A spatially constrained independent component analysis jointly informed by structural and functional connectivity.

Mahshid Fouladivanda[1], Armin Iraji[1], Lei Wu[1], Vince D Calhoun[1]

[1] Tri-institute Translational Research in Neuroimaging and Data Science (TReNDS Center), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, USA.

Abstract

There are a growing number of neuroimaging studies motivating joint structural and functional brain connectivity models. Connectivity features of different modalities provide insight into brain functional organization by leveraging complementary information, especially in brain disorders as schizophrenia. To this end, we proposed a multimodal ICA model to utilize information from both structural and functional brain connectivity as well as guided by spatial maps to estimate intrinsic connectivity networks (ICNs). Structural connectivity is estimated through whole-brain tractography on diffusion-weighted MRI (dMRI), while functional connectivity information is derived from resting-state functional MRI (rs-fMRI). The proposed structural-functional connectivity and spatially constrained ICA (sfCICA) model optimizes ICNs at the individual level using a multi-objective framework. We evaluated our method using synthetic data and a real dataset (including dMRI and rs-fMRI images from 300 schizophrenia patients and controls). The results demonstrated that our method enhances the functional coupling between ICs with higher structural connectivity in both synthetic and real data. Additionally, the resulting component maps showed improved modularity and enhanced network distinction, particularly in the patients group. Statistical analysis on the functional connectivity networks revealed more significant group differences when comparing the estimated structural-functional connectivity and spatially constrained ICNs with single modality ICNs. In summary, compared to an fMRI only method, the proposed joint approach for estimating ICNs showed multiple benefits from being jointly informed by structural and functional connectivity information. These findings suggest advantages in simultaneously learning from structural and functional connectivity information in brain network studies, effectively enhancing connectivity estimates based on structural connectivity.

Keywords: multi-modal independent components analysis, multi-objective model, diffusion MRI, resting state fMRI, spatial constrained, schizophrenia

Corresponding author: Vince. D. Calhoun (v.calhoun@gsu.edu)

1. Introduction

The human brain can be modeled as a distinct functional unit that exhibit a temporally synchronized pattern known as intrinsic functional networks (ICNs) [1, 2]. These ICNs interact with each other, collectively representing brain function. Consequently, accurate estimation of the ICNs is a critical step in studying brain networks since it helps to minimize the impact of inaccuracies in functional activities [2]. One approach for identifying ICNs and gaining insights into brain networks, including brain disorders like schizophrenia [3-5] is data-driven independent component analysis (ICA) [6, 7]. This method assumes spatial independence for the temporally synchronized functional units and recovers a set of maximally independent sources from a mixture of unknown source signals without any prior information [8], using single modality neuroimaging, particularly resting-state functional magnetic resonance imaging (rs-fMRI).

ICA has become widely used to partition the brain into spatially overlapping and distinct ICNs at the group level. Group ICNs are estimated by using fMRI images of all subjects and then subject level corresponding obtained. The most common approach is to apply ICA to the data, then use back reconstruction to estimate subject specific maps [8-10]. These approaches are fully data-driven; however, more recently, spatially constrained ICA approaches [11, 12] have been proposed to estimate components using a set of template components as priors. They can be used to build a fully automated approach (e.g., the NeuroMark pipeline [13]) that does not require post hoc component selection or network matching [13], and also automatically provides component ordering, providing modular functional network connectivity. These approaches are used in many studies, and offer a fully automated framework that can be integrated within a larger framework (e.g., a containerized version is available from http://trendscenter.org/software/gift, as well as a BIDSapp [14]) and which can be more easily compared across studies.

Current constrained ICA approaches are unimodal and does not leverage information from other modalities, particularly, dMRI images. dMRI images provides information about the physical paths in human brain, captured as structural connectivity (SC) which drive functional activities of different units. In this work we investigate a model that jointly estimates ICNs informed by multimodal data to provide a more complete view of the connectome, that is, one informed by both structural and functional connectivity. The idea is to jointly estimate components such that they are informed by multiple modalities. This has many advantages, for example, a connectivity domain ICA approach [15] called joint connectivity matrix ICA has been used to jointly parcellate structural and functional connectivity data, yielding a data-driven structure/function parcellation [16].

One of the longstanding aspiration in brain research is linking brain function to its underlying architecture which, it made the interplay between structural and functional connectivity (FC) as a fundamental in network neuroscience [17]. There have been a number of studies trying to fill this gap [18, 19], by developing multimodal fusion [20, 21] to incorporate complementary information across modalities into the analysis. Multimodal fusion can provide additional insights into brain structure and function that are impacted by psychopathology, including identifying which structural or functional aspects of pathology might be linked to human behavior or cognition [21] particularly in disorders related to both brain functional and structures as schizophrenia [22]. Moreover, information from different modalities can be jointly analyzed

to investigate ICNs using ICA. This has been done mostly by focusing on linking group-level spatial features, including joint ICA (jICA) [23], parallel ICA (pICA) [24], parallel ICA with reference (pICAR) [25], mCCA+jICA [26], mCCAR+jICA [27, 28], and linked ICA [29], or jointly optimizing for fMRI networks and covarying networks from structural MRI as in parallel group ICA+ICA. However, these approaches have not directly integrated structural and functional connectivity information [30, 31]. To our knowledge, there are limited works using SC to estimate ICNs as in [16] which performed joint structure/functional parcellation. In particular, there are no approaches to inform adaptive ICN estimation using both structural and functional connectivity. Using connectivity information constrained helps to drive the model in a way learning from SC information (real connections among different functional units) and FC (their functional interactions).

In this paper, we thus introduce a novel adaptive ICA model to directly investigate the ICNs by imposing constraints based on two different modalities at individual level. We build on a new multi-modal ICA model jointly embedding structural and functional connectivity. In addition, to increase generalizability of the estimated ICNs, prior spatial information was used in the multi-modal ICA model. Our model is an iterative pipeline to learn simultaneously from both fMRI and dMRI, and ICN spatial maps with the purpose of identifying joint structural-functional independent components. While prior work [16] applied a connectivity based model to perform joint data-driven parcellation of fMRI and dMRI, no approaches have tried to directly link the full spatiotemporal fMRI data with dMRI data in the context of a connectivity constraint at the individual subject level and within a spatially constrained model, thus providing a fully automated approach. The proposed structural-functional connectivity and spatial constrained ICA model (sfCICA) identifies ICNs which are simultaneously optimized to be maximally spatially independent and abused by structural and functional connectivity. Thus, the learning procedure to estimate the subject level ICNs was informed by SC and FC information of each subject. We first validate the approach with a simulation, and then we apply the proposed model to a real dataset, including controls and patients with schizophrenia. Results show that the proposed method can effectively imposed prior information on ICNs while jointly learned from structural-functional connectivity in both simulation and real human brain data. Moreover, the identified ICNs show an increased ability to discriminate healthy controls from schizophrenia.

2. Material and Methods

2.1. Structural-functional connectivity and spatially constrained ICA (sfCICA)

Model definition.

A multimodal ICA (sfCICA) model called structural and functional connectivity and spatial constrained ICA was introduced. Briefly, the sfCICA framework estimates subject-level ICNs imposing structural connectivity weights on TC distances (functional connectivity). The main advantage of the constrained in sfCICA framework is imposed the model to jointly learn from both structural and functional connectivity. Using sfCICA model, subject specific ICNs were estimated by multimodal information and using spatial information for each ICNs from a standard template.

In sum, we introduced a multi objective function as in **Eq.1**, it contains three terms: the first and second terms find independent ICNs which are learned from the spatial template. And the third term, constrains the learning procedure to account for both structural and functional connectivity.

$$\max \quad (J(S_l^k), F(S_l^k), -C(w_l^k)) \qquad (1)$$
$$\text{s. t. .} \left\| w_l^k \right\| \quad = 1$$

The sfCICA model was proposed in a multi-objective framework, aiming to estimate ICNs by optimizing their independence, as maximizing their non-Gaussianity measured by negentropy or kurtosis ($J(S_l^k)$) [32] (**Eq.2**).

$$\max \quad J(S_l^k) = (E[G(S_l^k)] - E[G(v)])^2 \qquad (2)$$
$$\text{s. t.} \left\| w_l^k \right\| \quad = 1$$

Where, S_l^k and w_l^k are respectively the spatial map and time course (TC) of the lth ICN (here we use Neuromark_fMRI_1.0 including 53 ICNs) for k^{th} subject estimated by maximizing the objective in equation (2) which $S_l^k = (w_l^k)^T * X^k$, G is any quadratic function, and v is a random variable.

In addition, similarity of the spatial maps with the prior spatial information was maximized by another cost function ($F(S_l^k)$) in **Eq.3**. So, independence and similarity to the prior spatial information (here a standard template used) were jointly maximized during optimization.

$$\max \quad F(S_l^k) = E[S_l S_l^k] \qquad (3)$$
$$\text{s. t.} \left\| w_l^k \right\| \quad = 1$$

where the S_l denoted spatial map of the lth components in the template.

Furthermore, we incorporate SC information to the model as a constraint on the FC of the estimated TC. To this end, minimizing the structural weighted distances of the paired TC would guide the model. SC weights captured from dMRI can reflect existing physical connectivity between functionally related gray matter regions, this means, more fiber bundles between two brain regions (here is ICNs) caused in more functional correlations and shorter distance between TC [33].

Inspired by [33] an objective function (min $\sum_{jl} M_{lj} \left\| FC_l \ - FC_j \ \right\|^2$) was reconstructed, where M_{lj} is SC weights and FC is correlation between pair ICNs (l, j). Note, FC between ICNs can be maximized by minimizing the distance of the ICNs TC. To this end, we propose the following cost function (**Eq.4**) in addition to the previous ones, to be simultaneously optimized in order to incorporate information from both structure and function as constraints.

$$\min \quad C(w_l \ ) = \sum_{j=1}^{L} M_{lj} \left\| w_l^k - w_j^k \right\|^2 \qquad (4)$$
$$\text{s. t.} \left\| w_l^k \right\| \quad = 1$$

To solve a multi-objective optimization problem, one usually needs to find the Pareto optimal set or its subset and critically evaluate which specific trade-off is more appropriate to the problem under study. In addition, optimizing a linear weighted sum of cost functions is a commonly used method for solving multi-objective optimization problems, which simplifies the problem [34]. In

our study, we employ a linear weighted method to solve the multi-objective problem described in **Eq.1**, allowing us to explore different points along the Pareto front by adjusting the weighting values in the linear weighted sum objective function.

$$\text{Cost} = \alpha J\left(S_i^k\right) + \beta\, F\left(S_i^k\right) - \gamma C\left(w_i^k\right) \qquad (5)$$
$$\text{s. t.}\ \left\|w_i^k\right\| \quad = 1$$

Where, α, β and γ are constant weight values. To ensure fair comparison and prevent the optimization from being biased towards a cost function with a larger magnitude, we normalize the cost functions in **Eq. (2-4)** as described [11] and weights (α, β and γ) are chosen equal (0.33). Then a gradient ascent method was used iteratively to converge an optimal solution.

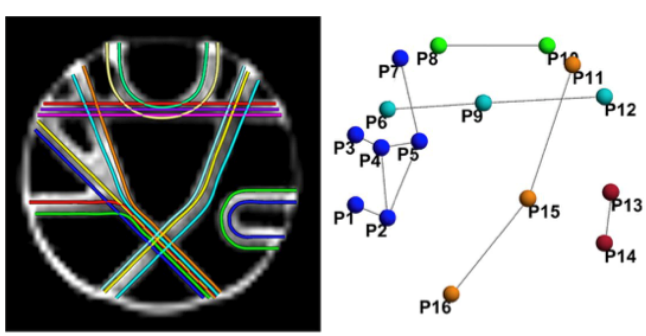
2.2. Evaluation analysis.

To evaluate the benefit of leveraging multi-modal information we compare our model with one of the spatial constrained ICA (CICA) models, known as a multi-objective optimization ICA model [13] which is simplify version of proposed framework when the multimodal information is not being utilized. To this end, we used results from synthetic and real data. More details of the simulation parameters for the synthetic data come in section **2.3.1**.

Then, we performed a set of statistical analysis. Paired t-test analysis was used to determine any significant changes between estimated ICNs from both sfCICA and CICA model by comparing corresponding FC matrices and distances of the TC. Moreover, for group comparison between SZ and HC, we performed generalized regression model (GLM) on FC matrices, and regress out effects of the age, sex, diagnosis, and imaging site effect as covariates for the FC matrices.

2.2.1. Synthetic Data

To evaluate the proposed method, we used a set of simulated functional MRI (fMRI) and diffusion MRI (dMRI) data. Simulated dMRI phantom date was from FiberCup [35, 36] (https://tractometer.org/fibercup/data/), including 3mm isotropic images in 64 uniformly distributed over sphere with b-value = 2000 m/s. It also provided 16 brain regions and fiber pathways among the regions to reconstruct SC, (shown in **Figure 1**). To keep consistence between the simulated dMRI and fMRI data, we used the same 16 predefined FiberCup regions to design synthetic fMRI images. For this purpose, we used the simTB toolbox [37], (https://trendscenter.org/software/simtb/).

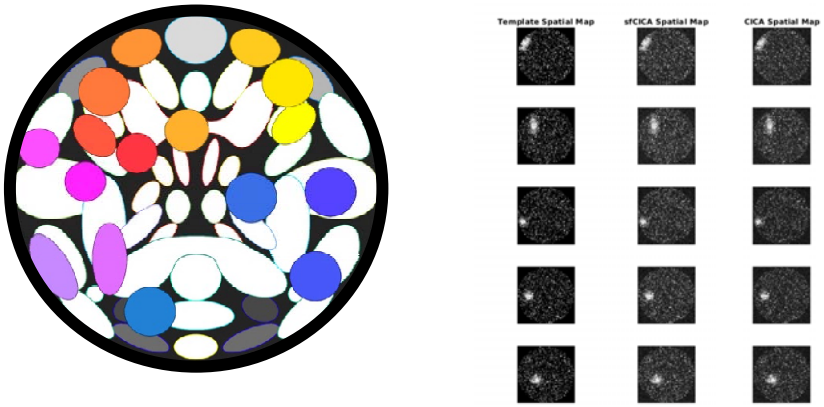


(a) (b)

Figure 1. Simulated FiberCup fiber pathways (a) and corresponding regions (b) for connectivity reconstruction was illustrated.

simTB employs a data generation model that assumes spatiotemporal separability, allowing the simulated fMRI data to be represented as the multiplication of spatial maps (SMs) and TCs. The generated data exhibit realistic dimensions, spatiotemporal activations, and noise characteristics similar to those observed in typical fMRI datasets [37]. Our fMRI simulation consists of 16 components for 30 subjects, which subjects are divided equally in 3 different groups with different design parameters (Group A, B and, C). Generally, all SMs have V = 64 × 64 voxels (V is the number of all voxels) and TCs are T = 3000 time points in length with a TR = 2 seconds. Spatial maps of the components represented in Figure (2), corresponding to the regions in FiberCup data (**Figure 1**).

Groups A, B and C differ in four ways, as follow. First, noise level was chosen differently per groups to determine the robustness of the proposed model. As in group A, mean 1.7, standard deviation 0.2, group B, mean 1.5, standard deviation 0.4 and, group C, mean 1.1, standard deviation 0.35. Then to make diversity among the ICNs from different groups we made some additional changes in their amplitudes and, shapes. Group A has weaker amplitude for ICN 6 than Group B and C and, Group C has weaker for ICN 8 and 11 than group A and B. Amplitudes of the ICNs (g_{ic}), modeled by normal distribution. Distribution of the ICN 6 g_{i6} is normal with mean 2.5, standard deviation 0.3 for group A, and mean 3.5, standard deviation 0.3 for group B. Similar normal distribution with mean 2.5, standard deviation 0.3 for ICN 8 and 11 in group C and mean 3.5, standard deviation 0.3 for group A and B. All the ICNs have different transitions in x and y directions, and rotation to make differences change in the ICNs. Also, ICNs in groups A and C expanded by 1.2 and 1.3 for the group B, contraction by 0.9 applied.



(a) (b)

Figure 2. Spatial maps of 16 ICNs used to generate simulated fMRI are illustrated in (a) and in (b), generated template used for spatial prior information and, estimated spatial maps from sfCICA and CICA methods.

2.3.Data acquisition and preprocessing

In this study rs-fMRI and dMRI images were used from the Function Biomedical Informatics Research Network (fBIRN) phase III datasets [38]. Data collection was at multiple sites of healthy control subjects (HC) or with schizophrenia (SZ) disorder. FBIRN dataset consists of rs-fMRI and dMRI images from 311 adults (age ranging from 18 to 65 years old), including 162

age-gender-matched healthy controls (HC) (Avg ± SD age: 37.0 ± 11.0) and 149 age-gender-matched schizophrenia (SZ) patients (Avg ± SD age: 37.9 ± 11.5).

Acquisition. Images were scanned at seven 3T Siemens Tim Trio System and one 3T GE Discovery MR750 scanner at multiple sites. The acquisition parameters of the rs-fMRI imaging were as follows: FOV of 220 × 220 mm (64 × 64 matrix), TR = 2 s, TE = 30 ms, flip angle (FA) = 77°, 162 volumes, 32 sequential ascending axial slices of 4 mm thickness and 1 mm skip. Subjects had their eyes closed during the resting state scan. For more details see [38]. For the dMRI images, scanning protocols were settled virtually to ensure equivalent acquisitions with 30 directions of diffusion gradient at b = 800 s/mm^2 and five measurements of b = 0 (b0) s/mm^2 [38].

rs-fMRI preprocessing. The rs-fMRI images, as in [39], were preprocessed via the statistical parametric mapping (SPM12, http://www.fil.ion.ucl.ac.uk/spm/) toolbox in MATLAB 2020b. The main steps of the preprocessing including removing first five fMRI scans, motion correction, slice timing correction, image registering to the standard Montreal Neurological Institute (MNI) space, spatial resampling (to 3 × 3 ×3 mm^3 isotropic voxels), spatial smoothing using Gaussian kernel with a full width at half maximum (FWHM) = 6 mm.

dMRI preprocessing. For preprocessing of the dMRI images, first all images were corrected for eddy current and motion distortions using eddy package (FSL 6.0 [40]) by motion induced signal dropout detection based on b0 volumes and replacement method [41]. Then, data quality check was performed by an in-house developed algorithm and visual inspection to exclude images with extreme head motion, signal dropout, or noise level [42, 43].

An additional step was carried out on preprocessed dMRI images to determine white matter fiber tracts for constructing SC. First voxel-wise diffusion tensor was modeled with linear regression under dtifit in FSL toolbox [40]. Next we used the estimated diffusion tensors to perform deterministic tractography in CAMINO toolbox [44] regarding three stopping criteria including: a fixed step size of 0.5mm, anisotropy threshold of 0.2 and angular threshold of 60° [44]. All fiber tracts were extracted in native space. Next, we warped the native fractional anisotropy (FA) maps into the standard MNI space using ANTs [45] and applied inverted spatial normalization to the Neuromark_fMRI_1.0 atlas (including N = 53 ICNs) [13] to obtain the corresponding atlas in native space. ICNs of the normalized Neuromark_fMRI_1.0 atlas used as nodes of the SC network which divided the brain into distinct regions. For each subject a SC network (M) with $N \times N$, was reconstructed by counting the number of streamlines connecting pair of nodes.

3. Results

Using the proposed sfCICA model, we computed 53 multi-modal ICNs for each subject during an iterative optimization procedure. Real data ICNs were also categorized in 7 different functional domains including: subcortical (SC), sensorimotor (SM), auditory (AUD), visual (VS), cognitive control (CC), default mode (DM) and cerebellum (CB) based on the predefined standard ICN template (NeuroMark_fMRI_1.0). Model evolution was based on the output ICNs and corresponding features of them such as FC using synthetic and real data. Analysis represented improvement on estimated ICNs using multi-modal information and supporting evidence of structural connectivity information direct influence on neural activities [46].

3.1. ICNs of the synthetic datasets show a structural and functional learning pattern.

To determine the applicability of the proposed model, we investigate the differences of distances among TC (a measure of the correlation) for both sfCICA and CICA model at individual level for different groups. The TC distance is the distance of an TC with all other remaining TCs of ICNs. Distance analysis on synthetic dataset represented decreasing in distances after performing SC constraints using sfCICA model in **Figure. 3**. Column **(a)** in **Figure. 3.** displays average paired distances differences between sfCICA and CICA model at component level distribution over subjects. Similarly, in all three group of synthetic datasets, negative average of the distance differences occurred in all the components except 14 and, 16 in group B and components 1 in group C (**Fig. 3**). The pink horizontal line is distance equal to zero in **Figure. 3.a**. The distance matrix of the average distance over all subjects is shown in **Figure. 3** column **(b)** and the corresponding group SC is shown in column (**c**).

Statistical analysis, using a paired t-test identified significant differences (p-value <0.05) between the average group distances of the sfCICA and CICA model for all components except for component 14 in group B, and components 1, 3, 7, 8, 11, 13, 14, 15 and, 16 in group C.

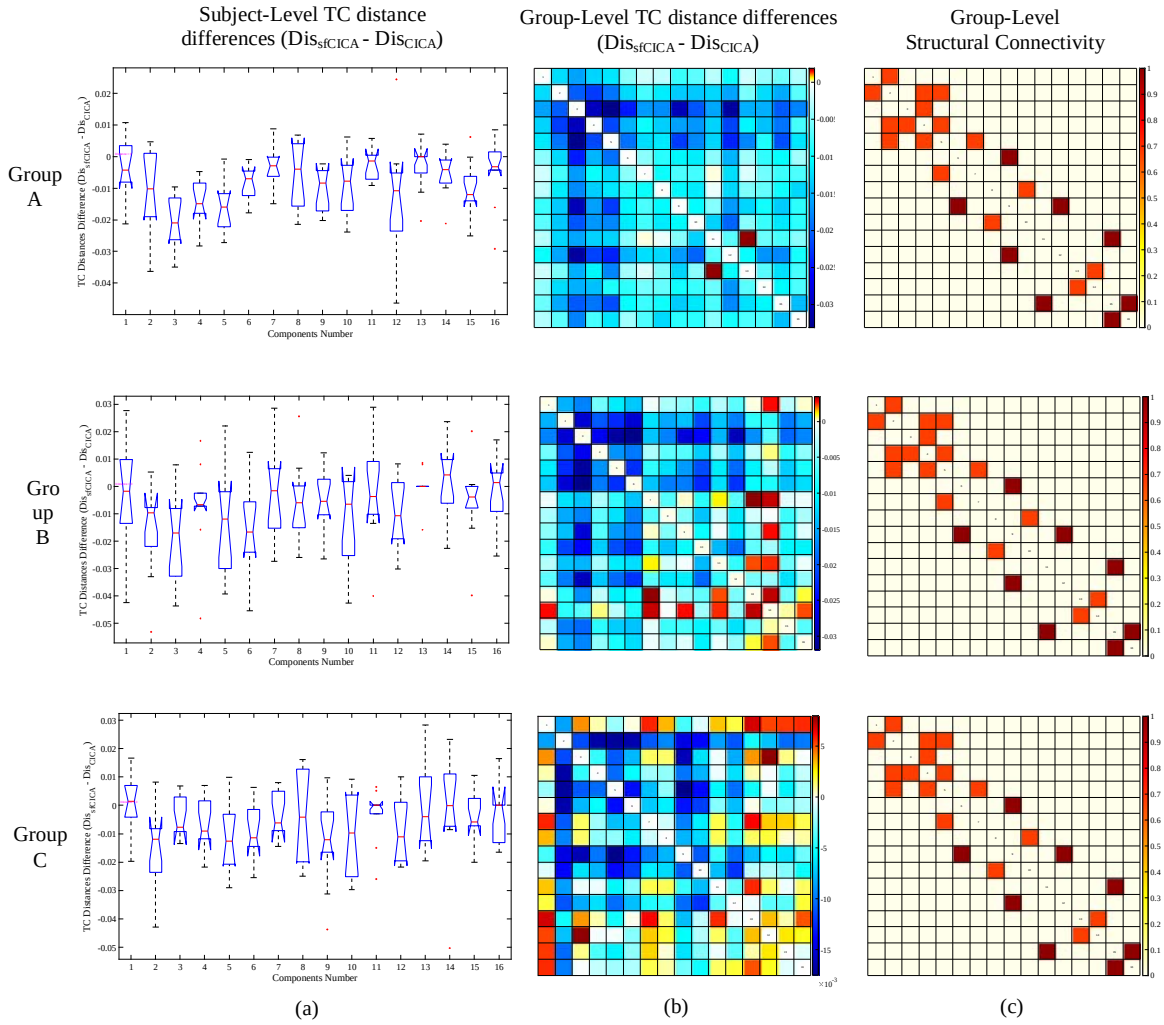


Figure 3. Column (**a**) shows distance differences of the TC are decreased in most of the components estimated by the sfCICA model in comparison to the components of the CICA model using three different synthetic datasets (represented at each row) at subject level. Column (**b**) illustrated the average distance differences of the TC over all

subjects for each component estimated by sfCICA and CICA model using three different group synthetic dataset in each row. Corresponding group level SC shown in the third column (c).

3.2. FC analysis of the synthetic date showed FC is informed by SC weights.

To highlight the effect of the SC information on FC estimated from ICNs using sfCICA model in synthetic data, we compare the similarity of the FCs with SC and residual (Res) FC for the proposed sfCICA with CICA model. We perform two different evaluations by computing the similarity of the FC matrix to determine the degree to which the SC impacted the resulting model weights, shown in **Fig. 4**. For this purpose, the FC matrix of the estimated ICNs from sfCICA were computed by using correlation between paired ICNs for individuals. Then similarity of the FCs was computed (1) with SC and, (2) with residual FC, all in subject-level. Residual FC is two FC matrix of sfCICA and CICA model. Similar procedure performed for the ICNs from CICA model and results from both sfCICA and CICA model were compared. Based on both evaluations, FC of shown more similarity with FC with SC and, residual matrix (Res). Results indicate the proposed sfCICA model is informed from structural information.

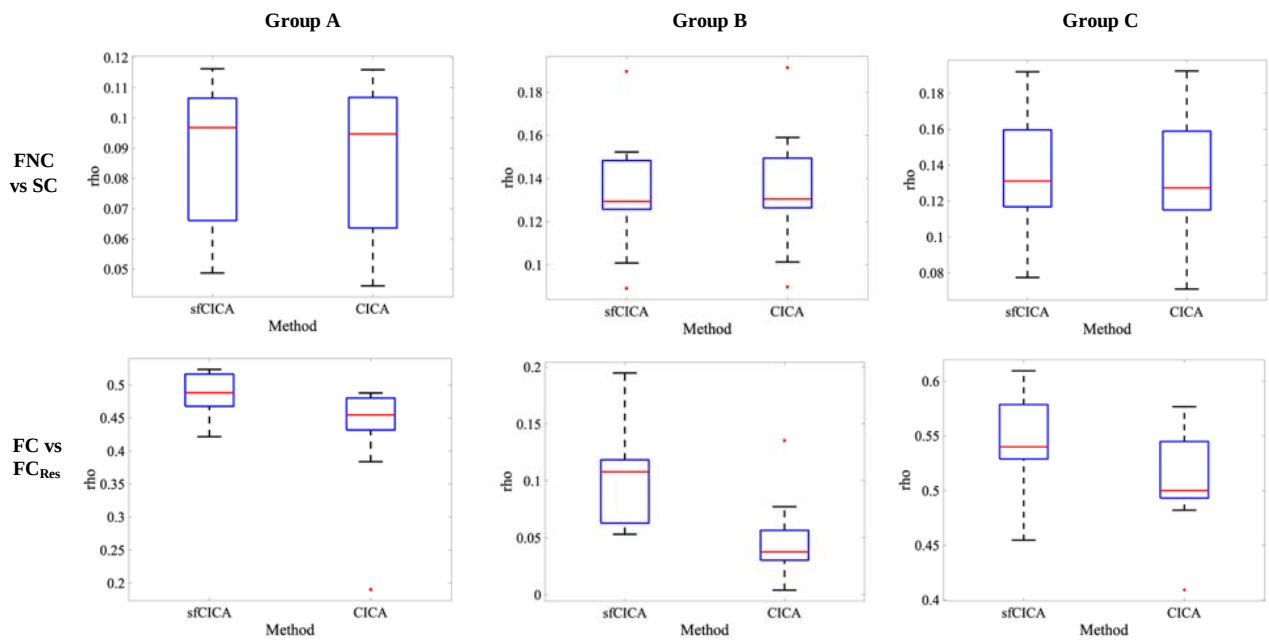


Figure 4. Similarity (Pearson correlation) between FC matrix of estimated ICNs from sfCICA model with SC and, residual (Res) FC matrix are represented in each row, and for three different groups of synthetic data at each column. Results show that similarity (correlation) with both SC and residual FC (FC_{Res}) sfCICA is higher than CICA, as expected.

3.3. Subject-level structural-functional ICNs show more modular and integrated networks.

Using the sfCICA model, ICNs guided by SC and FC, for each subject. In this paper, we used NeuroMark template [13] as prior spatial information, and the included information were evaluated by estimating the spatial similarity between estimated ICNs and corresponding

NeuroMark network template using Pearson's correlation for both sfCICA and CICA model. To analyze the ability of the proposed model for both healthy and disease cases we show the similarity results for each group (HC vs SZ) separately. **Figure 5.a** illustrates increasing spatial similarity for both HC and SZ cohorts for sfCICA model (Avg: 0.52) with respect to the CICA model (Avg: 0.47) using the FBIRN dataset. On average, a 10% increase in similarity occurred for the sfCICA model relative to the spatial similarity estimated from CICA model. Increasing spatial similarity with standard and reproduceable template over multiple subjects represented that informing the optimization process using SC could improve the results.
Moreover, the distance of the TCs between the estimated TC pairs was computed and represented at the individual level in **Figure 5.b** to determine the impact of the SC. It is observed that on average over all components of each subject the estimated distance was decreased compared to the CICA model. in addition, average distance over all subjects and components were decreased. In addition, the average distance of the TC is reduced in comparison to the CICA model in both HC and SZ subjects, that is the networks are more functionally integrated. Here, distance is referred to mean square distances between paired TC.

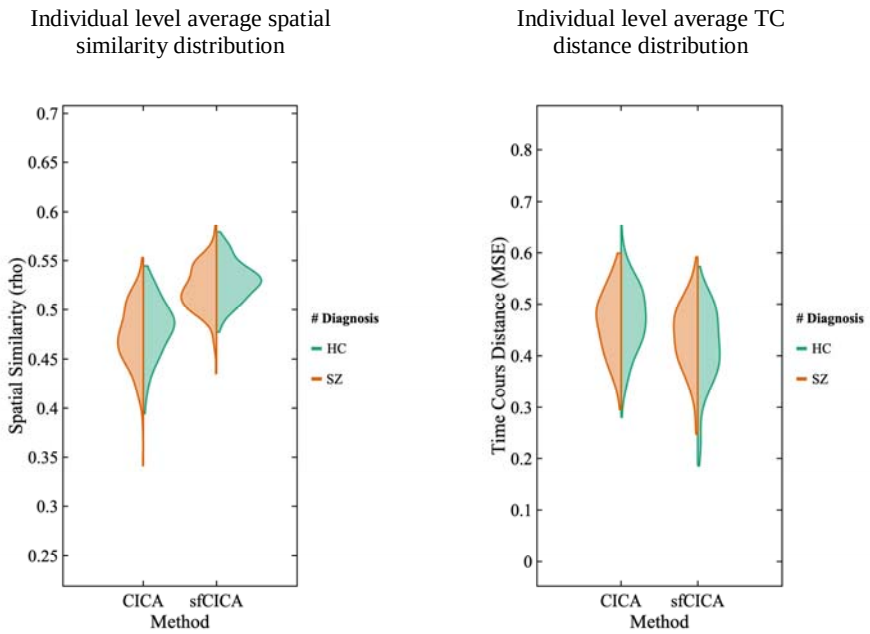


Figure 5. Spatial similarity and TC distance of the ICNs were represented in individual level over all components. In (a) distribution of the spatial similarity (Pearson correlation) with NeuroMark template used for prior spatial information at two HC (green) and SZ (red) groups were illustrated for both CICA and sfCICA model. Mean square distances of each TC from remaining TCs shown in (b), within each HC (green) and SZ (red) group using CICA and sfCICA model. sfCICA showed more spatial similarity and reduced TC distances.

3.4. Subject-level structural-functional ICNs show a structural and functional learning pattern for FBIRN dataset

In **Figure. 6. (a-b)**, the difference between estimated TC distances (a measure of the correlation) corresponding using sfCICA and CICA model was investigated per ICN at individual level. TC distances distribution at individual level of each ICN were separated per group HC/ SZ for both models (sfCICA and CICA). Both HC, and SZ groups have lower distances in sfCICA respect to

the CICA. It is shown in **Figure. 6.a** SZ group distance show lower distances with respect to the HC groups in both models. Thus, negative values in **Figure. 6.b** correspond to lower distances value among the ICNs of sfCICA model respect to the CICA. Interestingly, we observed an increase in negative correlations in the FBIRN dataset (**Figure. 6.b**). In addition, components 18, 29, 37, 40, 4, 47, 48, 51, 52, and 53 show on average positive SC (**Figure. 6.b**), should show increasing distance respect to the distance of the TC from CICA model and as a result a positive average pattern. **Figure. 6** shows that weights of the SC are influencing the distance of the TCs, i.e., where the SC is weak the distance is increased and stronger SC results in reduced distance among the fMRI TCs relative to the CICA model. Results suggest both structural weights and functional spatial maps impose complementary information on the estimated ICNs.

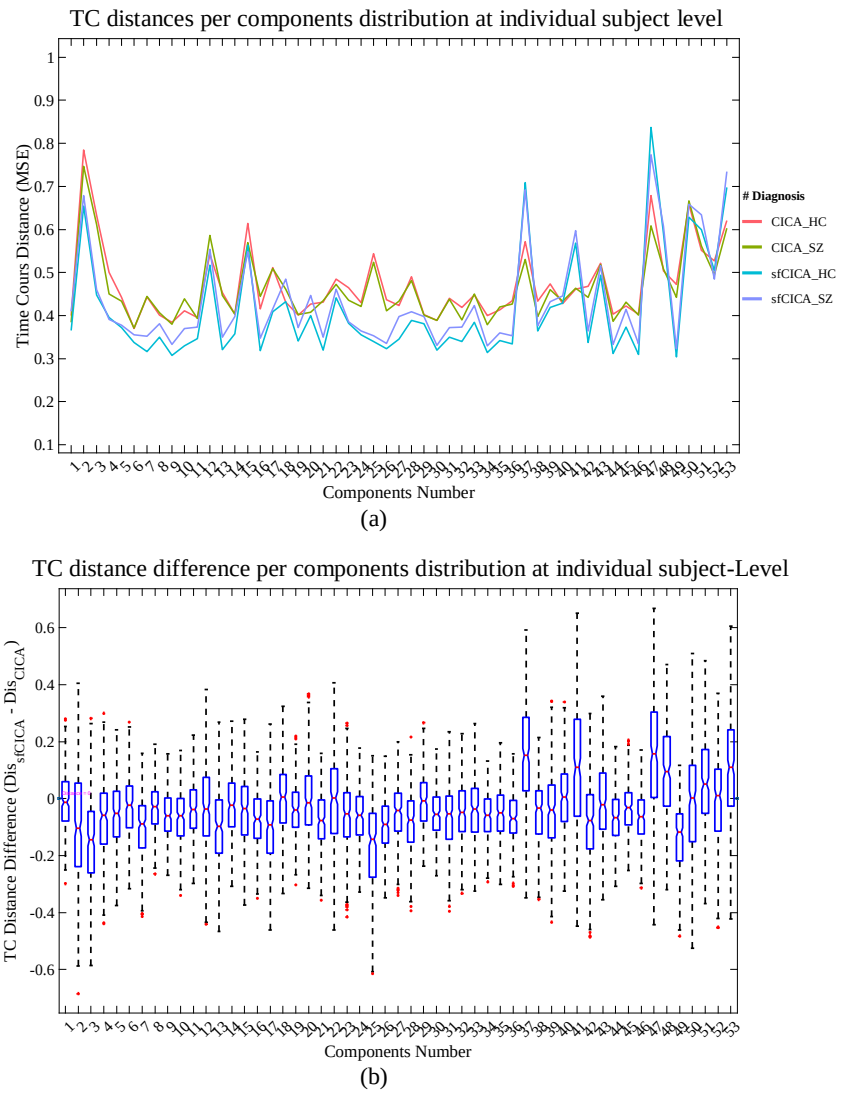


Figure 6. Distance differences of the TCs are decreased in most of the components estimated by sfCICA model in comparison to the components of the CICA model using FBIRN dataset at subject level, represented in (a). And (b) illustrated the average distance of the TC over all subjects for each component estimated by sfCICA and CICA model using the FBIRN dataset.

Moreover, regarding **Figure. 7.a**, it is observed that components with positive differences have high average distance in the CICA model. In other words, a similar pattern for the average distance is evident in both sfCICA and CICA model. By comparing both matrices in **Figure. 7**, we understand that the SC weights constraints helped to separate the ICNs by maximizing the distances. Further, analysis of the paired t-test between components over both models, components number 1, 6, 11, 12, 14, 15, 18, 20, 22, 29, 39, 40, 43, 45, 50, 52 represent nonsignificant changes among 53 components.

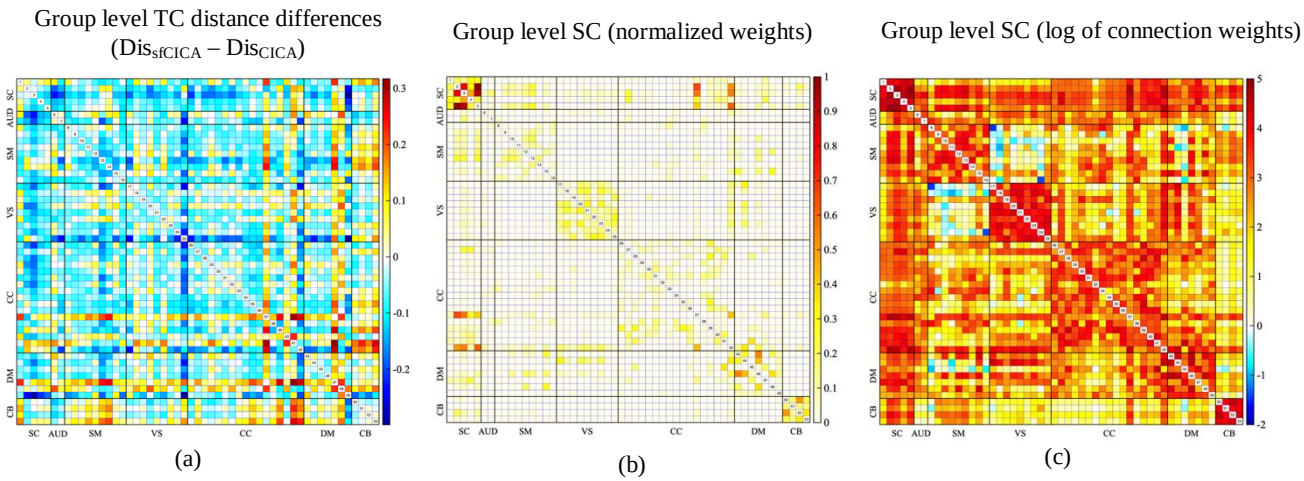


Figure 7. Distance differences of the TCs are decreased in most of the components estimated by sfCICA model in comparison to the components of the CICA model using FBIRN dataset at subject level and significant differences observed (P-value<0.05). (**a**) illustrated the differences between average pair-wise TC distances computed from sfCICA and CICA model using FBIRN dataset. Group level SC is represented in (b)-(c), in (b) normalized connection weights and (c) logarithmic scale ($log10(connection\ weight)$), weight of a connection is number of the tracts connecting two ICN.

4. Distinguished spatial patterns between estimated ICNs from the sfCICA and CICA model.

Importantly, we observed that some functional networks represented by IC 2, 22, 42, 7, 11, 21, 13, 14, 34, 37, 50,3, 27, 39 and, 42, show more modular and defined spatial maps in most subjects particularly in schizophrenia. This shows that the SC constraints are helping us to better distinguish the healthy and patient subjects in these ICNs. These functional networks mostly included subcortical, sensory networks (sensorimotor and, visual), default mode and cerebellum. **Figure. 8** shows three different components for subjects (S) 29, 30, 37 and 74 at each row. All the component's values were converted to z-scores and threshold by 3 (z-score>3). We also qualitatively observe more modular networks with respect to the CICA model, **(see Figure 8).**

Moreover, we performed paired t-test analysis between sfCICA and CICA using the spatial maps within each group (HC / SZ) separately. In healthy subjects, significant differences (p-value<0.05) were observed in most components except in components 45 and 51, in schizophrenia only components 9, 27 and, 51 did not show significant differences.

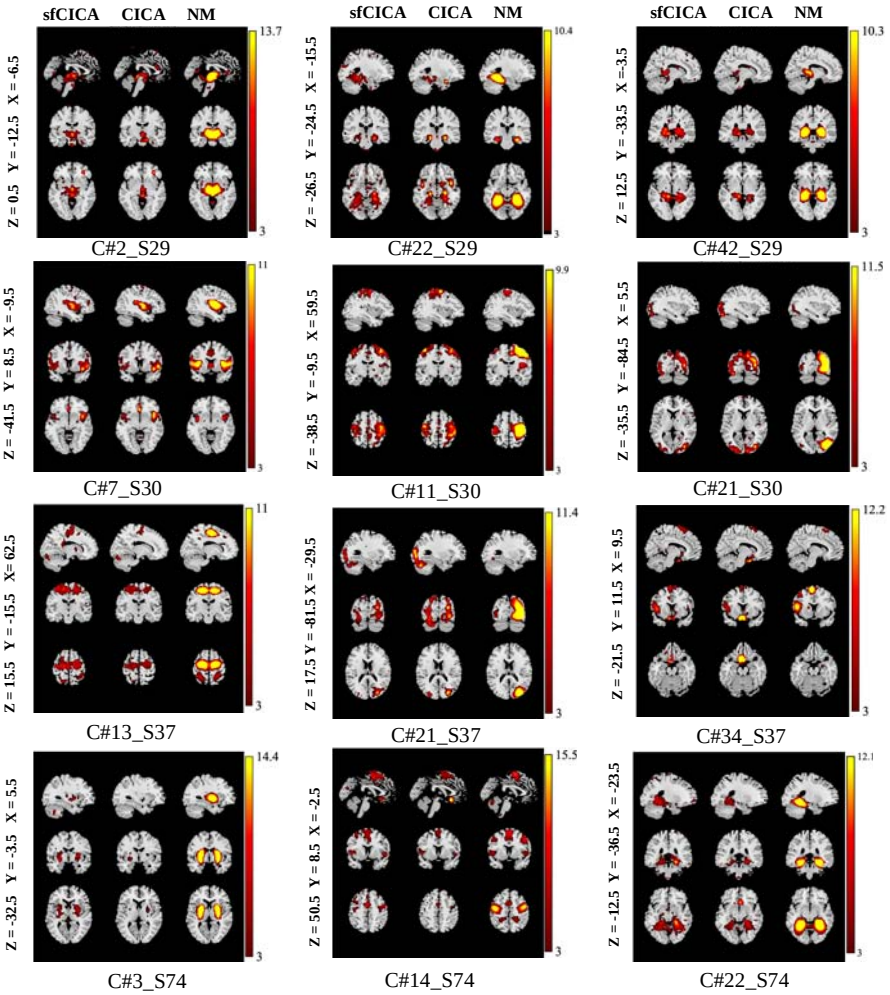


Figure 8. Each panel shows spatial maps for example ICNs (z-score>3) for sfCICA, CICA, and the NeuroMark template. sfCICA generally showed more modular characteristics in different subjects relative to the CICA model.

Paired t-test analysis was also performed on FC matrix to estimate the significancy of the computed ICNs using the proposed method respect to the CICA method. For this purpose, we first estimate the FC matrix for each subject using the Pearson correlation among the ICNs (TCs), in both sfCICA and CICA method. Estimated t-value (**Figure. 9**) represented significant differences in ICNs related to subcortical network and some ICNs in SM, CC, VS and CB network. Interestingly decision results of the hypothesis in paired t-test showed all the estimated ICNs rejected the null hypothesis, except one of them and the proposed model results were significantly different from the CICA model.

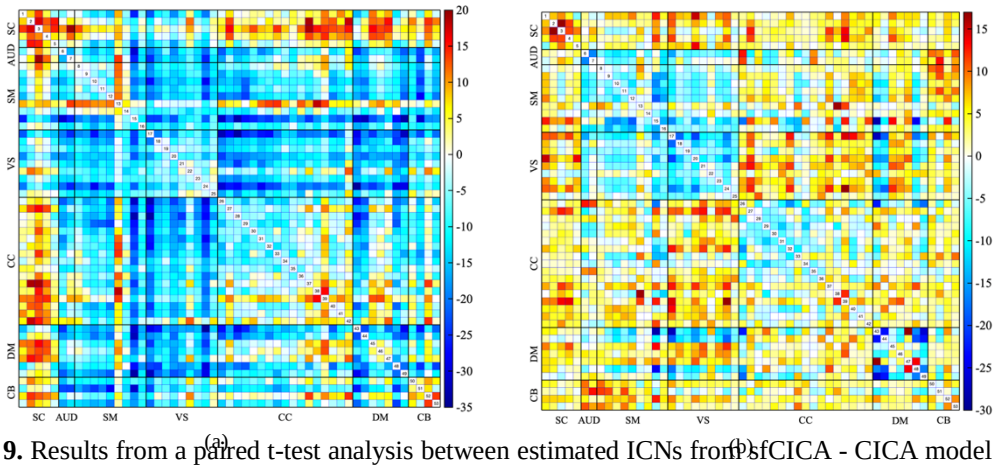


Figure 9. Results from a paired t-test analysis between estimated ICNs from sfCICA - CICA model using Pearson's correlation features. (**a**). Paired t-test values for all subjects. (**b**). Paired t-test value on the absolute values. It is observed that sfCICA is higher almost everywhere with respect to the CICA.

Moreover, paired t-test analysis was also performed regarding absolute value for the FCs and illustrated in **Figure. 8.b**. Moreover, regarding paired t-test analysis on absolute value of the FCs in **Figure. 9.b,** we observed reduced t-values for the SM, VS, CC and, DMN networks. It is replicating decreasing of the FC association between two model and can be explained based on the hypothesis that impact of the relationship between SC and FC on cognitive performance may depend on the functional domain.

5. Static FC is informed by SC weights.

Results from SC and static FC were estimated jointly by incorporating information from the SC weights on the correlation of the ICNs TCs over all time (i.e., the FC matrix). We perform two different evaluations by computing the similarity of the FC matrix to determine the degree to which the SC impacted the resulting model weights, shown in **Figure. 10**.

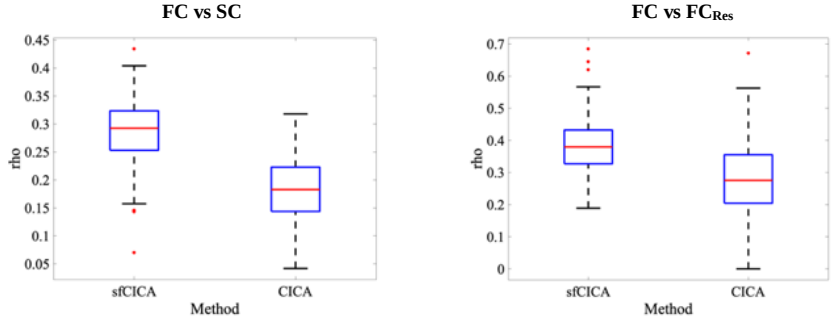


Figure 10. Similarity (Pearson correlation) between FC matrix of estimated ICNs from sfCICA model with SC and, residual (Res) FC matrix are represented in each row that is separated for FBIRN dataset and three different groups of synthetic data. Results show that similarity (correlation) with both SC and residual FC (FC_{Res}) sfCICA is higher than CICA, as expected.

For this purpose, the FC matrix of the estimated ICNs from sfCICA were computed by using correlation between paired ICNs for individuals. Then similarity of the FCs was computed (1) with SC and, (2) with residual FC, all in subject-level. Residual FC shows differences between FC of ICNs from sfCICA and CICA model. Similar procedure performed for the ICNs from CICA model and results from both sfCICA and CICA model were compared. Based on both evaluations, FC of the FBIRN dataset shown more similarity with FC with SC and, residual matrix (Res). Results indicate the proposed sfCICA model is informed from structural information.

6. Structural- functional connectivity informed ICNs reveals significant differences between controls and schizophrenia patients.

In **Figure. 11** (**a**)-(**b**), the estimated average FC across all subjects is displayed. In addition, we evaluated the group-level FC patterns across both sfCICA and CICA model using statistical paired t-test analysis. In overall, significant differences were observed between the average FCs (p-value<0.05). Significant increasing observed among some of the ICNs related to subcortical (SC), default mode network (DM) and, cerebellum (CB) networks across all subjects using sfCICA model in **Figure. 11** (**a**)-(**b**) however, sensory related networks (auditory (AUD), visual (VS) and, sensorimotor (SM)) and cognitive control (CC) network are highlighted in both.

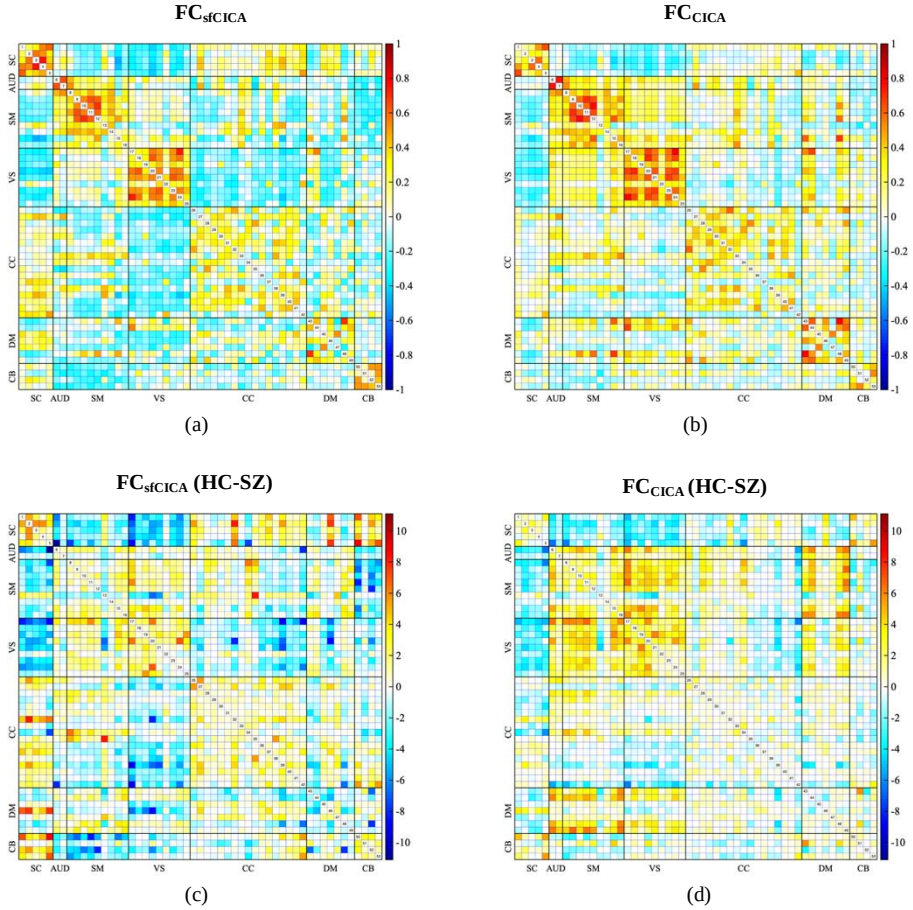


Figure 11. Group FC, estimated over all subject using estimated ICNs from sfCICA (**a**) and CICA (**b**) methods, shown significant differences (Paired t-test: p-value<0.05). In (b) and (c) results of the GLM analysis represented

significant differences between sfCICA and CICA model by regressing out diagnostic, age, imaging site and sex effect. Results were then corrected by FDR correction analysis (q<0.05) and -10*log(p-value) *sign(t-value) were represented. ICs were categorized in seven different domains including subcortical network (SC), auditory network (AUD), sensorimotor network (SM), visual network (VS), cognitive control network (CC), default mode network (DM) and cerebellum (CB). Results show that SC related networks are more significant in sfCICA model and sensory networks (AUD, VS and SM) are negatively significant.

Furthermore, to investigate the ability of the proposed method to find group differences (healthy/ patience) in FC patterns using sfCICA and CICA model, we performed generalized linear regression using multiple covariates (**Figure. 11** (**c**)-(**d**)). Results showed more significant uncorrelated (negative) correlation across all networks. The significance of the correlation in sensory networks decreased and the increment of the correlation occurred in SC related ICNs and networks, using sfCICA model. Results might replicate the power of SC which has been ignored in functional information regarding CICA model. It is suggested that SC information highlighted the role of the SC in identifying a jointly structural and functional ICA output. Values are $-10 \times \log(p_value) \times sign\,(t_value)$.

4. Discussion

In this study, joint structural-functional data-driven method learning from functional information and constrained by SC weights is proposed to characterize intrinsic functional brain patterns for healthy and schizophrenia patients.

Most previous ICA models are focused on functional information or functional information with prior spatial constraints [11, 13]. Recent work, has shifted to multimodal based models, using estimated features of different modalities and jointly analyzed them or determined the covariate changes [16]. Joint-ICA, parallel ICA, link-ICA and, cmICA models used fMRI/ dMRI/ sMRI features as connectivity, fractional anisotropy, and structural measure to determine functional networks. The proposed method is a learning procedure simultaneously using raw functional information and SC. It enables to directly constraining functional information changes during time.

Less work has considered dMRI connectivity, as performed in cmICA [16]. In cmICA [16], the authors use multimodal fMRI and dMRI by integrating connectivity features estimated from each modality to determine ICNs shared between SC and FC. Thus, there might be a mismatch between features from different modality, even in cmICA has been tried to solve it but feature contribution of each modality is asymmetric [16]. In contrast, our method is based on both direct and indirect features and the contribution of each modality is in a way to enhance each other. It means, where SC weights are strong, functional data gets more correlated and, in case of weaker connections less correlations occurred. Notably, the sfCICA model relies on two main concepts of an ICA analysis, maximizing independence and similarity, and tries to estimate intrinsic brain functional networks which are optimized for both structural and functional connectivity information. The proposed joint multimodal information integration allows us to better characterize healthy or disordered brain activity and connectivity [15, 16].

Characterization of the intrinsic brain activity known as functional network modeling has been widely used in human brain studies, particularly because of its importance in brain disorder

research [47]. One of the greatest advancements in this area uncovered by rs-fMRI images using ICA analysis [8] and continued by different modality of image.

Synthetic data analysis used to evaluate replicability and noise robustness of the proposed model by simulating three different group as represented in **Figure 3 and 4**. Each group has a different level of SNR and for each group a similar pattern, constrained to SC was observed. Our results highlighted the effect of the SC constrained on TCs as their distances were significantly decreased where there exists more SC. Additionally, by changing the functional information and keeping the SC constant, the learning procedure is dependent on both information, conferring the advantage of both modalities when estimating the ICNs.

Empirical results of our proposed multimodal model on real data showed that our model improves results in both healthy and schizophrenia groups, relative to the CICA model. Estimated components were significantly more spatially similar (**Figure. 5.a**) to the reproducible template (NeuroMark template [13]) in both healthy and schizophrenia subjects. This may enhance the replicability of the components using different datasets [48].

Moreover, the proposed sfCICA model aims to estimate a unified and integrated brain network using regional functional information instead of FC as in [33] which is the first time this has been studies in an ICA analysis. Overall, results in **Figure. 5.b** showed distance of the TCs were decreased which can be interpreted as more functionally correlated and unified ICNs. While considering each component specifically (**Figure. 6**), components related to cognitive control, default mode and cerebellum network were less constrained. In other words, weaker SC with all other ICNs, for instances, CC network represents weakest connection with CC, SM, and cerebellum network. Also, DMN represents increasing distances with its network and respectively has weakest intra network SC which is consistent with previous findings [16, 49, 50], indicating advantages of multimodal approaches using a combination of structural and functional features, which also facilitate interpretability of the results [49, 50].

On the other hand, there is a hypothesis that SC can directly impact on FC, as higher SC leads to more FC [46, 51]. Statistical analysis on FC of the estimated multi-modal (sfCICA) ICNs and single model ICNs (CICA) in **figure 9.b**, represented no association between the FCs of the two model in specific domains including SM, AUD, VS, CC and DM. Interestingly results were in accordance with previous finding [52, 53]. There is evidence that SC and FC have no coupling in SM and VS, that means either SC is not dense, or it may come from dysfunction characteristics of the SZ subjects accompanied by weaker SC in SM, anterior cingulum cortex (ACC, related to DM domain) [52]. So, our results apparently revealed it in the estimated ICNs features (FC) using sfCICA model. Regarding the VS domain, little is known about it, however it has been shown that connectivity of the visual system in SZ would be reduced [54] which is consistent with our findings. Overall, findings suggested that integrating multi-modal information to estimate ICNs could help to find more accurate intrinsic functional domains in comparison to single modality models as CICA.

On the other hand, our proposed sfCICA model enhances differences between schizophrenia patients and controls. In schizophrenia both structural and functional networks are affected [55, 56], our results showed a broad range of impacted networks. Referring to FC matrices (**Figure.**

10), results are consistent with, but extend, previous work. For example, the cognitive control network. Based on previous reports, dysfunction has been shown in the CC network, this was also observed in our statistical results using the single modality CICA model. The proposed sfCICA model showed enhanced functional correlation among the CC network with itself and other domains (e.g., visual). Overall, findings suggested the contribution of SC information could enhance our ability to identify dysfunctional connections. Also, ICNs related to memory and high order cognitive control functions related to DMN networks have been reported as affected regions in SZ patients [56-59]. Consistent with previous findings our results shows that functional correlations of DMN are reduced in positive connections while stronger in negative ones [56], representing effectiveness of our model.

This study showed that identified ICNs are restricted to prior spatial maps, and they modified spatially and temporally during optimization process. In our case, the spatial template provides a spatial and included replicable ICNs and it then adaptive to individual subjects as a functionally defined template. Our multimodal sfCICA model uncovers differences in functional activity of different regions and their interactions are affected by structural information. To include differences in spatial maps and enable the comparability of results, a multimodal (structural/functional) template is needed. On the other hand, SC was identified prior to optimization, and it remains constant over the running. Although TC and spatial maps are updated, the model can be extended to modify structural weights throughout the optimization process. Beyond this, we can also incorporate the dynamics of brain function. There is also additional functional information which can be considered in sfCICA model.

5. Conclusion

In this work we proposed a novel structural- functional connectivity information and spatial guided model, sfCICA, aiming to estimate functional networks in the human brain. The main motivation for the proposed model is that structural and functional information can jointly influence the model estimation. Results showed that integrating SC information in sfCICA enhances the sensitivity to group differences. The resulting ICNs were also more spatially focal besides more synchronization in the sfCICA model compared to CICA. Results were consistent with those founds using different synthetic images with different functional information and noise levels. In sum, our findings demonstrate the effectiveness of the proposed model in simulations and real data. The approach is applicable to the study of a wide range of areas such as brain development, aging, health, and disorder.

References

1. Genon, S., et al., *How to Characterize the Function of a Brain Region.* Trends Cogn Sci, 2018. **22**(4): p. 350-364.
2. Iraji, A., et al., *Space: A Missing Piece of the Dynamic Puzzle.* Trends Cogn Sci, 2020. **24**(2): p. 135-149.
3. Iraji, A., et al., *The spatial chronnectome reveals a dynamic interplay between functional segregation and integration.* Hum Brain Mapp, 2019. **40**(10): p. 3058-3077.
4. Iraji, A., et al., *Moving beyond the 'CAP' of the Iceberg: Intrinsic connectivity networks in fMRI are continuously engaging and overlapping.* Neuroimage, 2022. **251**.
5. Iraji, A., et al., *Multi-spatial-scale dynamic interactions between functional sources reveal sex-specific changes in schizophrenia.* Netw Neurosci, 2022. **6**(2): p. 357-381.
6. Calhoun, V.D. and T. Adalı, *Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery.* IEEE Rev Biomed Eng, 2012. **5**: p. 60-73.
7. Calhoun, V.D. and N. de Lacy, *Ten Key Observations on the Analysis of Resting-state Functional MR Imaging Data Using Independent Component Analysis.* Neuroimaging Clin N Am, 2017. **27**(4): p. 561-579.
8. Calhoun, V.D., et al., *A method for making group inferences from functional MRI data using independent component analysis.* Hum Brain Mapp, 2001. **14**(3): p. 140-51.
9. Allen, E.A., et al., *A baseline for the multivariate comparison of resting-state networks.* Front Syst Neurosci, 2011. **5**: p. 2.
10. Erhardt, E.B., et al., *Comparison of multi-subject ICA methods for analysis of fMRI data.* Hum Brain Mapp, 2011. **32**(12): p. 2075-95.
11. Du, Y. and Y. Fan, *Group information guided ICA for fMRI data analysis.* Neuroimage, 2013. **69**: p. 157-97.
12. Lin, Q.H., et al., *Semiblind spatial ICA of fMRI using spatial constraints.* Hum Brain Mapp, 2010. **31**(7): p. 1076-88.
13. Du, Y.H., et al., *NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders.* Neuroimage-Clinical, 2020. **28**.
14. Harms, M.P., et al., *Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects.* Neuroimage, 2018. **183**: p. 972-984.
15. Iraji, A., et al., *The connectivity domain: Analyzing resting state fMRI data using feature-based data-driven and model-based methods.* Neuroimage, 2016. **134**: p. 494-507.
16. Wu, L. and V. Calhoun, *Joint connectivity matrix independent component analysis: Auto-linking of structural and functional connectivities.* Hum Brain Mapp, 2023. **44**(4): p. 1533-1547.
17. Park, H.J. and K. Friston, *Structural and functional brain networks: from connections to cognition.* Science, 2013. **342**(6158): p. 1238411.
18. Batista-García-Ramó, K. and C.I. Fernández-Verdecia, *What We Know About the Brain Structure-Function Relationship.* Behav Sci (Basel), 2018. **8**(4).
19. Suárez, L.E., et al., *Linking Structure and Function in Macroscale Brain Networks.* Trends Cogn Sci, 2020. **24**(4): p. 302-315.

20. Calhoun, V.D. and J. Sui, *Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness.* Biol Psychiatry Cogn Neurosci Neuroimaging, 2016. **1**(3): p. 230-244.
21. Sui, J., et al., *Function-structure associations of the brain: evidence from multimodal connectivity and covariance studies.* Neuroimage, 2014. **102 Pt 1**: p. 11-23.
22. Calhoun, V., *Data-driven approaches for identifying links between brain structure and function in health and disease.* Dialogues in Clinical Neuroscience, 2018. **20**(2): p. 87-99.
23. Calhoun, V.D., et al., *A method for multitask fMRI data fusion applied to schizophrenia.* Human Brain Mapping, 2006. **27**(7): p. 598-610.
24. Liu, J., et al., *Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA.* Hum Brain Mapp, 2009. **30**(1): p. 241-55.
25. Chen, J., et al., *Guided exploration of genomic risk for gray matter abnormalities in schizophrenia using parallel independent component analysis with reference.* Neuroimage, 2013. **83**: p. 384-96.
26. Sui, J., et al., *Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model.* Neuroimage, 2011. **57**(3): p. 839-55.
27. Qi, S., et al., *MicroRNA132 associated multimodal neuroimaging patterns in unmedicated major depressive disorder.* Brain, 2018. **141**(3): p. 916-926.
28. Shile, Q., et al., *Supervised multimodal fusion and its application in searching joint neuromarkers of working memory deficits in schizophrenia.* Annu Int Conf IEEE Eng Med Biol Soc, 2016. **2016**: p. 4021-4026.
29. Groves, A.R., et al., *Linked independent component analysis for multimodal data fusion.* Neuroimage, 2011. **54**(3): p. 2198-2217.
30. Qi, S.L., et al., *Three-way parallel group independent component analysis: Fusion of spatial and spatiotemporal magnetic resonance imaging data.* Human Brain Mapping, 2022. **43**(4): p. 1280-1294.
31. Qi, S.L., et al., *Parallel group ICA plus ICA: Joint estimation of linked functional network variability and structural covariation with application to schizophrenia.* Human Brain Mapping, 2019. **40**(13): p. 3795-3809.
32. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis.* IEEE Trans Neural Netw, 1999. **10**(3): p. 626-34.
33. Zhu, Q., et al., *Multimodal Brain Network Jointly Construction and Fusion for Diagnosis of Epilepsy.* Front Neurosci, 2021. **15**: p. 734711.
34. Klamroth, K. and T. Jørgen, *Constrained optimization using multiple objective programming.* Journal of Global Optimization, 2007. **37**: p. 325-355.
35. Fillard, P., et al., *Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom.* Neuroimage, 2011. **56**(1): p. 220-34.
36. Poupon, C., et al., *New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models.* Magn Reson Med, 2008. **60**(6): p. 1276-83.
37. Erhardt, E.B., et al., *SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability.* Neuroimage, 2012. **59**(4): p. 4160-7.
38. Keator, D.B., et al., *The Function Biomedical Informatics Research Network Data Repository.* Neuroimage, 2016. **124**(Pt B): p. 1074-1079.
39. Iraji, A., et al., *Canonical and Replicable Multi-Scale Intrinsic Connectivity Networks in 100k+ Resting-State fMRI Datasets.* bioRxiv, 2022: p. 2022.09.03.506487.

40. Jenkinson, M., et al., *FSL.* Neuroimage, 2012. **62**(2): p. 782-90.
41. Andersson, J.L.R. and S.N. Sotiropoulos, *An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging.* Neuroimage, 2016. **125**: p. 1063-1078.
42. Caprihan, A., et al., *Source-based morphometry analysis of group differences in fractional anisotropy in schizophrenia.* Brain Connect, 2011. **1**(2): p. 133-45.
43. Wu, L., et al., *Connectivity-based whole brain dual parcellation by group ICA reveals tract structures and decreased connectivity in schizophrenia.* Hum Brain Mapp, 2015. **36**(11): p. 4681-701.
44. Cook, P.A., et al. *Camino: Open-Source Diffusion-MRI Reconstruction and Processing.* 2006.
45. Avants, B.B., N. Tustison, and G. Song, *Advanced normalization tools (ANTS).* Insight j, 2009. **2**(365): p. 1-35.
46. Litwińczuk, M.C., et al., *Combination of structural and functional connectivity explains unique variation in specific domains of cognitive function.* Neuroimage, 2022. **262**: p. 119531.
47. Wang, Z., et al., *Brain functional network modeling and analysis based on fMRI: a systematic review.* Cogn Neurodyn, 2021. **15**(3): p. 389-403.
48. Duda, M., et al., *Reliability and clinical utility of spatially constrained estimates of intrinsic functional networks from very short fMRI scans.* Human Brain Mapping, 2023. **44**(6): p. 2620-2635.
49. Glasser, M.F., et al., *A multi-modal parcellation of human cerebral cortex.* Nature, 2016. **536**(7615): p. 171-178.
50. Van Essen, D.C., et al., *Cerebral cortical folding, parcellation, and connectivity in humans, nonhuman primates, and mice.* Proceedings of the National Academy of Sciences, 2019. **116**(52): p. 26173-26180.
51. Zhao, J., et al., *Structure-function coupling in white matter uncovers the abnormal brain connectivity in Schizophrenia.* Translational Psychiatry, 2023. **13**(1): p. 214.
52. Li, P., et al., *Association between functional and structural connectivity of the corticostriatal network in people with schizophrenia and unaffected first-degree relatives.* J Psychiatry Neurosci, 2020. **45**(6): p. 395-405.
53. Tsai, Y.H., et al., *Modular organization of brain resting state networks in patients with classical trigeminal neuralgia.* Neuroimage Clin, 2019. **24**: p. 102027.
54. Reavis, E.A., et al., *Structural and Functional Connectivity of Visual Cortex in Schizophrenia and Bipolar Disorder: A Graph-Theoretic Analysis.* Schizophrenia Bulletin Open, 2020. **1**(1).
55. Li, P., et al., *Altered brain network connectivity as a potential endophenotype of schizophrenia.* Scientific reports, 2017. **7**(1): p. 5483.
56. Qi, S., et al., *Parallel group ICA+ ICA: Joint estimation of linked functional network variability and structural covariation with application to schizophrenia.* Human brain mapping, 2019. **40**(13): p. 3795-3809.
57. Aleman, A., et al., *Memory impairment in schizophrenia: a meta-analysis.* American Journal of Psychiatry, 1999. **156**(9): p. 1358-1366.
58. Minzenberg, M.J., et al., *Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia.* Archives of general psychiatry, 2009. **66**(8): p. 811-822.

59. Potkin, S., et al., *Working memory and DLPFC inefficiency in schizophrenia: the FBIRN study.* Schizophrenia bulletin, 2009. **35**(1): p. 19-31.