



Group Similarity Constraint Functional Brain Network Estimation for Mild Cognitive Impairment Classification

Xin Gao^{1†}, Xiaowen Xu^{2,3†}, Xuyun Hua^{4,5}, Peijun Wang^{2,3}, Weikai Li^{6,1*} and Rui Li^{7*}

¹ Shanghai Universal Medical Imaging Diagnostic Center, Shanghai, China, ² Tongji University School of Medicine, Tongji University, Shanghai, China, ³ Department of Medical Imaging, Tongji Hospital, Shanghai, China, ⁴ Yueyang Hospital of Integrated Chinese and Western Medicine, Shanghai, China, ⁵ Shanghai University of Traditional Chinese Medicine, Shanghai, China, ⁶ College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, ⁷ School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an, China

OPEN ACCESS

Edited by:

Bertrand Thirion,
Institut National de Recherche en
Informatique et en Automatique
(INRIA), France

Reviewed by:

Xiaoyun Liang,
Australian Catholic University,
Australia
Meiling Li,
Athinoula A. Martinos Center
for Biomedical Imaging and Harvard
Medical School, United States

*Correspondence:

Weikai Li
leeweikai@outlook.com
Rui Li
liruizixing@163.com

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 12 November 2019

Accepted: 14 February 2020

Published: 10 March 2020

Citation:

Gao X, Xu X, Hua X, Wang P, Li W
and Li R (2020) Group Similarity
Constraint Functional Brain Network
Estimation for Mild Cognitive
Impairment Classification.
Front. Neurosci. 14:165.
doi: 10.3389/fnins.2020.00165

Functional brain network (FBN) provides an effective biomarker for understanding brain activation patterns and a diagnostic criterion for neurodegenerative diseases detections. Unfortunately, it remains challenges to estimate the biologically meaningful or discriminative FBNs accurately, because of the poor quality of functional magnetic resonance imaging data or our limited understanding of human brain. In this study, a novel FBN estimation model based on group similarity prior was proposed. In particular, we extended the FBN estimation model to tensor form and incorporated the tensor trace-norm regularizer to formulate the group similarity constraint. To verify the proposed method, we conducted experiments on identifying mild cognitive impairments (MCIs) from normal controls (NCs) based on the estimated FBNs. Experimental results illustrated that our method is effective in modeling FBNs. Consequently, we achieved 91.97% classification accuracy, outperforming the state-of-the-art methods. The *post hoc* analysis further demonstrated that more biologically meaningful functional brain connections were obtained using our proposed method.

Keywords: functional brain network, functional magnetic resonance imaging, group similarity constraint, mild cognitive impairment, Pearson's correlation, partial correlation

INTRODUCTION

As a neurodegenerative disorder, Alzheimer's disease (AD) is one of the most common causes of dementia (Wee et al., 2012). According to a recent report (Bain et al., 2008), the incidence of AD doubles every 5 years after age 60. AD seriously interferes with patients' daily life, affects their memory and ability to communicate, and eventually causes their deaths. Unfortunately, there is no effective treatment for AD thus far. Hence, it is quite important to delay the onset and progression of AD during its early stages via pharmacological and behavioral interventions.

Mild cognitive impairment (MCI) is often considered as a critical time window and treatment period for the prediction or delaying the conversion in AD (Wee et al., 2012). In some recent statistical studies, nearly 10–15% patients with MCI develop probable AD each year

(Grundman et al., 2004; Misra et al., 2009). The early detection and accurate diagnosis of MCI is considered a significant means of slowing AD progression (Alzheimer's Association, 2017).

As a successful non-invasive technique, functional magnetic resonance imaging (fMRI) provides an effective method of measurement for revealing brain activities and patterns (Brunetti et al., 2006; Kevin et al., 2008; Jin et al., 2010). However, because spontaneous brain activity is random and asynchronous across subjects and scanners, it remains a challenge to identify MCI patients from normal controls (NC) by directly using the imaging information. Furthermore, high-order FBN-based statistical information provides new perspectives for discovering brain activity and connection patterns, thus improving our ability to understand brain information (Smith et al., 2011; Sporns, 2011; Wee et al., 2012; Stam, 2014; Rosa et al., 2015). In addition, various research has shown that the changing of functional brain networks are closely related to various neurological and psychological diseases such as AD (Supekar et al., 2008; Huang et al., 2009; Liu F. et al., 2012), MCI (Fan and Browndyke, 2010; Wee et al., 2012, 2014; Yu et al., 2016), autism spectrum disorder (ASD) (Theije et al., 2011; Gotts et al., 2012), Parkinson's disease (PD) (Baggio et al., 2014), etc. All of these depend heavily on the quality of the final estimated FBN. Hence, improved FBN reliability is crucial to such estimates (Li et al., 2019a).

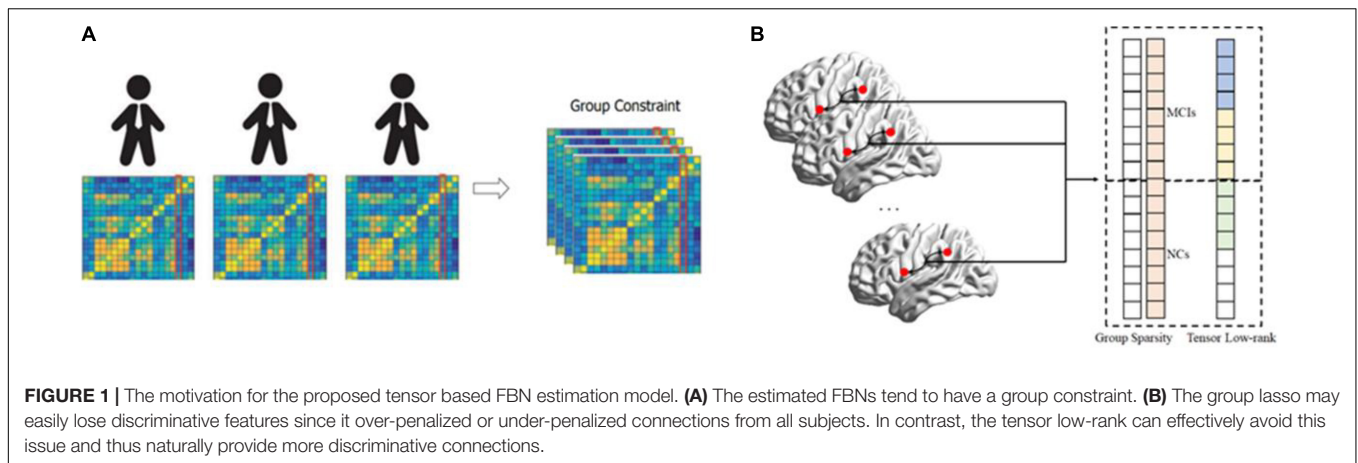
According to a FBN research review (Smith et al., 2011), correlation-based methods such as Pearson's correlation (PC) (Li et al., 2017) and sparse representation (SR) (Lee et al., 2011; Zhou et al., 2014), are generally more sensitive than complex, high-order methods. However, due to the influence of noise in the observed data, correlation-based brain networks inevitably exhibit dense connections and thus contain substantial noise or false connections. One solution is to introduce sparse priors, as is done in the thresholding and SR (LASSO) methods. Actually, the topological structure of an FBN involves more than just sparsity (Sporns, 2011). Several studies (Lee et al., 2011; Qiao et al., 2016; Wee et al., 2016; Yu et al., 2016; Li et al., 2017, 2019b) have been focused on incorporating additional biological priors into FBNs to make them more discriminative. In practice, sparsity, modularity, group-sparsity, low-rank, and scale-free priors are commonly used (Lee et al., 2011; Qiao et al., 2016; Wee et al., 2016; Yu et al., 2016; Li et al., 2017). Moreover, priors can also be obtained from data quality (Li et al., 2019a) and other high-quality data (Li et al., 2019b). Note that most of the biological/data priors can be formulated into a regularized framework. This illustrates that a reliable FBN estimation model should both fit the data well and effectively encode brain organization priors (Qiao et al., 2016).

Despite the advantages of existing FBN estimation methods, it is currently still an open field to estimate FBNs due to the complex of human brains and the poor quality of the observed data. In this paper, we focus on the group similarity prior of FBN (Wee et al., 2014), as shown in **Figure 1A**. In contrast, most current FBN estimation methods focus primarily on a single participant and rarely consider inter-group information from cross-participants, which result in different network topological structures across subjects. This

performance inevitably makes comparisons between subjects difficult and thus can degrade the generalization performance of trained classifiers. Besides, the existing group constraint methods are mainly based primarily on the group sparsity penalty (i.e., $l_{2,1}$ -norm) to mitigate inter-subject variability (Wee et al., 2014; Yu et al., 2016). However, the specific information from individuals can be ignored, due to the additional $l_{2,1}$ -norm can often over-penalized or under-penalized connections of estimated FBN as shown in **Figure 1B**. In addition, some researchers have focused on group-fused multiple graphical-lasso schemes (Liang et al., 2016, 2018), which alleviating the issue of group sparsity constraints in some extent. As mentioned above, existing group-based FBN estimation approaches still have great potential.

In this paper, we use the regularization framework to incorporate the group similarity constraint into the FBN estimation model. In detail, we formulate the group similarity prior as a tensor low-rank (TLR) regularizer and incorporate it into the FBN estimation model. In addition, we further incorporate the sparse and tensor low-rank (STLR) for better FBN estimation. Since the low-rank is NP-hard, we optimize its upper limit (i.e., the trace norm penalization) for better calculation efficiency. In particular, we adopt Parallel factor analysis (PARAFAC) to calculate its eigenvalues (Liu X. et al., 2012) and design a proximal operator to estimate the FBN with the group similarity constraint. In the end, we incorporate the trace norm regularizer into the SR and PC models to create a simple test platform. To verify the proposed methods, we adopt an estimated FBN for MCI identification. In fact, the proposed method uses the group similarity constraints to shrink the FBN solution space, and thus can estimate more discriminative FBNs effectively. The highlights of this paper include:

1. We incorporate the group similarity constraint into the FBN estimation model using a low-rank regularizer. In addition, we further relax it into a trace norm regularizer and design an optimization algorithm to estimate FBNs with group similarity.
2. We use group similarity-based FBNs to separate MCIs from NCs. The experimental results show that the proposed scheme outperforms the baseline methods. Moreover, the proposed methods can provide more biologically meaningful connections than existing methods.
3. We provide an effective FBN estimation module useful in modeling the group similarity prior. The module is flexible enough to incorporate into other FBN estimation models. The experimental results show that the proposed module can effectively improve the MCI classification accuracies of the estimated FBNs.
4. We identified the most significant functional connections and the most discriminative brain regions using the proposed FBN estimation model. This analysis of functional connectivity and graph theory attributes can be used to discover biologically meaningful biomarkers and further elucidate the topological properties of a brain network that is experiencing MCI.



The remainder of this paper is organized as follows. In Section 2, we introduce our data preparation methodology. The proposed methods, i.e., the group similarity-based FBN estimation schemes, including the motivations, models, and algorithms are introduced. In Section 3, we evaluate the proposed methods using MCI identification experiments. A discussion and conclusions are presented in Sections 4 and 5, respectively.

MATERIALS AND METHODS

Data Acquisition

For this study, we adopted publicly available neuroimaging data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (Jack et al., 2010)¹. ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and non-profit organizations. The initial goal of ADNI was to define biomarkers for use in clinical trials and to determine the best way to measure the treatment effects of AD therapeutics.

In particular, 137 participants including 68 MCIs and 69 NCs were adopted in this experiment, which was similar to a previous study (Zhou et al., 2018). The scanning parameters included: TR/TE = 3000/30 mm, flip angle = 80, imaging matrix = 64 × 64, 48 slices, 140 volumes, and voxel thickness = 3.3 mm. The SPM8 toolbox² and DPARSFA (version 2.2) (Chao-Gan and Yu-Feng, 2010) were used to preprocess the fMRI data using the well-accepted pipelines. The preprocessing pipeline included removing the first 10 volumes, slice timing, realigning, normalizing, spatially smoothing, temporally detrending, regressing out covariates (ventricle and WM signals, as well as six head-motion parameters), and temporally filtering. We followed previous work (Chen et al., 2016, 2017) to alleviate head motion effects and artifacts and excluded subjects with more than 2.5 min (50 frames) of data with FD > 0.5 from further analysis (Power et al., 2012). Finally,

we used the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) to partition pre-processed blood-oxygen level dependent (BOLD) signals into 116 ROIs with 137 volumes. Finally, for k th participants, we put these volumes into a data matrix $X^{(k)} \in R^{N \times T}$ (X for short). For more details, please refer to Zhou et al. (2018).

Functional Brain Network Estimation

After obtaining fMRI data matrix X from the R-fMRI data, we performed FBN estimation. As mentioned above, correlation-based FBN estimation methods have been demonstrated to be more sensitive than some complex higher-order methods (Smith et al., 2011). Therefore, this paper focuses on correlation-based methods and adopts them as a baseline. In particular, we first defined the data matrix (i.e., the BOLD signal matrix) $X \in R^{N \times T}$, where T is the number of volumes and N is the number of ROIs. The fMRI time series associated with the i th ROI is represented by $x_i \in R^T$, $i = 1, \dots, N$.

Correlation-Based Methods

As the simplest and most widely used FBN estimation schemes, PC-based FBN estimation methods account for a large proportion of FBN studies (Smith et al., 2013). The FBN edge weights $W = (W_{ij}) \in R^{N \times N}$ can be calculated via PC as follows:

$$W_{ij} \propto \frac{(x_i - \bar{x}_i)^T (x_j - \bar{x}_j)}{\sqrt{(x_i - \bar{x}_i)^T (x_i - \bar{x}_i)} \sqrt{(x_j - \bar{x}_j)^T (x_j - \bar{x}_j)}}, \quad (1)$$

In Eq. (1), $x_i - \bar{x}_i$ is a centralized counterpart of x_i . Due to the effects of the noise in the fMRI data, PC always generates dense FBNs. Thus, a thresholding scheme is often used to make the PC-based FBNs sparse by filtering out the noisy or weak connections. The PC based FBN can be expressed as follows:

$$W_{ij}^{(new)} \propto \begin{cases} W_{ij}, & W_{ij} > \text{threshold} \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where $W_{ij}^{(new)}$ denotes the connection value between nodes i and j after thresholding.

¹<http://adni.loni.ucla.edu>

²<http://www.fil.ion.ucl.ac.uk/spm/>

When one compares PC measures to full-correlation cross ROIs, one notes that the interaction among multiple ROIs is neglected due to the confounding effect. In contrast, a partial correlation is proposed by regressing out the confounding effects from other ROIs. However, the partial correlation-based methods can easily be ill-posed due to the singularity of the covariance matrix $\Sigma \times \mathbf{X}^T \mathbf{X}$. One simple solution is to incorporate an l_1 -norm regularizer into the partial correlation model (Lee et al., 2011), thus naturally incorporating the FBN sparsity prior (SR). The SR model is as follows:

$$\min_{W_{ij}} \sum_{i=1}^n \left\| x_i - \sum_{j \neq i} W_{ij} x_j \right\|^2 + \lambda \sum_{j \neq i} |W_{ij}|, \quad (3)$$

The matrix form is proposed as follows:

$$\min_{\mathbf{W}} \|\mathbf{X} - \mathbf{X}\mathbf{W}\|_F^2 + \lambda \|\mathbf{W}\|_1$$

s.t. $W_{ii} = 0, \forall i = 1, \dots, n,$ (4)

Note that the l_1 -norm regularizer in Eq. (4) plays a key role in achieving a sparse, stable solution (Lee et al., 2011).

Regularization Framework for FBN Estimation

Based on the above description, both PC- and SR-based FBN estimation models can be summarized into the regularized FBN learning framework. We can naturally incorporate a regularized term and statistical information into the objective function in order to construct a new FBN estimation platform. Specifically, the platform can be formulated using a matrix-regularized learning framework as follows:

$$\min_{\mathbf{W}} f(\mathbf{X}, \mathbf{W}) + \lambda R(\mathbf{W}), \text{ s.t. } \mathbf{W} \in \Delta, \quad (5)$$

where $f(\mathbf{X}, \mathbf{W})$ models the FBN statistical information and $R(\mathbf{W})$ is the regularization term used to incorporate FBN biological priors and stabilize solutions. In addition, some specific constraints such as symmetry or positive semi-definiteness may be included in Δ to shrink the \mathbf{W} search space. This provides an effective way of obtaining a better FBN. The λ is a hyper-parameter that controls the balance between the first (data-fitting) and second (regularization) terms.

Most of the recently proposed FBN estimation models (Higgins et al., 2018; Li et al., 2018; Wang et al., 2018; Zhou et al., 2018) can be unified under this regularized framework by re-designing the two terms in Eq. (5). Popular data-fitting terms include $\|\mathbf{W} - \mathbf{X}^T \mathbf{X}\|_F^2$ used in Eq. (2) and $\|\mathbf{X} - \mathbf{X}\mathbf{W}\|_F^2$ used in Eq. (4), while popular regularization terms include l_1 -norm (Huang et al., 2010), trace norm, their combination (Qiao et al., 2016), etc.

Sparse and Low-Rank-Based FBN Estimation

Before we introduce the proposed method, we would like to review the sparse and low-rank-based (SLR) FBN estimation model briefly (Qiao et al., 2016). The sparsity and low-rank regularizers, i.e., the l_1 -norm and trace norm) causes sparse and similar connections across each brain region, naturally

incorporating estimated FBN modularity priors. The SLR FBN estimation model is given as follows:

$$\min_{\mathbf{W}} \|\mathbf{X} - \mathbf{X}\mathbf{W}\|_F^2 + \lambda \|\mathbf{W}\|_1 + \gamma \|\mathbf{W}\|_*, \quad (6)$$

where \mathbf{X} is the BOLD signal data, \mathbf{W} represents the estimated FBNs, $\lambda \|\mathbf{W}\|_1$ is the sparsity regularizer and $\gamma \|\mathbf{W}\|_*$ is the low-rank regularizer.

Group Sparsity-Based FBN Estimation

However, the abovementioned FBN estimation models are unable to deal with inter-subject variability problems because the FBN is estimated at an individual level, which easily causes different network topological structures across subjects. To mitigate the effects of inter-subject variability, Wee et al. proposed a group-constrained sparse linear regression model (Wee et al., 2014) that applied the idea of joint feature selection in group-lassos to regression problems (Yuan and Lin, 2006). In particular, a group sparsity regularizer (GSR, i.e., $l_{2,1}$ -norm) was incorporated into the FBN estimation model. The GSR FBN estimation model is given as follows:

$$\min_{\mathbf{W}_j} \sum_{k=1}^N \|\mathbf{X}_j^k - \mathbf{X}^k \mathbf{W}_j^k\|_F^2 + \lambda \|\mathbf{W}\|_{2,1}, \quad (7)$$

where \mathbf{X}_j^k is the BOLD signal of the j th ROI and k th participant, \mathbf{X}^k is the data matrix of k participants, \mathbf{W}_j^k represents the functional connections of the j th ROI and k th participant, and $\lambda \|\mathbf{W}\|_{2,1}$ is the group sparsity regularizer. Relative to the SR method, this minimizes inter-subject variability via an additional l_2 -norm regularizer across all subjects. However, these methods may penalize too much for estimated FBNs. For example, if a functional connection is removed from the MCIs but exists in NCs and the weight of this connection in the NCs is slightly larger than in MCIs, the GSR method tends to force removal of this connection from the NCs. In addition, if the number of NCs or the weight of this connection in the NCs is substantially larger than in the MCIs, the GSR method tends to force this connection to exist in the MCIs. Thus, a GSR can lose discriminative information from estimated FBNs, as shown in **Figure 1B**.

Methods

To incorporate group constraints easily and directly, we first extended the existing matrix regularization framework to tensor form as follows:

$$\min_{\mathbf{W}} \sum_{k=1}^n f(\mathbf{X}^{(k)}, \mathbf{W}^{(k)}) - \lambda R(\mathbf{W}) \quad (8)$$

where $\mathbf{X}^{(k)} \in \mathbb{R}^{N \times T}$ represents the input data of the k th participant, ROI is the number of predefined ROIs, T is the duration of data observation, and n is the number of participants. $\mathbf{W} \in \mathbb{R}^{N \times N \times K}$ represents the estimated FBNs and $\mathbf{W}^{(k)} \in \mathbb{R}^{N \times N}$ represents the corresponding FBN of the k th participant. K is the number of participants. Obvious, in Eq. (8), \mathbf{W} is 3-dimensional tensors. As with the matrix regularization framework, in Eq. (8), $\sum_{k=1}^n f(\mathbf{X}^{(k)}, \mathbf{W}^{(k)})$ is

the data-fitting term and $R(W)$ is the regularization term in tensor format.

As shown in **Figure 1**, the abovementioned $l_{2,1}$ -norm penalty excessively punishes the estimated FBNs, which leads to interference across various groups in the data. To alleviate this issue, this paper uses the tensor regularization framework to relax the $l_{2,1}$ -norm penalty and naturally introduce the tensor low-rank (TLR) regularizer to formulate the group similarity prior. The proposed tensor low-rank-based FBN estimate can be formulated as follows:

$$\min_W \sum_{k=1}^K f(\mathbf{X}^{(k)}, \mathbf{W}^{(k)}) + \lambda R(\mathbf{W})_{\text{lowrank}} \quad (9)$$

For the regularized terms in Eq. (9), $R(\mathbf{W})_{\text{lowrank}}$ indicates the rank of tensor \mathbf{W} , which can be represented by number of non-zero elements in the eigenvalue of \mathbf{W} . Unfortunately, the low-rank regularizer is non-convex with respect to W and is NP-hard to solve. Thus, we relax it to trace-norm $\|\mathbf{W}\|_*$ and obtain the following optimization model:

$$\min_W \sum_{k=1}^K f(\mathbf{X}^{(k)}, \mathbf{W}^{(k)}) + \lambda \|\mathbf{W}\|_* \quad (10)$$

Here, we aim to capture the partial correlation of the observed fMRI data due to its empirical effectiveness. In particular, we adopted SR as a testing platform since the PC method suffers from confounding effects. In particular, we used $\|\mathbf{X}_j^k - \mathbf{X}^k \mathbf{W}_j^i\|_F^2$ as the data-fitting term to formulate the inverse covariance structure (i.e., partial correlation) in the data, and added a l_1 -norm penalty to encode the sparse priors, resulting in the following Sparse and Tensor Low-Rank (STLR) optimization model.

$$\min_W \sum_{k=1}^K \|\mathbf{X}^k - \mathbf{X}^k \mathbf{W}^k\|_F^2 + \lambda \|\mathbf{W}\|_* + \gamma \|\mathbf{W}\|_1 \quad (11)$$

where λ and γ are hyper-parameters used to control the balance between the three terms in the objective function. It should also be noted that the data fitting term can be designed as $\|\mathbf{W}^k - \mathbf{X}^{kT} \mathbf{X}^k\|_F^2$ to capture full correlation statistics. In addition, when $\gamma = 0$, the proposed method reduces to the network learning model based on the traditional sparse regression FBN estimation method given in Eq. (4). As we can see in Eq. (11), when $\lambda = 0$, Eq. (8) will be reduced to TLR.

Algorithm

Because the l_1 -norm and trace penalties exist, the proposed scheme is convex but non-differentiable. This leads to a non-trivial problem. Fortunately, several approaches have been proposed for dealing with such issues (Donoho and Elad, 2003; Meinshausen and Bühlmann, 2006; Tomioka and Sugiyama, 2009). In this paper, we use the proximal method (Combettes and Pesquet, 2011) to solve the proposed optimal FBN estimation model because of its simplicity and efficiency. The details are given as follows:

First, we address the STLR or TLR data-fitting term (i.e., $\sum_{k=1}^n \|\mathbf{X}^k - \mathbf{X}^k \mathbf{W}^i\|_F^2$), whose gradient with respect to W^k is $\nabla_{\mathbf{W}^k} f(\mathbf{X}^k, \mathbf{W}^k) = \mathbf{X}^{kT} \mathbf{X}^k \mathbf{W}^k - \mathbf{X}^{kT} \mathbf{X}^k$. For each iteration, we

first update the W according to the gradient descent criterion:

$$\mathbf{W}_t^k = \mathbf{W}_{t-1}^k - a_t \nabla f(\mathbf{X}^k, \mathbf{W}_{t-1}^k), \quad (12)$$

where a_t denotes the gradient descent step size. The initial value of the step size a_t was set to 0.001 and subsequently adaptively updated based on the line search scheme proposed by Nemirovski (Nesterov, 1983) using the SLEP toolbox³.

Second, we address the regularization term $\|\mathbf{W}\|_1$. According to the definition of a proximal operator (Combettes and Pesquet, 2011), the proximal operator of $\|\mathbf{W}\|_1$ is equivalent to the following soft thresholding operation on W ,

$$\text{prox}_{\lambda \|\cdot\|_1}(\mathbf{W}) = [\text{sign}(W_{ij}) \times \max(W_{ij} - \lambda, 0)]_{n \times n}, \quad (13)$$

Similarly, the proximal operator $\lambda \|\mathbf{W}\|_*$ corresponds to a shrinkage operation on the singular values of W as follows.

$$\text{prox}_{\lambda \|\cdot\|_*}(\mathbf{W}) = \sum_{r=1}^R \max(\lambda_r - \lambda, 0) a_{ir}^{(1)} a_{jr}^{(2)} a_{kr}^{(3)} \quad (14)$$

Here, a_{ir}, a_{jr}, a_{kr} is a vector in unit norm space and λ_r is the corresponding eigenvalue based on the parallel factor analysis (PARAFAC) (Liu X. et al., 2012). Then, the final algorithm is given in **Table 1**:

TABLE 1 | The Algorithm for Estimating the FBN based on STLR/TLR.

Input: \mathbf{X}, λ
Output: \mathbf{W}
Initialize \mathbf{W}
while <i>not converged</i>
$\mathbf{W}_{k1} = \text{gradient}(\mathbf{W});$
$\mathbf{W}_{k1} = \text{prox}_{\lambda \ \cdot\ _*}(\mathbf{W}_k);$
$\mathbf{W}_{k1} = \text{prox}_{\lambda \ \cdot\ _1}(\mathbf{W}_k);$
End

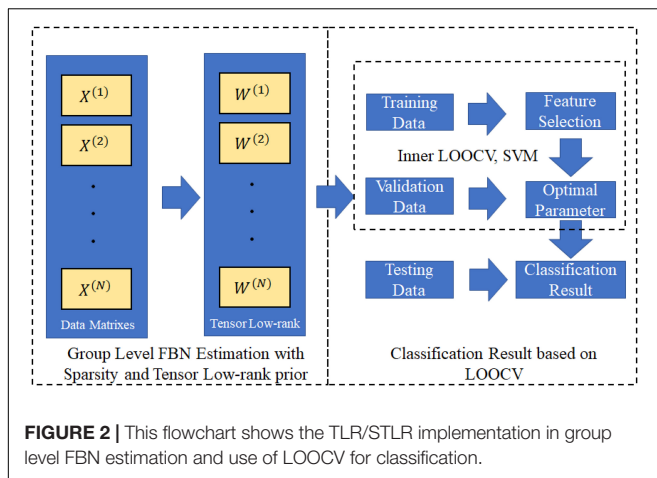
EXPERIMENT

Experimental Setting

After obtaining the FBNs of all subjects, the main task remaining was to use the aforementioned FBNs to train a classifier that could separate ASDs from NCs. Since the FBN matrix was symmetric, we used only its upper triangular elements as classification input features. Even so, the feature dimensions remained too high to train a classifier with good generalization due to the limited training sample availability in this study. Therefore, we performed feature filtering before classification training. Specifically, the classification pipeline included the following two main steps. A flow chart is given in **Figure 2**. First, we estimated FBNs for each individual using PC⁴, SR,

³<http://yelabs.net/software/SLEP/>

⁴In order to improve PC flexibility and conduct fair comparisons, we introduced a hard-thresholding PC parameter by reducing a portion of weak connections.



SLR, GSR, TLR and STLR. The estimated FBNs are shown in **Figure 3**. After we obtained the estimated FBNs, we sought to determine how to use these connections to separate MCIs from NCs. It should be noted that both the feature selection and classifier design have large influences on the final accuracy (Wee et al., 2014). Because of this and because our focus was FBN estimation, we adopted the simplest feature selection method (t -test with p -value < 0.01) and used the most popular linear SVM classifier with default parameter $C = 1$ (Chang and Lin, 2007).

Due to the small sample size, we used the leave one out cross validation (LOOCV) strategy to verify method performance. In this strategy, only one subject was left out of testing while the others were used to train the models and determine optimal parameters. To choose optimal parameters, an inner LOO cross-validation was further conducted on the training data using a grid-search strategy. In the outer loop, we chosen the training and testing dataset to re-select feature and re-train the model by the selected parameters. More specifically, for the regularized parameters λ and γ , the candidate value range was $[2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$. For the hard threshold of $PC_{\text{threshold}}$, we used 20 sparsity levels with a range of [5%, 10%, ..., 95%, 100%]. For example, 90% means that 10% of the weak edges were filtered out of the FBN. In the outer loop, we used a training and testing dataset to re-select features and re-train the model based on the selected parameters.

Network Visualization

For visual comparison of the FBNs constructed using the PC, SR, SLR, GSR, TLR and STLR methods, we first show the FBN adjacency matrices. W is shown as constructed via various methods in **Figure 3**. It can be observed from **Figure 3** that the PC-based FBNs (i.e., **Figure 3A**) are quite different from the SR-based FBNs (i.e., **Figures 3B–F**) since they use a different data-fitting term [i.e., the first term in Eq. (5)]. Moreover, the topology of the FBN as estimated via SLR is similar to those produced by STLR and TLR because (1) both methods employ the same data-fitting term and (2) the low-rank and sparse regularity behind

SLR (i.e., the trace norm in the matrix scheme) are based on the STLR result (i.e., the trace norm in the tensor scheme).

MCI Identification

The MCI versus NC classification results from the ADNI dataset are given in **Table 2**. The proposed STLR method achieved the best accuracy in this experiment. In addition, the SLR and GSR results are also provided in **Figure 4** and **Table 2**.

A set of quantitative measurements, including accuracy, sensitivity and specificity, were used to evaluate the classification performances of six different methods (PC, SR, SLR, GSR, TLR and STLR). The mathematical definitions of these three measures follow:

$$\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TruePositive} + \text{FalsePositive} + \text{TrueNegative} + \text{FalseNegative}}, \quad (15)$$

$$\text{Sensitivity} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}, \quad (16)$$

$$\text{Specificity} = \frac{\text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive}}, \quad (17)$$

Here, *TruePositive* is the number of positive subjects that are correctly classified in the ASD identification task. Similarly, *TrueNegative*, *FalsePositive*, and *FalseNegative* are the quantities of their respective, corresponding subjects. In addition, the ROC of these methods is provided in **Figure 4**.

Sensitivity to Network Model Parameters

Regardless of the FBN estimation method used, the classification accuracy is sensitive to various parameters (e.g., regularized SLR, GSR, TLR, and STLR parameters). Therefore, in our above classification experiments, we determine classification results with various parameters based on LOOCV. In **Figure 5**, we show the classification accuracies that correspond to various STLR parametric combinations. **Figure 5** shows that we achieve the best accuracy (92.70%) with $\lambda = 2^{-3}$ (for sparsity) and $\gamma = 2^4$ (for tensor low-rank).

Most Discriminant Brain Regions and Consensus Connections

As the selected connections in each inner loop might be different. We recorded all selected features during the training

TABLE 2 | Classification performance of various FBN estimation methods on the ADNI dataset.

Method	Accuracy	Sensitivity	Specificity
PC	67.15	72.06	62.32
SR	78.10	79.41	76.81
SLR	80.29	80.88	79.71
GSR	83.21	88.24	78.26
TLR	85.40	86.96	83.82
STLR	91.97	92.65	91.30

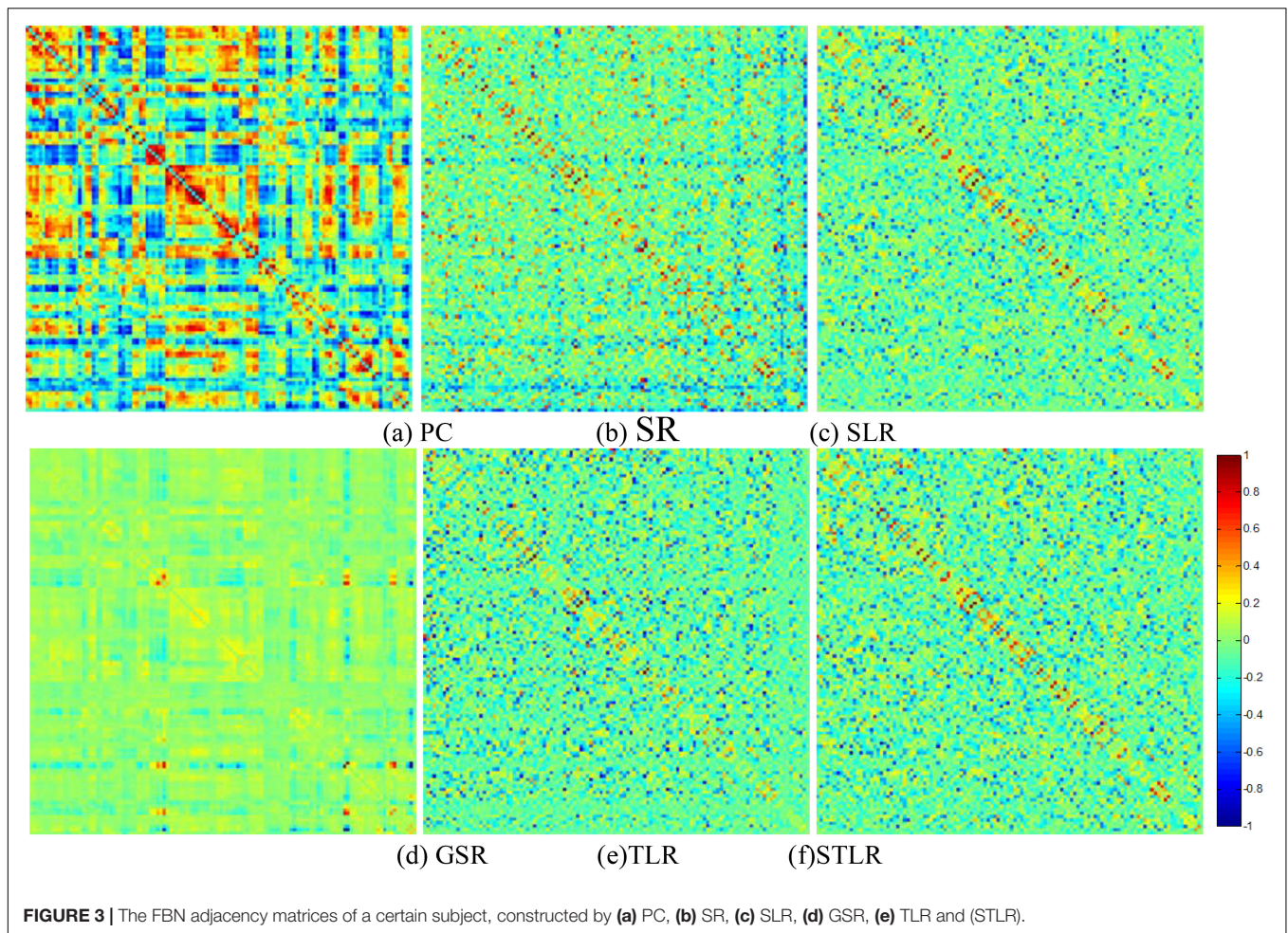


FIGURE 3 | The FBN adjacency matrices of a certain subject, constructed by (a) PC, (b) SR, (c) SLR, (d) GSR, (e) TLR and (f) STLR.

process. The statistics of the selected connections include a mean of 131.08 and variance 4.15. In addition, we further record the consensus connections for the classification model in each inner LOOCV loops. As mentioned above, we select the consensus connections with p -value < 0.01 in each loops, resulting in the 82 consensus connections shown in **Figure 6**. The thickness of an arc indicates the discriminative power of an edge, and is inversely proportional to the estimated p -values. The arc colors in **Figure 6** are randomly generated to differentiate ROIs. In particular, there are 19 functional network connections that show decreased functional connectivity in patients with MCI.

In addition, we provide the most significant MCI and NC brain regions. In particular, the top 20 brain regions (without the cerebellum) with the largest number of discriminative connections ($p < 0.01$) are given in **Table 3**. The experimental results demonstrate that the brain regions that most discriminate between MCIs and NCs are distributed primarily in the thalamus, middle temporal gyrus, hippocampus, parahippocampal gyrus, inferior parietal (which corresponds to the subcortical network), Default mode network (DMN), dorsal attention network, and fronto-parietal task control network.

STLR on an Independent Dataset

To evaluate the purposed scheme further, we re-selected 50 independent participants (including 27 MCIs and 23 NCs) from the ADNI dataset to create an independent test dataset. Following the same preprocessing pipelines as mentioned above, the model was pre-trained on the aforementioned dataset with $\lambda = 2^{-3}$ (for sparsity) and $\gamma = 2^4$ (for tensor low-rank). Consequently, it achieved 86.00% accuracy, 91.30% sensitivity, and 85.19% specificity, which further demonstrated the effectiveness of the proposed method.

Altered Functional Network Topological Properties in MCI Patients

Based on the FBN estimated via STLR with $\lambda = 2^{-3}$ (for sparsity) and $\gamma = 2^4$ (for tensor low-rank), several global graph theory metrics, including clustering coefficients (C_p), the shortest path length (L_p), the normalized clustering coefficient (γ), the normalized characteristic path length (λ), as well as small-world (σ) and global efficiencies (E_{global}), were calculated to uncover the topological properties of functional networks in MCI and NC groups (**Table 4**). Moreover, we employed Modified Greedy strategy to calculate the modularity scores

of the estimated FBNs (Newman, 2006). As expected, both groups fit $\gamma = C_p^{\text{real}} / C_p^{\text{rand}} > 1$, $\lambda = L_p^{\text{real}} / L_p^{\text{rand}} \approx 1$ and $\sigma = \gamma/\lambda > 1$. Thus, the functional networks of MCI patients

and NCs exhibit small-world topological attributes (Watts and Strogatz, 1998). This means that the brain networks of the two groups maintain complex, efficient neural architectures that optimize the balance between local specialization and global integration (Sporns and Zwi, 2004; Achard and Bullmore, 2007; Sporns, 2012). Further comparisons suggest that the small-world σ -values of MCI patients are lower than those of NCs, which indicates the disruption of the “economic small-world” (i.e., reductions in the effective segregation and integration of information in the brain network). Furthermore, we found the C_p -values and modularities (Q -values) in MCI patients to be significantly lower than those in NC groups ($P < 0.01$). These changes in C_p and modularity suggest reduced local information processing network segregation in MCI patients. Although there is no significant difference between MCIs and NCs in L_p and E_{global} , the lower values of these two global topological attributes in MCIs indicate decreased network integration.

Using the definition of “hubs” (Sporns, 2011), we identified hub nodes in MCI patients and NCs. As shown in **Table 5**, the common MCI and NC hub regions are located primarily in the bilateral superior temporal, bilateral heschl, right middle frontal, and left angular gyrus. Most are distributed in the DMN, auditory network, fronto-parietal task control network, and dorsal attention network. Moreover, it is notable that some hub nodes are present only in MCI patients but absent in NCs. These are the several hubs found in the right middle temporal and left middle frontal gyrus. In addition,

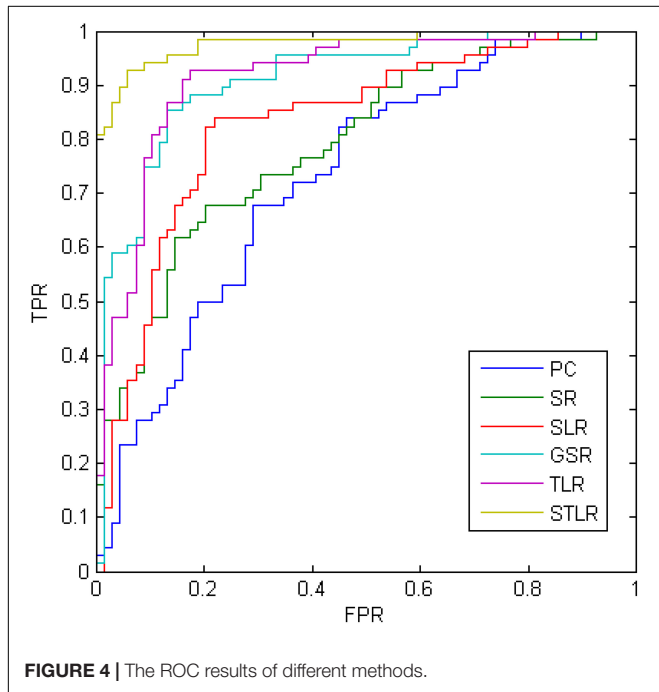


FIGURE 4 | The ROC results of different methods.

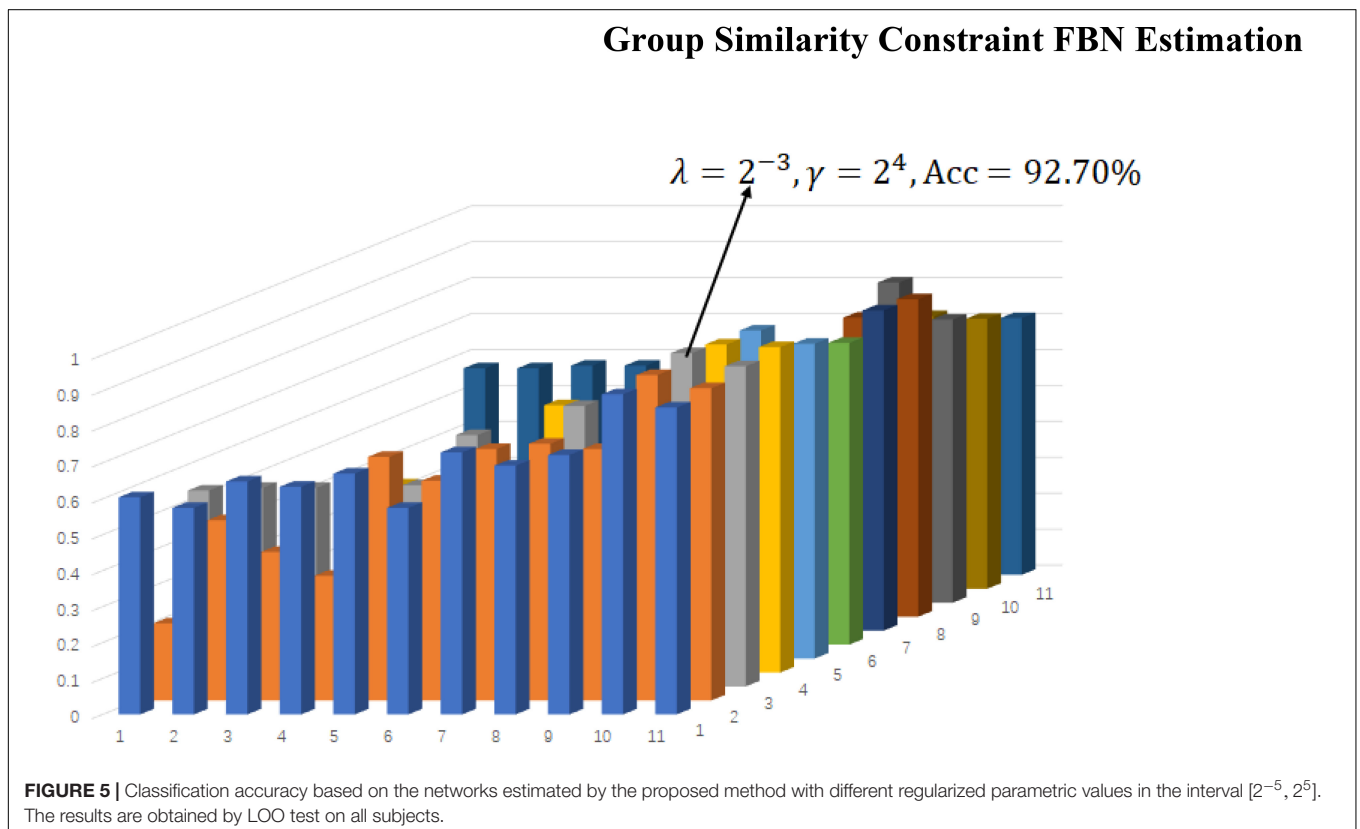
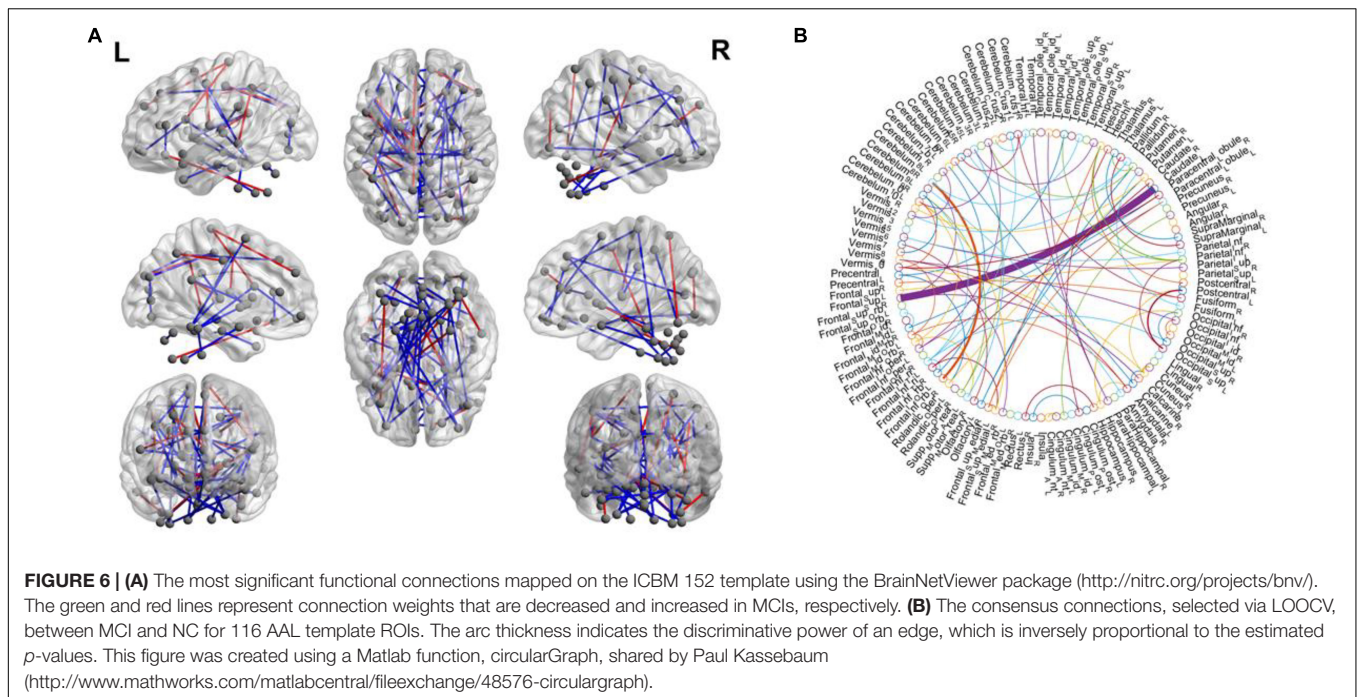


FIGURE 5 | Classification accuracy based on the networks estimated by the proposed method with different regularized parametric values in the interval $[2^{-5}, 2^5]$. The results are obtained by LOO test on all subjects.



there are some hub nodes found in HCs but not in MCI patients. They are located on the right inferior parietal and right middle frontal gyrus. These discriminative brain regions are distributed mainly in the DMN, fronto-parietal task control, and dorsal attention networks. Differences in subnetworks and corresponding brain regions play important roles in differential diagnosis of MCI relative to NC status.

DISCUSSION

The human brain is one of the most complex systems in the world. To ensure efficient brain information interactions, the FBN should have more “structures” than sparsity (Smith et al., 2011; Sporns, 2011). In this work, we incorporated a tensor low-rank regularizer to model the group similarity priors of the estimated FBNs. The MCI versus NC classification capabilities of the proposed models were verified using the ADNI dataset. Based on the results, we give the following brief discussion.

1. STLR-based methods were more accurate than baseline and state-of-art methods on our dataset. One possible reason is that the STLR scheme naturally incorporates additional information from inter-group subjects, and thus can produce clearer or more discriminative FBNs. It should also be noted that the proposed scheme is a flexible module. In addition to using SR-based models, it can be easily adopted using other FBN estimation models such as PC-based, Bayesian, or Granger causal-based networks. Also, we can incorporate other biological group priors into the tensor-based FBN estimation models.

2. The most discriminative functional connections and the corresponding predominating brain regions were discussed. By projecting brain regions with significant brain network functional connectivity differences and graph theory metrics to subnetworks, we found that the differences between MCI patients and NCs were distributed mainly in the DMN, dorsal

TABLE 3 | The top 20 brain regions (without the cerebellum) with largest number of discriminative connections.

AAL Number	Corresponding brain regions	Sub-networks
77	Thalamus_L	Subcortical network
85	Temporal_Mid_L	Dorsal attention network
6	Frontal_Sup_Orb_R	Default mode network
9	Frontal_Mid_Orb_L	Default mode network
38	Hippocampus_R	Default mode network
39	ParaHippocampal_L	Default mode network
61	Parietal_Inf_L	Dorsal attention network
62	Parietal_Inf_R	Dorsal attention network
70	Paracentral_Lobule_R	Sensory/somatomotor hand
71	Caudate_L	Fronto-parietal task control
72	Caudate_R	Fronto-parietal task control
75	Pallidum_L	Subcortical network
11	Frontal_Inf_Oper_L	Executive control network
13	Frontal_Inf_Tri_L	Executive control network
24	Frontal_Sup_Medial_R	Fronto-parietal task control
42	Amygdala_R	Subcortical network
45	Cuneus_L	Visual network
47	Lingual_L	Default mode network
73	Putamen_L	Saliency network
78	Thalamus_R	Subcortical network

TABLE 4 | Statistical result of topological properties between MCIs and NCs.

	MCI	NC
Cp*	0.216 ± 0.016	0.262 ± 0.022
Lp	9.459 ± 2.956	9.886 ± 1.712
γ	1.147 ± 0.069	1.186 ± 0.099
λ	1.056 ± 0.018	1.062 ± 0.019
σ	1.086 ± 0.052	1.117 ± 0.088
Eglobal	0.032 ± 0.003	0.033 ± 0.002
Q*	0.202 ± 0.000	0.959 ± 0.000

*Significant with FDR (0.05).

TABLE 5 | Hubs in MCI and NCs defined with the degree.

	AAL Number	Corresponding brain regions	Sub-networks	
MCI	88	Temporal_Pole_Mid_R	Default mode network	
	79	Heschl_L	Auditory network	
	10	Frontal_Mid_Orb_R	Default mode network	
	80	Heschl_R	Auditory network	
	81	Temporal_Sup_L	Auditory network	
	65	Angular_L	Default mode network	
	87	Temporal_Pole_Mid_L	Default mode network	
	9	Frontal_Mid_Orb_L	Fronto-parietal task control	
	83	Temporal_Pole_Sup_L	Auditory network	
	84	Temporal_Pole_Sup_R	Auditory network	
	NC	79	Heschl_L	Auditory network
		81	Temporal_Sup_L	Auditory network
		65	Angular_L	Default mode network
		66	Angular_R	Default mode network
87		Temporal_Pole_Mid_L	Default mode network	
10		Frontal_Mid_Orb_R	Default mode network	
80		Heschl_R	Auditory network	
62		Parietal_Inf_R	Dorsal attention network	
84		Temporal_Pole_Sup_R	Auditory network	
83		Temporal_Pole_Sup_L	Auditory network	

AAL, the automated anatomical labeling atlas.

attention, frontoparietal task, executive control, and auditory networks. Of these, the DMN had the most significant discriminative ability. Changes in these subnetworks were consistent with the results of previous cognitive function studies such as those on spatial attention (Rolle et al., 2017), executive function (Liao et al., 2019), and auditory function (Bi et al., 2018) that reference subnetworks in MCI patients. Moreover, the DMN has been regarded as the core part of a functional center (Liu et al., 2019) that is involved in episodic memory and is thought to be the major cognitive domain impaired during early-stage AD (Eyler et al., 2019). That the DMN contains the most distinguishing information for MCI identification was verified using our proposed methods. Furthermore, in our study, we located the predominant brain regions (i.e., the thalamus, middle temporal gyrus, hippocampus, parahippocampal gyrus, and inferior parietal and middle frontal gyri) for MCI diagnosis.

3. Brain network patterns are altered in MCI patients. Our study found that the global topological properties of MCI

patients and NCs fit the small-world attribute. That is, the brain networks of MCI and NC groups conform to “economic small-world” classification, which uses rapid, real-time information processing across separate brain regions to maximize efficiency with minimal cost and to render resilience against pathological attacks (Sporns and Zwi, 2004; Sporns, 2012; Liao et al., 2017). Statistical analysis suggested that the value of small-world σ was lower in MCI patients than in NCs, which indicated disruption of brain network integration and segregation. This MCI small-world result is consistent with previous research (Yu et al., 2018). Moreover, the significantly decreased C_p and modularity values noted in MCI patients further verified the reduction in brain network functional segregation. Lower C_p - and Q -values suggest less concentrated clustering of local connections and a weaker capacity for specialized processing within densely interconnected groups of brain regions in MCI patients (Rubinov and Sporns, 2010).

However, since the proposed scheme is a simple attempt to model the group similarity prior, there are several limitations in the proposed methods that should be improved upon via future work.

1. In this paper, we provide only simple verification to validate the effectiveness of the TL scheme and do not consider other factors (e.g., Atlas selection and data preprocessing). Therefore, we simply adopt the commonly used AAL atlas to define ROI. In the future, we would like to consider using a functional template (e.g., Power264) to alleviate this issue.
2. In this paper, we use only the tensor low-rank module to formulate the group similarity prior. In fact, the brain has a highly complex structure, and group similarity can be formulated into other formats. Therefore, we will use more abundant biological/physical priors to construct appropriate regular terms and further improve the current group-constraint model in future studies.
3. The global graph theory metrics (i.e., C_p , L_p , and small-world) were areas of focus in our study. However, nodal and other graph theory metrics could also be used to describe the complex topological mechanisms of brain networks. In future research, more graph theory metrics, such as the nodal shortest path length, local efficiency, and participant coefficient of modularity may be used to elaborate upon more specific local brain network topological properties.

CONCLUSION

Human brain patterns still need deep exploration. Thus, providing better brain descriptions remains challenging and meaningful. Inspired by the group similarity priors, we introduced the tensor based FBN estimation scheme. In particular, we proposed TLR and STLR to estimated FBN. More specially, we used the PARAFAC decomposition to capture

FBNs with low-rank topologies. Finally, we applied the estimated FBNs to classification. The results illustrate that the introduction of the group similarity constraint can effectively improve baseline method performance. The *post hoc* analysis further showed that more biologically meaningful functional brain connections were obtained by incorporating the group similarity prior.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <http://adni.loni.ucla.edu>.

AUTHOR CONTRIBUTIONS

All authors developed the proposed algorithm and architecture. WL and RL designed the evaluation experiments. XG, PW, and XH preprocessed the fMRI. XX and XH analyzed and interpreted the results of the data. All authors contributed to preparation of the article, figures, and charts.

REFERENCES

- Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3:e17. doi: 10.1371/journal.pcbi.0030017
- Alzheimer's Association, (2017). 2017 Alzheimer's disease facts and figures. *Alzheimers Dement.* 13, 325–373.
- Baggio, H. C., Sala-Llonch, R., Segura, B., Marti, M. J., Valdeoriola, F., Compta, Y., et al. (2014). Functional brain networks and cognitive deficits in Parkinson's disease. *Hum. Brain Mapp.* 35, 4620–4634. doi: 10.1002/hbm.22499
- Bain, L. J., Jedrzejewski, K., Morrison-Bogorad, M., Albert, M., Cotman, C., Hendrie, H., et al. (2008). Healthy brain aging: a meeting report from the Sylvan M. Cohen Annual Retreat of the University of Pennsylvania Institute on Aging. *Alzheimers Dement.* 4, 443–446.
- Bi, X.-A., Sun, Q., Zhao, J., Xu, Q., and Wang, L. (2018). Non-linear ICA analysis of resting-state fMRI in mild cognitive impairment. *Front. Neurosci.* 12:413. doi: 10.3389/fnins.2018.00413
- Brunetti, M., Belardinelli, P., Gratta, C. D., Pizzella, V., Penna, S. D., Ferretti, A., et al. (2006). Human brain activation elicited by the localization of sounds delivering at attended or unattended positions: an fMRI/MEG study. *Cogn. Process.* 7, 116–117.
- Chang, C. C., and Lin, C. J. (2007). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 389–396.
- Chao-Gan, Y., and Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* 4:13. doi: 10.3389/fnsys.2010.00013
- Chen, X., Zhang, H., Gao, Y., Wee, C. Y., Li, G., and Shen, D. (2016). High-order resting-state functional connectivity network for MCI classification. *Hum. Brain Mapp.* 37, 3282–3296. doi: 10.1002/hbm.23240
- Chen, X., Zhang, H., Zhang, L., Shen, C., Lee, S. W., and Shen, D. (2017). Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification. *Hum. Brain Mapp.* 38, 5019–5034. doi: 10.1002/hbm.23711
- Combettes, P. L., and Pesquet, J. C. (2011). "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, eds H. H. Bauschke, R. S. Burachik, P. L.

FUNDING

This work was partly supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1500501, KJ1600512, and KJ1600518), Natural Science Foundation Project of CQCSTC (cstc2018jcyjA27560398), Shanghai Municipal Commission of Health and Family Planning Science and Research Subjects (201740010), Shanghai Municipal Commission of Economy and Informatization, Special Fund for Artificial Intelligence Innovation and Development (XX-RGZN-01-19-4224), National Natural Science Foundation of China (Grant No. 81830059), and Shanghai Municipal Commission of Health and Family Planning Smart Medical Special Research Project (Grant No. 018ZHYL0105).

ACKNOWLEDGMENTS

We appreciate Charlesworth for correcting the English language usage, grammar, punctuation, and spelling. This manuscript has been released as a Pre-Print at Gao et al. (2019).

- Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, (New York, NY: Springer).
- Donoho, D. L., and Elad, M. (2003). Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization. *Proc. Natl. Acad. Sci.* 100, 2197–2202.
- Eyler, L. T., Elman, J. A., Hatton, S. N., Gough, S., Mischel, A. K., Hagler, D. J., et al. (2019). Resting state abnormalities of the default mode network in mild cognitive impairment: a systematic review and meta-analysis. *J. Alzheimers Dis.* 70, 107–120. doi: 10.3233/JAD-180847
- Fan, Y., and Browndyke, J. N. (2010). MCI diagnosis via manifold based classification of functional brain networks. *Alzheimers Dement.* 6:S16.
- Gao, X., Xu, X., Li, W., and Li, R. (2019). Group similarity constraint functional brain network estimation for mild cognitive impairment classification. *bioRxiv* [Preprint] doi: 10.1101/734574
- Gotts, S. J., Simmons, W. K., Milbury, L. A., Wallace, G. L., Cox, R. W., and Martin, A. (2012). Fractionation of social brain circuits in autism spectrum disorders. *Brain A J. Neurol.* 135(Pt 9), 2711–2725. doi: 10.1093/brain/aw160
- Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., et al. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch. Neurol.* 61, 59–66.
- Higgins, I., Kundu, S., and Guo, Y. (2018). Integrative Bayesian analysis of brain functional networks incorporating anatomical knowledge. *Neuroimage* 181, 263–278. doi: 10.1016/j.neuroimage.2018.07.015
- Huang, S., Li, J., Sun, L., Liu, J., Wu, T., Chen, K., et al. (2009). "Learning brain connectivity of Alzheimer's disease from neuroimaging data," in *Paper Presented at the Advances in Neural Information Processing Systems 22: Conference on Neural Information Processing Systems 2009. Proceedings of A Meeting Held 7-10 December 2009*, Vancouver, BC.
- Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., et al. (2010). Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *Neuroimage* 50, 935–949. doi: 10.1016/j.neuroimage.2009.12.120
- Jack, C. R., Bernstein, M. A., Fox, N. C., Paul, T., Danielle, H., Bret, B., et al. (2010). The Alzheimer's Disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691.

- Jin, H. L., Durand, R., Gradinaru, V., Feng, Z., Goshen, I., Kim, D. S., et al. (2010). Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature* 465, 788–792. doi: 10.1038/nature09108
- Kevin, W., Doug, W., Matthias, S., and Gerhard, S. (2008). Correspondence of visual evoked potentials with FMRI signals in human visual cortex. *Brain Topogr.* 21, 86–92. doi: 10.1007/s10548-008-0069-y
- Lee, H., Lee, D. S., Kang, H., Kim, B. N., and Chung, M. K. (2011). Sparse brain network recovery under compressed sensing. *IEEE Trans. Med. Imaging* 30, 1154–1165. doi: 10.1109/TMI.2011.2140380
- Li, H., Zhu, X., and Fan, Y. (2018). Identification of multi-scale hierarchical brain functional networks using deep matrix factorization. *Med. Image Comput. Assist. Interv.* 11072, 223–231. doi: 10.1007/978-3-030-00931-1_26
- Li, W., Qiao, L., Zhang, L., Wang, Z., and Shen, D. (2019a). Functional brain network estimation with time series self-scrubbing. *IEEE J. Biomed. Health Inform.* 23, 2494–2504. doi: 10.1109/JBHI.2019.2893880
- Li, W., Wang, Z., Zhang, L., Qiao, L., and Shen, D. (2017). Remodeling Pearson's correlation for functional brain network estimation and autism spectrum disorder identification. *Front. Neuroinform.* 11:55. doi: 10.3389/fninf.2017.00055
- Li, W., Zhang, L., Qiao, L., and Shen, D. (2019b). Towards a better estimation of functional brain network for mild cognitive impairment identification: a transfer learning view. *IEEE J. Biomed. Health Inform.* doi: 10.1109/JBHI.2019.2934230 [Epub ahead of print].
- Liang, X., Connelly, A., and Calamante, F. (2016). A novel joint sparse partial correlation method for estimating group functional networks. *Hum. Brain Mapp.* 37, 1162–1177. doi: 10.1002/hbm.23092
- Liang, X., Vaughan, D. N., Connelly, A., and Calamante, F. (2018). A novel group-fused sparse partial correlation method for simultaneous estimation of functional networks in group comparison studies. *Brain Topogr.* 31, 364–379. doi: 10.1007/s10548-017-0615-6
- Liao, X., Vasilakos, A. V., and He, Y. (2017). Small-world human brain networks: perspectives and challenges. *Neurosci. Biobehav. Rev.* 77, 286–300. doi: 10.1016/j.neubiorev.2017.03.018
- Liao, Y.-Y., Chen, I.-H., Lin, Y.-J., Chen, Y., and Hsu, W.-C. (2019). Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. *Front. Aging Neurosci.* 11:162. doi: 10.3389/fnagi.2019.00162
- Liu, C., Yen, C. C.-C., Szczupak, D., Frank, Q. Y., Leopold, D. A., and Silva, A. C. (2019). Anatomical and functional investigation of the marmoset default mode network. *Nat. Commun.* 10:1975. doi: 10.1038/s41467-019-09813-7
- Liu, F., Wee, C. Y., Chen, H., and Shen, D. (2012). "Inter-modality relationship constrained multi-task feature selection for AD/MCI classification," in *Paper Presented at the Medical Image Computing & Computer-assisted Intervention: Miccai International Conference on Medical Image Computing & Computer-assisted Intervention*, Heidelberg.
- Liu, X., Bourennane, S., and Fossati, C. (2012). Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis. *IEEE Trans. Geosci. Remote Sens.* 50, 3717–3724.
- Meinshausen, N., and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *Ann. Stat.* 34, 1436–1462.
- Misra, C., Fan, Y., and Davatzikos, C. (2009). Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *Alzheimers Dement.* 44, 1415–1422. doi: 10.1016/j.neuroimage.2008.10.031
- Nesterov, Y. E. (1983). A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR* 269, 543–547.
- Newman, M. E. (2006). "Modularity and community structure in networks," in *Paper Presented at the APS March Meeting*, Ridge, NY.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018
- Qiao, L., Han, Z., Kim, M., Teng, S., Zhang, L., and Shen, D. (2016). Estimating functional brain networks by incorporating a modularity prior. *Neuroimage* 141, 399–407. doi: 10.1016/j.neuroimage.2016.07.058
- Rolle, C. E., Anguera, J. A., Skinner, S. N., Voytek, B., and Gazzaley, A. (2017). Enhancing spatial attention and working memory in younger and older adults. *J. Cogn. Neurosci.* 29, 1483–1497. doi: 10.1162/jocn_a_01159
- Rosa, M. J., Portugal, L., Hahn, T., Fallgatter, A. J., Garrido, M. I., Shawe-Taylor, J., et al. (2015). Sparse network-based models for patient classification using fMRI. *Neuroimage* 105, 493–506. doi: 10.1016/j.neuroimage.2014.11.021
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., et al. (2013). Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* 17, 666–682. doi: 10.1016/j.tics.2013.09.016
- Sporns, O. (2011). *Networks of the Brain*. Cambridge, MA: MIT Press.
- Sporns, O. (2012). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125.
- Sporns, O., and Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics* 2, 145–162.
- Stam, C. J. (2014). Modern network science of neurological disorders. *Nat. Rev. Neurosci.* 15, 683–695. doi: 10.1038/nrn3801
- Supekar, K., Menon, V., Rubin, D., Musen, M., and Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Comput. Biol.* 4:e1000100. doi: 10.1371/journal.pcbi.1000100
- Theije, C. G. M. D., Wu, J., Silva, S. L. D., Kamphuis, P. J., Garssen, J., Korte, S. M., et al. (2011). Pathways underlying the gut-to-brain connection in autism spectrum disorders as future targets for disease management. *Eur. J. Pharmacol.* 668(Suppl. 1), S70–S80. doi: 10.1016/j.ejphar.2011.07.013
- Tomioka, R., and Sugiyama, M. (2009). Dual augmented lagrangian method for efficient sparse reconstruction. *IEEE Signal Process. Lett.* 16, 1067–1070. doi: 10.1364/AO.51.005676
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- Wang, Y., Lin, K., Qi, Y., Lian, Q., Feng, S., Pan, G., et al. (2018). Estimating brain connectivity with varying length time lags using recurrent neural network. *IEEE Trans. Biomed. Eng.* 65, 1953–1963. doi: 10.1109/TBME.2018.2842769
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442.
- Wee, C. Y., Yap, P. T., and Shen, D. (2016). Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks. *CNS Neurosci. Ther.* 22, 212–219. doi: 10.1111/cns.12499
- Wee, C. Y., Yap, P. T., Zhang, D., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2012). Identification of MCI individuals using structural and functional connectivity networks. *Neuroimage* 59, 2045–2056. doi: 10.1016/j.neuroimage.2011.10.015
- Wee, C. Y., Yap, P. T., Zhang, D., Wang, L., and Shen, D. (2014). Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct. Funct.* 219, 641–656. doi: 10.1007/s00429-013-0524-8
- Yu, K., Wang, X., Li, Q., Zhang, X., Li, X., and Li, S. (2018). Individual morphological brain network construction based on multivariate euclidean distances between brain regions. *Front. Hum. Neurosci.* 12:204. doi: 10.3389/fnhum.2018.00204
- Yu, R., Zhang, H., An, L., Chen, X., Wei, Z., and Shen, D. (2016). *Correlation-Weighted Sparse Group Representation for Brain Network Construction in MCI Classification*. Cham: Springer International Publishing.

- Yuan, M., and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc.* 68, 49–67.
- Zhou, L., Wang, L., and Ogunbona, P. (2014). “Discriminative sparse inverse covariance matrix: application in brain functional network classification,” in *Paper Presented at the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH.
- Zhou, Y., Qiao, L., Li, W., Zhang, L., and Shen, D. (2018). Simultaneous estimation of low- and high-order functional connectivity for identifying mild cognitive impairment. *Front. Neuroinform.* 12:3. doi: 10.3389/fninf.2018.00003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Gao, Xu, Hua, Wang, Li and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.