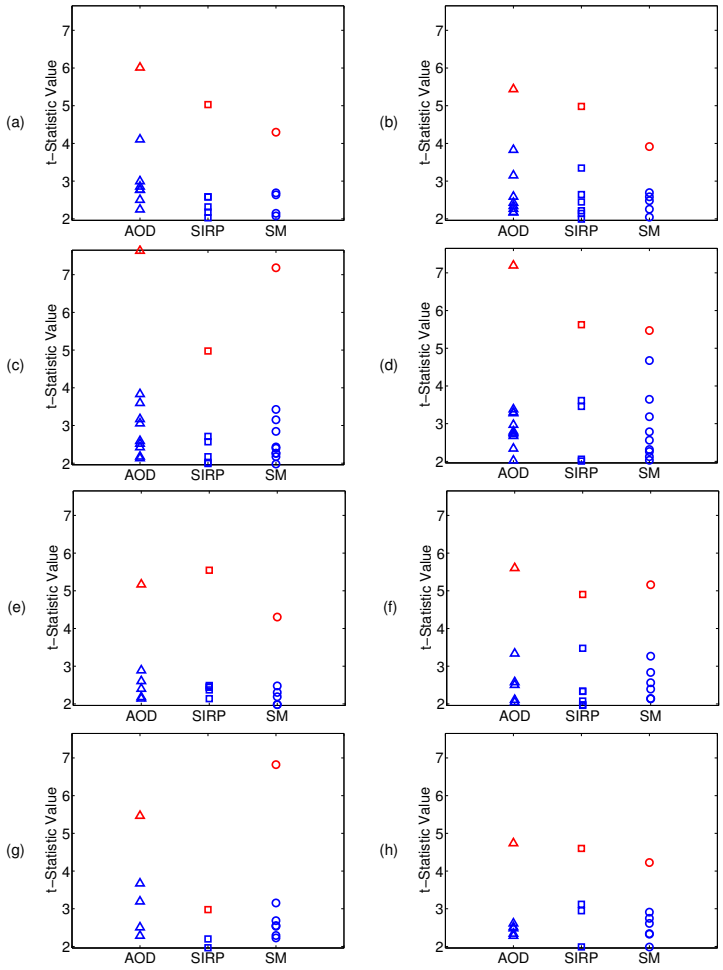

Figure 4: Comparison of the t-statistics derived from the correlation between the subject weights and MCCB scores from the GDMs compared with those from the original components. The t-statistics of the GMDs are shown in red, while the original components are shown in blue. (a) and (b) show this relationship for the speed of processing score for ICA and IVA, respectively. (c) and (d) show this relationship for the verbal learning score for ICA and IVA, respectively. (e) and (f) show this relationship for the verbal working memory score for ICA and IVA, respectively. (g) and (h) show this relationship for the nonverbal working memory score for ICA and IVA, respectively.

important to note that since a subset of the possible contrasts that could be derived from each task were used in this work, the present study should be seen as an initial example to demonstrate the power of GDMs to relate different factorization methods on real fMRI data rather than a general statement about the tasks themselves.

4 Conclusions

The increasing use of fMRI data to study neural function and its disruption due to psychiatric conditions, such as schizophrenia, has lead to the rise of a wide variety of methods to analyze such data. This leads to the issue of how to objectively compare the performance of different methods, without the need for a tedious factor alignment step. In this paper, we have presented a novel technique, GDMs, to relate the results of different fMRI analysis methods based upon either their ability to detect differences between patients and controls or to correlate behavioral variables as well as to highlight which regions differ between the methods. We applied this technique to the decompositions of fMRI data from three tasks derived using individual analyses using ICA as well as a joint analysis using IVA. Our results show that GDMs are an effective technique to compare the results of different factorization methods on real fMRI data. Additionally, through our comparison of the results using individual ICAs and IVA, we find that a joint analysis using IVA is able to determine regions that are more similar across tasks and can be more discriminative between patients and controls than those found using individual analyses using ICA. It is important to note that though we construct GDMs based upon highlighting group differences and associations with behavioral variables, GDMs could also be constructed from the output of a regression analysis, which can account for the effects of confounding variables and be used when there are less clearly defined groups. We should also note that though developed in the context of comparing different decompositions of multi-task fMRI data, GDMs can be applied to data from different neuroimaging modalities, such as electroencephalogram or diffusion tensor imaging, or even in the context of multimodal fusion.

Acknowledgment

This work was supported by the following grants: NIH-NIBIB R01 EB 005846, NSF-CCF 1618551, and NSF 1539067.

References

Abolghasemi, V., Ferdowsi, S., & Sanei, S. (2015). Fast and incoherent dictionary learning algorithms with application to fMRI. *Signal, Image*

and Video Processing, *9*(1), 147–158.

Abraham, A., Dohmatob, E., Thirion, B., Samaras, D., & Varoquaux, G. (2013). Extracting brain regions from rest fMRI with total-variation constrained dictionary learning. In *International conference on medical image computing and computer-assisted intervention* (pp. 607–615).

Adalı, T., Anderson, M., & Fu, G.-S. (2014, May). Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. *IEEE Signal Processing Magazine*, *31*(3), 18–33. doi: 10.1109/MSP.2014.2300511

Adalı, T., Levin-Schwartz, Y., & Calhoun, V. D. (2015). Multi-modal data fusion using source separation: Application to medical imaging. *Proceedings of the IEEE*, *103*(9), 1494-1506.

Afshin-Pour, B., Hossein-Zadeh, G.-A., Strother, S. C., & Soltanian-Zadeh, H. (2012). Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework. *NeuroImage*, *60*(4), 1970–1981.

Alter, O., Brown, P. O., & Botstein, D. (2003). Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms. *Proceedings of the National Academy of Sciences*, *100*(6), 3351–3356.

August, S. M., Kiwanuka, J. N., McMahon, R. P., & Gold, J. M. (2012). The MATRICS consensus cognitive battery (MCCB): clinical and cognitive correlates. *Schizophrenia Research*, *134*(1), 76–82.

Bell, A., & Sejnowski, T. (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129–1159.

Calhoun, V. D., & Adalı, T. (2009, Sep.). Feature-based fusion of medical imaging data. *IEEE Trans. Inf. Technol. Biomed.*, *13*(5), 711–720. doi: 10.1109/TITB.2008.923773

Calhoun, V. D., Adalı, T., Kiehl, K. A., Astur, R., Pekar, J. J., & Pearlson, G. D. (2006). A method for multitask fMRI data fusion applied to schizophrenia. *Human Brain Mapping*, *27*, 598–610.

Calhoun, V. D., Adalı, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, *14*(3), 140–151.

Calhoun, V. D., Adalı, T., Stevens, M. C., Kiehl, K. A., & Pekar, J. J. (2005). Semi-blind ICA of fMRI: a method for utilizing hypothesis-derived time courses in a spatial ICA analysis. *NeuroImage*, *25*(2), 527–538.

Calhoun, V. D., & Allen, E. (2013). Extracting intrinsic functional networks with feature-based group independent component analysis. *Psychometrika*, *78*(2), 243–259.

Chen, S., Ross, T. J., Chuang, K.-S., Stein, E. A., Yang, Y., & Zhan, W.

(2010). A new approach to estimating the signal dimension of concatenated resting-state functional MRI data sets. *Magnetic Resonance Imaging*, *28*(9), 1344–1352.

Cordes, D., & Nandy, R. R. (2006). Estimation of the intrinsic dimensionality of fMRI data. *NeuroImage*, *29*(1), 145–154.

Davidson, I., Gilpin, S., Carmichael, O., & Walker, P. (2013). Network discovery via constrained tensor analysis of fMRI data. In *Proceedings of the 19th acm sigkdd international conference on knowledge discovery and data mining* (pp. 194–202). doi: 10.1145/2487575.2487619

Degras, D., & Lindquist, M. A. (2014). A hierarchical model for simultaneous detection and estimation in multi-subject fMRI studies. *NeuroImage*, *98*, 61–72.

Deserno, L., Sterzer, P., Wüstenberg, T., Heinz, A., & Schlagenhauf, F. (2012). Reduced prefrontal-parietal effective connectivity and working memory deficits in schizophrenia. *Journal of Neuroscience*, *32*(1), 12–20.

Eklund, A., & Nichols, T. (2017). How open science revealed false positives in brain imaging. *Significance*, *14*(1), 30–33.

Engberg, A. M. E., Andersen, K. W., Mørup, M., & Madsen, K. H. (2016). Independent vector analysis for capturing common components in fMRI group analysis. In *2016 international workshop on pattern recognition in neuroimaging (prni)* (pp. 1–4).

Erhardt, E. B., Rachakonda, S., Bedrick, E. J., Allen, E. A., Adalı, T., & Calhoun, V. D. (2011). Comparison of multi-subject ICA methods for analysis of fMRI data. *Human Brain Mapping*, *12*, 2075–2095.

Ferdowsi, S., Abolghasemi, V., & Sanei, S. (2010). A constrained NMF algorithm for BOLD detection in fMRI. In *2010 ieee international workshop on machine learning for signal processing (mlsp)* (pp. 77–82).

Fitzgerald, P. B., Srithiran, A., Benitez, J., Daskalakis, Z. Z., Oxley, T. J., Kulkarni, J., & Egan, G. F. (2008). An fMRI study of prefrontal brain activation during multiple tasks in patients with major depressive disorder. *Human Brain Mapping*, *29*(4), 490–501.

Gollub, R. L., Shoemaker, J. M., King, M. D., White, T., Ehrlich, S., Sponheim, S. R., ... Andreasen, N. C. (2013). The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics*, *11*(3), 367–388.

Gopal, S., Miller, R. L., Michael, A., Adalı, T., Cetin, M., Rachakonda, S., ... Calhoun, V. D. (2016). Spatial variance in resting fMRI networks of schizophrenia patients: an independent vector analysis. *Schizophrenia Bulletin*, *42*(1), 152–160.

Hotelling, H. (1947). Multivariate quality control. *Techniques of Statistical Analysis*.

Hui, M., Li, R., Chen, K., Jin, Z., Yao, L., & Long, Z. (2013, May). Improved estimation of the number of independent components for functional magnetic resonance data by a whitening filter. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 629-641. doi: 10.1109/JBHI.2013.2253560

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. John Wiley & Sons, Inc.

Kiehl, K. A., & Liddle, P. F. (2001). An event-related functional magnetic resonance imaging study of an auditory oddball task in schizophrenia. *Schizophrenia Research*, *48*(2), 159–171.

Kuang, L.-D., Lin, Q.-H., Gong, X.-F., Cong, F., Sui, J., & Calhoun, V. D. (2015). Multi-subject fMRI analysis via combined independent component analysis and shift-invariant canonical polyadic decomposition. *Journal of Neuroscience Methods*, *256*, 127–140.

Kuang, L. D., Lin, Q. H., Gong, X. F., Fan, J., Cong, F. Y., & Calhoun, V. D. (2013, July). Multi-subject fMRI data analysis: Shift-invariant tensor factorization vs. group independent component analysis. In *2013 ieee china summit and international conference on signal and information processing* (p. 269-272). doi: 10.1109/ChinaSIP.2013.6625342

Lahat, D., Adalı, T., & Jutten, C. (2015, Sep.). Multimodal data fusion: An overview of methods, challenges, and prospects. *Proceedings of the IEEE*, *103*(9), 1449-1477. doi: 10.1109/JPROC.2015.2460697

Lee, J.-H., Lee, T.-W., Jolesz, F. A., & Yoo, S.-S. (2008). Independent vector analysis (IVA): Multivariate approach for fMRI group study. *NeuroImage*, *40*(1), 86–109. doi: 10.1016/j.neuroimage.2007.11.019

Lee, K., Tak, S., & Ye, J. C. (2011, May). A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion. *IEEE Transactions on Medical Imaging*, *30*(5), 1076–1089. doi: 10.1109/TMI.2010.2097275

Levin-Schwartz, Y., Calhoun, V. D., & Adalı, T. (2017, July). Quantifying the interaction and contribution of multiple datasets in fusion: Application to the detection of schizophrenia. *IEEE Transactions on Medical Imaging*, *36*(7), 1385-1395. doi: 10.1109/TMI.2017.2678483

Levin-Schwartz, Y., Song, Y., Schreier, P. J., Calhoun, V. D., & Adalı, T. (2016). Sample-poor estimation of order and common signal subspace with application to fusion of medical imaging data. *NeuroImage*, *134*, 486–493.

Li, X.-L., & Adalı, T. (2010, Dec.). Independent component analysis by entropy bound minimization. *IEEE Transactions on Signal Processing*, *58*(10), 5151–5164. doi: 10.1109/TSP.2010.2055859

Li, Y.-O., Adalı, T., & Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*, *28*(11), 1251–1266.

Liu, J., Pearlson, G., Windemuth, A., Ruano, G., Perrone-Bizzozero, N. I., & Calhoun, V. D. (2009). Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. *Human Brain Mapping*, *30*(1), 241–255. doi: 10.1002/hbm.20508

Löfstedt, T., & Trygg, J. (2011). OnPLS - a novel multiblock method for the modelling of predictive and orthogonal variation. *Journal of Chemometrics*, *25*(8), 441–455.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Lohmann, G., Volz, K. G., & Ullsperger, M. (2007). Using non-negative matrix factorization for single-trial analysis of fMRI data. *NeuroImage*, *37*(4), 1148–1160.

Ma, S., Calhoun, V. D., Phlypo, R., & Adalı, T. (2014). Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis. *NeuroImage*, *90*, 196-206.

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. *Human Brain Mapping*, *6*, 160–188.

Meng, X., Jiang, R., Lin, D., Bustillo, J., Jones, T., Chen, J., ... Calhoun, V. D. (2017). Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *NeuroImage*, *145*, 218–229.

Michael, A. M., Baum, S. A., Fries, J. F., Ho, B.-C., Pierson, R. K., Andreasen, N. C., & Calhoun, V. D. (2009). A method to fuse fMRI tasks through spatial correlations: Applied to schizophrenia. *Human Brain Mapping*, *30*(8), 2512–2529. doi: 10.1002/hbm.20691

Mijovic, B., Vanderperren, K., Novitskiy, N., Vanrumste, B., Stiers, P., den Bergh, B. V., ... Vos, M. D. (2012). The "why" and "how" of joint ICA: Results from a visual detection task. *NeuroImage*, *60*(2), 1171-1185.

Moeller, F., LeVan, P., & Gotman, J. (2011). Independent component analysis (ICA) of generalized spike wave discharges in fMRI: Comparison with general linear model-based EEG-fMRI. *Human Brain Mapping*, *32*(2), 209–217.

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., ... Marder, S. R. (2008). The MATRICS consensus cognitive battery, part 1: test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213.

Plis, S. M., Sui, J., Lane, T., Roy, S., Clark, V. P., Potluru, V. K., ... Calhoun, V. D. (2014). High-order interactions observed in multitask intrinsic networks are dominant indicators of aberrant brain function in schizophrenia. *NeuroImage*, *102, Part 1*, 35-48. doi:

http://dx.doi.org/10.1016/j.neuroimage.2013.07.041

Ramezani, M., Marble, K., Trang, H., Johnsrude, I. S., & Abolmaesumi, P. (2015). Joint sparse representation of brain activity patterns in multi-task fMRI data. *IEEE Transactions on Medical Imaging*, *34*(1), 2–12. doi: 10.1109/TMI.2014.2340816

Ramezani, M., Rasoulian, A., Hollenstein, T., Harkness, K., Johnsrude, I., & Abolmaesumi, P. (2014). Joint source based analysis of multiple brain structures in studying major depressive disorder. *Proc. SPIE*, *9034*, 1-6. doi: 10.1117/12.2042275

Rasmussen, P. M., Abrahamsen, T. J., Madsen, K. H., & Hansen, L. K. (2012). Nonlinear denoising and analysis of neuroimages with kernel principal component analysis and pre-image estimation. *NeuroImage*, *60*(3), 1807–1818.

Roels, S. P., Bossier, H., Loeys, T., & Moerkerke, B. (2015). Data-analytical stability of cluster-wise and peak-wise inference in fMRI data analysis. *Journal of Neuroscience Methods*, *240*, 37–47.

Schöpf, V., Windischberger, C., Kasess, C. H., Lanzenberger, R., & Moser, E. (2010, Dec. 01). Group ICA of resting-state data: a comparison. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *23*(5), 317–325. doi: 10.1007/s10334-010-0212-0

Shen, L., Thompson, P. M., Potkin, S. G., Bertram, L., Farrer, L. A., Foroud, T. M., ... Saykin, A. J. (2014). Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. *Brain Imaging and Behavior*, *8*(2), 183–207.

SPM5. (2011). *Statistical Parametric Mapping.* http://www.fil.ion.ucl.ac.uk/spm/software/spm5.

Stone, J. V., Porrill, J., Porter, N. R., & Wilkinson, I. D. (2002). Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. *NeuroImage*, *15*(2), 407–421.

Strother, S. C., Anderson, J., Hansen, L. K., Kjems, U., Kustra, R., Sidtis, J., ... Rottenberg, D. (2002). The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. *NeuroImage*, *15*(4), 747–771.

Sui, J., Adalı, T., Pearlson, G., Yang, H., Sponheim, S. R., White, T., & Calhoun, V. D. (2010). A CCA+ICA based model for multi-task brain imaging data fusion and its application to schizophrenia. *NeuroImage*, *51*, 123-134.

Sui, J., Adalı, T., Pearlson, G. D., Clark, V. P., & Calhoun, V. D. (2009). A method for accurate group difference detection by constraining the mixing coefficients in an ICA framework. *Human Brain Mapping*, *30*(9), 2953–2970.

Sui, J., Pearlson, G. D., Du, Y., Yu, Q., Jones, T. R., Chen, J., ... Calhoun, V. D. (2015). In search of multimodal neuroimaging biomarkers of

cognitive deficits in schizophrenia. *Biological Psychiatry*, *78*(11), 794–804.

Svensén, M., Kruggel, F., & Benali, H. (2002). ICA of fMRI group study data. *NeuroImage*, *16*(3), 551–563.

Thirion, B., & Faugeras, O. (2003). Dynamical components analysis of fMRI data through kernel PCA. *NeuroImage*, *20*(1), 34–49.

van Deun, K., Smilde, A. K., Thorrez, L., Kiers, H. A. L., & van Mechelen, I. (2013). Identifying common and distinctive processes underlying multiset data. *Chemometrics and Intelligent Laboratory Systems*, *129*, 40-51. doi: http://dx.doi.org/10.1016/j.chemolab.2013.07.005

Wan, J., Zhang, Z., Rao, B. D., Fang, S., Yan, J., Saykin, A. J., & Shen, L. (2014). Identifying the neuroanatomical basis of cognitive impairment in Alzheimer's disease by correlation-and nonlinearity-aware sparse Bayesian learning. *IEEE Transactions on Medical Imaging*, *33*(7), 1475–1487.

Wang, Z. (2011). Fixed-point algorithms for constrained ICA and their applications in fMRI data analysis. *Magnetic Resonance Imaging*, *29*(9), 1288–1303.

Welvaert, M., & Rosseel, Y. (2014). A review of fMRI simulation studies. *PloS one*, *9*(7), 1–10.

Xie, X., Cao, Z., Weng, X., & Jin, D. (2009). Estimating intrinsic dimensionality of fMRI dataset incorporating an AR(1) noise model with cubic spline interpolation. *Neurocomputing*, *72*(4), 1042–1055.

Yourganov, G., Chen, X., Lukic, A. S., Grady, C. L., Small, S. L., Wernick, M. N., & Strother, S. C. (2011). Dimensionality estimation for optimal detection of functional networks in BOLD fMRI data. *NeuroImage*, *56*(2), 531–543.

Zhao, S., Han, J., Lv, J., Jiang, X., Hu, X., Zhao, Y., ... Liu, T. (2015). Supervised dictionary learning for inferring concurrent brain networks. *IEEE Transactions on Medical Imaging*, *34*(10), 2036–2045.

Zhao, X., Glahn, D., Tan, L. H., Li, N., Xiong, J., & Gao, J.-H. (2004). Comparison of TCA and ICA techniques in fMRI data processing. *Journal of Magnetic Resonance Imaging*, *19*(4), 397–402.

Zhong, Y., Wang, H., Lu, G., Zhang, Z., Jiao, Q., & Liu, Y. (2009). Detecting functional connectivity in fMRI using PCA and regression analysis. *Brain Topography*, *22*(2), 134–144.