



## RESEARCH ARTICLE

# Longitudinally consistent estimates of intrinsic functional networks

Qingyu Zhao<sup>1</sup> | Dongjin Kwon<sup>1,2</sup> | Eva M. Müller-Oehring<sup>1,2</sup> | Anne-Pascale Le Berre<sup>1</sup> | Adolf Pfefferbaum<sup>1,2</sup> | Edith V. Sullivan<sup>1</sup> | Kilian M. Pohl<sup>2</sup><sup>1</sup>Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California<sup>2</sup>Center for Health Sciences, SRI International, Menlo Park, California**Correspondence**Kilian M. Pohl, Center for Health Sciences, SRI International, Menlo Park, CA 94025.  
Email: kilian.pohl@sri.com**Funding information**

National Institutes of Health, Grant/Award Numbers: AA017923, AA010723, AA013521, AA017168, AA021697

**Abstract**

Increasing numbers of neuroimaging studies are acquiring data to examine changes in brain architecture by investigating intrinsic functional networks (IFN) from longitudinal resting-state functional MRI (rs-fMRI). At the subject level, these IFNs are determined by cross-sectional procedures, which neglect intra-subject dependencies and result in suboptimal estimates of the networks. Here, a novel longitudinal approach simultaneously extracts subject-specific IFNs across multiple visits by explicitly modeling functional brain development as an essential context for seeking change. On data generated by an innovative simulation based on real rs-fMRI, the method was more accurate in estimating subject-specific IFNs than cross-sectional approaches. Furthermore, only group-analysis based on longitudinally consistent estimates identified significant developmental effects within IFNs of 246 adolescents from the National Consortium on Alcohol and NeuroDevelopment in Adolescence (NCANDA) study. The findings were confirmed by the cross-sectional estimates when the corresponding group analysis was confined to the developmental effects. Those effects also converged with current concepts of neurodevelopment.

**KEYWORDS**

intrinsic functional networks, longitudinal analysis, resting-state fMRI

## 1 | INTRODUCTION

The functional architecture of the brain has been widely studied using resting state-functional MRI (rs-fMRI), which measures spontaneous, low frequency fluctuations in the blood-oxygen-level dependent (BOLD) signal (Biswal, 2012; Buckner, Krienen, & Yeo, 2013). Research based on rs-fMRI has revealed a number of intrinsic functional networks (IFNs) that consist of anatomically separate but functionally connected brain regions demonstrating synchronous BOLD fluctuations at rest (Rosazza & Minati, 2011; Shen, 2015). Key networks associated with maturing of functional organization (Oron Semper, Murillo, & Bernacer, 2016; Power, Fair, Schlaggar, & Petersen, 2010) are the default mode network, sensory and motor networks, the executive control network, and visual networks (Fair et al., 2008; Gu et al., 2015; Sherman et al., 2014). Abnormalities in such functional development have been linked to neurological disorders or psychiatric diseases (Chase, 2014; Park et al., 2017). To further understanding of the maturation of functional networks, an increasing number of

neuroimaging studies are acquiring longitudinal rs-fMRI (Alexander et al., 2017; Kooijman et al., 2016; Nora et al., 2018). For example, the multi-site National Consortium on Alcohol and NeuroDevelopment in Adolescence (NCANDA) study (Brown et al., 2015; Müller-Oehring et al., 2018; Pfefferbaum et al., 2016; Sullivan et al., 2016) is investigating adolescent neurodevelopment and its deviation due to alcohol use by acquiring rs-fMRI of each study participant before initiating appreciable drinking and following them annually.

Existing longitudinal studies have applied cross-sectional approaches to identify subject-specific IFNs; that is, IFNs are estimated independently for rs-fMRI of each visit (Damaraju et al., 2014; Gao, Alcauter, Smith, Gilmore, & Lin, 2015; Odish et al., 2015; van der Horn et al., 2017). Popular cross-sectional approaches (Harrison et al., 2015; van den Heuvel & Sporns, 2013) for this task are based on expert-guided, seed-based analysis (Fox et al., 2005) and data-driven *spatial Independent Component Analysis (ICA)* (Calhoun et al., 2001; Calhoun, Adali, Pearson, & Pekar, 2001b; Schmithorst & Holland, 2004). Of these two approaches, spatial ICA is generally more

sensitive for detecting subtle differences in functional connectivity within an IFN (Koch et al., 2010). ICA encodes each IFN as an independent component, which is represented by a *spatial map* coupled with its characteristic BOLD signal. The spatial map captures the consistency between the characteristic BOLD signal and the BOLD signal at each voxel. Voxels with high consistency define regions with coherent engagement in oscillatory activity so that a spatial map provides a natural measure of functional connectivity among voxels within an IFN. To compute spatial maps of each subject, spatial ICA is first applied to the rs-fMRI images of a cohort resulting in a set of *group-level* IFNs (Beckmann, Mackay, Filippini, & Smith, 2009; Calhoun, Adali, Pearson, & Pekar, 2001a, 2001b). The group-level IFNs are then modified according to each individual rs-fMRI using methods such as dual-regression (Beckmann et al., 2009), back-reconstruction (Calhoun, Adali, Pearson, & Pekar, 2001a), or Group-Information-Guided ICA (GIG-ICA) (Du & Fan, 2013). Specific to each subject and visit, the resulting IFNs are then analyzed by statistical group analysis in order to identify, for example, *group-level* developmental effects (Gao et al., 2015; Stoffers et al., 2015).

While successfully applied to cross-sectional (Lee, Smyser, & Shimony, 2013) and longitudinal studies (Damaraju et al., 2014; Odish et al., 2015), subject-specific IFNs estimated by cross-sectional approaches do not consider intra-subject dependencies. Accounting for intra-subject dependencies is crucial for accurately extracting IFNs from longitudinal data as the rs-fMRIs of an individual can highly vary across visits due to factors such as thermal noise, scanner drift, and head motion (Bright & Murphy, 2013, 2015). These complex factors can induce biologically implausible changes within the sequences of IFNs estimated from the longitudinal rs-fMRI of a subject. The resulting large variances often observed within subject-specific IFNs complicate the discovery of the underlying developmental effects within IFNs of a cohort. Previous studies usually rely on group-level statistical models to account for such variances (Madhyastha et al., 2017). The method presented herein aims to reveal the gain of modeling intra-subject dependencies within the procedure of subject-specific IFN estimation prior to the group-level analysis, an uncharted topic.

Specifically, we present *longitudinal ICA*, which estimates consistent subject-specific IFNs from longitudinal rs-fMRI of at least three visits. To do so, we expand GIG-ICA to extract IFNs at each visit and to model explicitly biologically-plausible variation in IFNs across visits. We hypothesize that longitudinal ICA produces more precise subject-specific IFNs than the ones generated separately for each visit. To support this hypothesis, longitudinal ICA is compared with cross-sectional approaches on synthetic datasets with known ground-truth IFNs and on the rs-fMRI of 246 no-to-low drinking adolescents of the NCANDA study (Müller-Oehring et al., 2018) that participated in three annual visits. The comparison between approaches is based on a statistical group analysis of the corresponding subject-specific IFN-estimates supporting the hypothesis of the NCANDA study that IFNs change consistently with neuromaturation concepts during adolescence (Müller-Oehring et al., 2018).

Our contributions to the field are threefold: (a) We propose a novel longitudinal approach called longitudinal ICA, which is grounded in the cross-sectional GIG-ICA for estimating subject-specific IFNs across multiple visits; (b) We propose a procedure for validating

approaches (cross-sectional and longitudinal) in estimating subject-specific IFNs from longitudinal rs-fMRI independent from the hypothesis of the statistical group analysis. The procedure is based on a novel longitudinal rs-fMRI simulation, which creates synthetic ground-truth data from real rs-fMRI of the NCANDA study; (c) To the best of our knowledge, this is the first study that applies ICA-based approaches for investigating developmental effects of IFNs using longitudinal rs-fMRI of more than two visits.

## 2 | METHODS AND MATERIALS

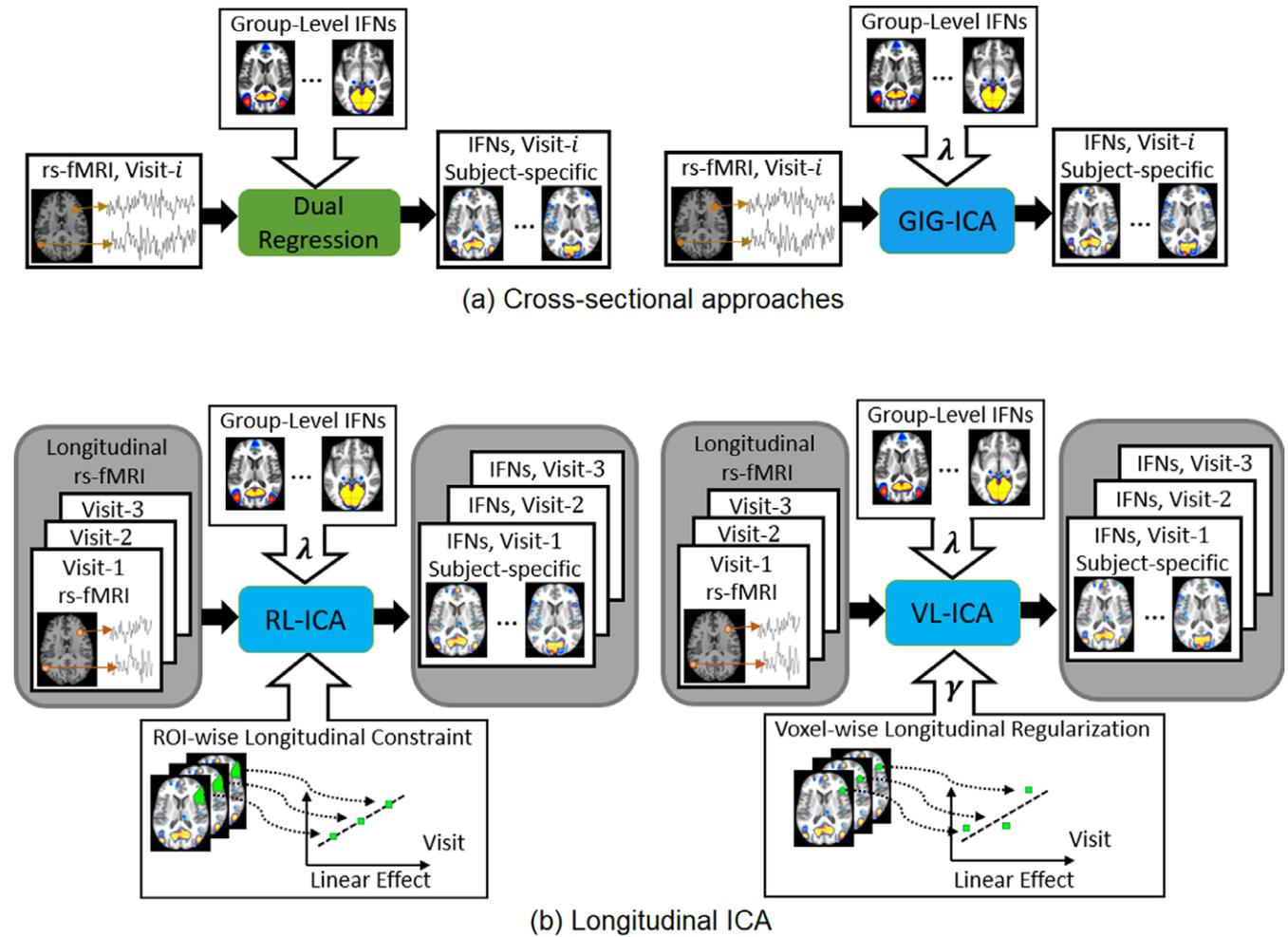
Brain functional development during the entire adolescent period may exhibit linear and nonlinear group-level trajectories (Casey, Jones, & Hareb, 2008). As for the NCANDA study, the longitudinal rs-fMRI of each adolescent was only collected in a two-year span, during which changes within a subject-specific IFN across visits are expected to be linear (Vetter et al., 2017). Section 2.1 introduces two implementations of longitudinal ICA that model these linear changes at different spatial scales (see also Figure 1): voxel-wise longitudinal ICA (VL-ICA) ensures that the variation is linear with respect to each voxel and ROI-wise longitudinal ICA (RL-ICA) with respect to each brain region-of-interest (ROI). Section 2.2 describes the experimental setup for testing the accuracy of estimating subject-specific IFNs under different imaging characteristics (e.g., noise level). The tests are based on synthetic datasets generated by the SimTB software (Erhardt, Allen, Wei, Eichele, & Calhoun, 2012) with benchmarks created by existing cross-sectional approaches Dual Regression (DUAL) and GIG-ICA (see Figure 1). Section 2.3 describes the application of the four approaches to the NCANDA dataset. The experimental setup consists of a data-driven parameter exploration of longitudinal ICA based on the novel simulation and a statistical group analysis testing for group-level developmental effects across subject-specific IFNs.

### 2.1 | Estimating subject-specific IFNs

For technically inclined readers, we now present the mathematical models of the proposed longitudinal implementations. To do so, the cross-sectional spatial ICA and its derivative GIG-ICA are first introduced. We then derive VL-ICA by adding a soft constraint to the formulation of GIG-ICA. The soft constraint encourages the change across visits at each voxel of a subject-specific IFN to be linear. In contrast, RL-ICA solves the minimization problem defined by GIG-ICA but confines the search space to subject-specific IFNs, whose change with respect to the average functional connectivity within each ROI is linear across visits.

#### 2.1.1 | Spatial ICA

Let  $\mathbf{X} \in \mathbb{R}^{M \times L}$  be the rs-fMRI image of one subject at one visit, where  $M$  is the number of voxels and  $L$  is the length of the time course. Spatial ICA assumes that the rs-fMRI image  $\mathbf{X}$  is a linear mixture of independent components. Let  $\mathbf{S} := [s_1, s_2, \dots, s_C]$  denote an  $M \times C$  matrix representing  $C$  independent components and  $\mathbf{A} \in \mathbb{R}^{C \times L}$  be the “mixing” matrix of those components, then the linear relationship defining ICA is



**FIGURE 1** (a) Two commonly-used cross-sectional approaches that estimate subject-specific IFNs for each visit independently. (b) The two variants of longitudinal ICA (RL-ICA, VL-ICA) jointly estimate IFNs across visits while explicitly modeling developmental effects [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

$$\mathbf{X} = \mathbf{S} \cdot \mathbf{A}. \quad (1)$$

Each independent component in  $\mathbf{S}$  can correspond either to an IFN or an artifact (such as physiological noise). If the  $i^{\text{th}}$  component corresponds to an IFN,  $s_i$  is also known as the spatial map of that network. Without loss of generality, we assume the  $C$  components all correspond to functional networks.

Now directly determining those components only from  $\mathbf{X}$  is an ill-posed problem. Instead, subject-specific PCA (principal component analysis) first reduces the dimension of  $\mathbf{X}$ . Afterwards, the reduced rs-fMRI sequence is prewhitened, that is, its covariance matrix is the identity ( $E[\mathbf{X}^T \mathbf{X}] = \mathbf{I}$ ) (Calhoun, Adali, Pearson, & Pekar, 2001a; Du & Fan, 2013). For simplicity, let  $\mathbf{X} \in \mathbb{R}^{M \times L}$  now denote the “prewhitened” rs-fMRI image with the reduced time course being of length  $L$ . Then the components  $\mathbf{S}$  can be recovered by computing an orthogonal demixing matrix  $\mathbf{W}$ , such that

$$\mathbf{S} = \mathbf{X}\mathbf{W}, \mathbf{W}^T \mathbf{W} = \mathbf{I} \quad (2)$$

and  $\mathbf{W} := [w_1, \dots, w_C]$  contains  $C$  demixing vectors, that is,  $s_i = \mathbf{X}w_i$ .  $\mathbf{W}$  is optimized such that the resulting components in  $\mathbf{S}$  achieve maximum statistical independence (Hyvarinen & Oja, 2000). Let  $J$  denote

such a function for measuring statistical independence. The optimization over  $\mathbf{W}$  can be expressed in the form of

$$\underset{\mathbf{W}}{\operatorname{argmin}} -J(\mathbf{W}), \text{ s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I}. \quad (3)$$

In this study, we chose the specific implementation of  $J$  as defined in (Du & Fan, 2013), which measures the negentropy (Hyvarinen & Oja, 2000) of a random variable as a proxy of statistical independence.

### 2.1.2 | Group-information-guided ICA

The goal of GIG-ICA is to simultaneously minimize the original spatial-ICA objective (Equation 3) and the dissimilarity between subject-specific estimates  $\mathbf{S}$  and the group-level estimates  $\tilde{\mathbf{S}} := [\tilde{s}_1, \dots, \tilde{s}_C]$ . Let  $D$  denote a function for measuring the dissimilarity, also called the *group-level regularization*. GIG-ICA then optimizes the following objective function,

$$\underset{\mathbf{W}}{\operatorname{argmin}} -J(\mathbf{W}) + \lambda \cdot D(\mathbf{W}), \text{ s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I}. \quad (4)$$

The parameter  $\lambda \in \mathbb{R}^+$  weighs the importance of the group-level regularization  $D$ , which in our implementation is defined with respect to the Frobenius norm ( $\|\cdot\|_F^2$ ) normalized by the number of voxels  $M$ :

$$\mathcal{D}(\mathbf{W}) = \frac{1}{M} \left\| \mathbf{X}\mathbf{W} - \hat{\mathbf{S}} \right\|_F^2. \quad (5)$$

### 2.1.3 | Implementations of longitudinal ICA

Longitudinal ICA now assumes that each IFN changes linearly across  $V$  visits, for example, with respect to the sequence of the  $i^{\text{th}}$  independent component  $\{s_i^1, \dots, s_i^V\}$ . Note, these sequences of independent components are now derived from longitudinal rs-fMRI  $\{\mathbf{X}^1, \dots, \mathbf{X}^V\}$ , where rs-fMRI  $\mathbf{X}^j$  is the time series acquired of a subject at the  $j^{\text{th}}$  visit.

#### VL-ICA: voxel-wise regularization

VL-ICA assumes the functional connectivity within an IFN at each voxel changes linearly across visits. Let  $t_j$  denote the time between visit  $j-1$  and visit  $j$ . Then the voxel-wise linear model can be expressed as

$$\frac{s_i^j - s_i^{j-1}}{t_j} = \frac{s_i^{j+1} - s_i^j}{t_{j+1}} \quad (6)$$

for  $j \in \{2, \dots, V-1\}$  and  $i \in \{1, \dots, C\}$ .

To rewrite of the above constraint into a matrix form, let  $\hat{\mathbf{W}} := [\mathbf{W}^1, \dots, \mathbf{W}^V]^T$  be the demixing matrix over all visits such that  $\mathbf{S}^j = \mathbf{X}^j \hat{\mathbf{W}}^j$ . Furthermore, let  $\mathbf{I}^M$  denote an  $M \times M$  identity matrix and let  $\mathbf{B}_{\text{vox}} \in \mathbb{R}^{M(V-2)LV}$  be

$$\mathbf{B}_{\text{vox}} := \begin{bmatrix} t_3 \mathbf{I}^M - (t_2 + t_3) \mathbf{I}^M & t_2 \mathbf{I}^M & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & t_V \mathbf{I}^M - (t_{V-1} + t_V) \mathbf{I}^M & t_{V-1} \mathbf{I}^M \end{bmatrix} \times \text{diag}(\mathbf{X}^1, \dots, \mathbf{X}^V). \quad (7)$$

Based on Equation 2, the constraint of Equation 6 with respect to all spatial maps can be rewritten as

$$\mathbf{B}_{\text{vox}} \hat{\mathbf{W}} = \mathbf{0}. \quad (8)$$

The idea behind VL-ICA is to optimize the GIG-ICA objective function (Equation 4) for each individual visit while satisfying the longitudinal linear constraint of Equation 8. In practice, the number of voxels is much larger than the time course length ( $M \gg L$ ), which means Equation 8 is an over-determined homogeneous linear system that often does not have a solution. To overcome this issue, Equation 8 is integrated into the optimization as an additional energy term, which we refer to as the voxel-wise *longitudinal regularization*. As such, VL-ICA optimizes a weighted sum of the GIG-ICA objective and the longitudinal regularization:

$$\begin{aligned} \underset{\hat{\mathbf{W}}}{\text{argmin}} E(\hat{\mathbf{W}}) &= \sum_{j=1}^V \left[ -J(\mathbf{W}^j) + \lambda D(\mathbf{W}^j) \right] + \gamma \frac{1}{M} \left\| \mathbf{B}_{\text{vox}} \hat{\mathbf{W}} \right\|_F^2, \\ \text{s.t. } \mathbf{W}^j \mathbf{W}^j &= \mathbf{I}, \text{ for } j = 1, \dots, V, \end{aligned} \quad (9)$$

where  $\gamma$  is the weighting parameter balancing the importance of the longitudinal regularization over the GIG-ICA objective function. Note, the longitudinal regularization is also normalized by  $M$  (i.e., the number of voxels), which is consistent with the definition of  $J$  (Du & Fan, 2013; Hyvarinen & Oja, 2000) and  $D$  (Equation 5).

#### RL-ICA: ROI-wise constraints

To avoid setting  $\gamma$  in Equation 9, the second implementation enforces the linear constraint with respect to each ROI (instead of voxel). To be specific, given a predefined parcellation of the brain with  $P$  ROIs and assuming that the voxels in each ROI share similar BOLD fluctuations, let  $k_p \in \mathbb{R}^M$  denote a binary indicator function associated with the  $p^{\text{th}}$  ROI and  $\mathbf{K} := [k_1, \dots, k_P]^T$  the collection of those binary indicator functions. RL-ICA assumes the average functional connectivity in each ROI changes linearly, that is,

$$\frac{k_p^T (s_i^j - s_i^{j-1})}{t_j} = \frac{k_p^T (s_i^{j+1} - s_i^j)}{t_{j+1}} \quad (10)$$

for  $j \in \{2, \dots, V-1\}$ ,  $i \in \{1, \dots, C\}$  and  $p \in \{1, \dots, P\}$ . Now define  $\mathbf{B}_{\text{ROI}} \in \mathbb{R}^{P(V-2)LV}$  as

$$\mathbf{B}_{\text{ROI}} := \begin{bmatrix} t_3 I^P - (t_2 + t_3) I^P & t_2 I^P & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & t_V I^P - (t_{V-1} + t_V) I^P & t_{V-1} I^P \end{bmatrix} \times \text{diag}(\mathbf{K}\mathbf{X}^1, \dots, \mathbf{K}\mathbf{X}^V). \quad (11)$$

then the previous constraint over all spatial maps can be rewritten as

$$\mathbf{B}_{\text{ROI}} \hat{\mathbf{W}} = \mathbf{0}. \quad (12)$$

Note that if  $P(V-2) < LV$ , Equation 12 is an under-determined system, which can be integrated as a hard constraint into the optimization:

$$\begin{aligned} \underset{\hat{\mathbf{W}}}{\text{argmin}} E(\hat{\mathbf{W}}) &= \sum_{j=1}^V \left[ -J(\mathbf{W}^j) + \lambda D(\mathbf{W}^j) \right], \\ \text{s.t. } \mathbf{B}_{\text{ROI}} \hat{\mathbf{W}} &= \mathbf{0}, \text{ and } \mathbf{W}^j \mathbf{W}^j = \mathbf{I}, \text{ for } j = 1, \dots, V \end{aligned} \quad (13)$$

By specifying the linear model as a hard constraint, RL-ICA depends on the same set of parameters as GIG-ICA.

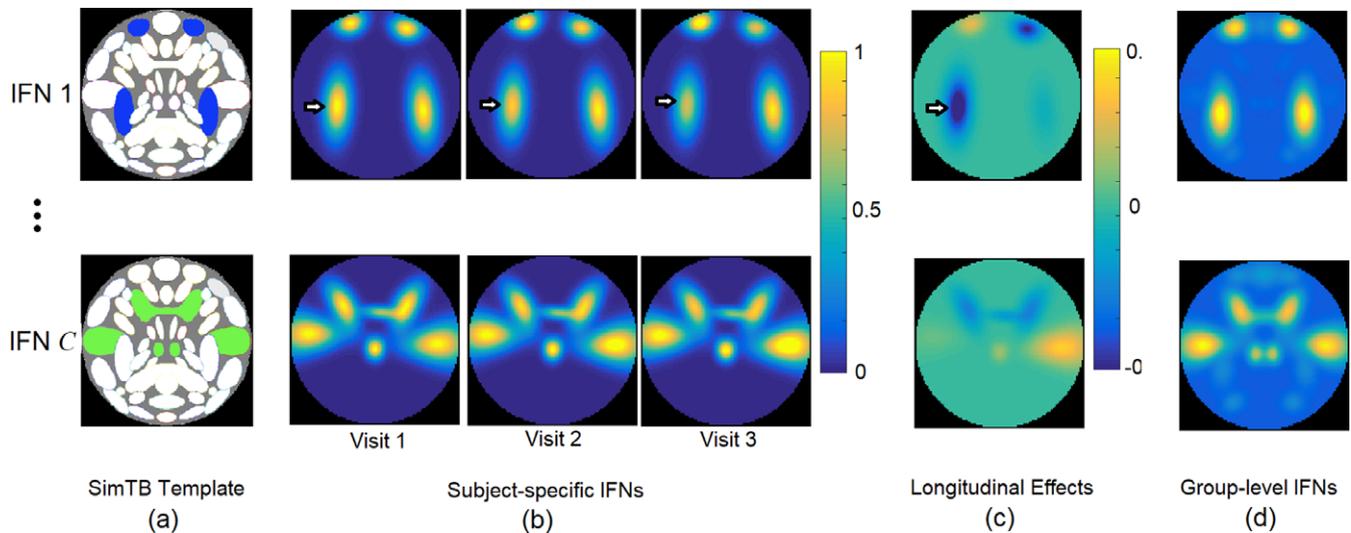
#### Solving the objectives of GIG-ICA-based approaches

The numerical algorithm described in Appendix A<sup>1</sup> efficiently determines the solution to the objective functions of GIG-ICA, VL-ICA, and RL-ICA by adopting a multi-splitting method (Lai & Osher, 2014). The algorithm can simultaneously optimize over multiple independent components to guarantee their statistical independence, which in our experience is superior to the traditional “one-unit-pursuit” strategy adopted in the original GIG-ICA implementation (Du & Fan, 2013).

## 2.2 | Synthetic experiments based on SimTB

To test the robustness of the implementations, 22 synthetic datasets were generated based on SimTB simulation (Allen et al., 2014; Erhardt et al., 2012). Each dataset consisted of 50 longitudinal rs-fMRIs or “subjects”. Each longitudinal rs-fMRI contained three synthetic “visit-specific” rs-fMRI images. Specifically, SimTB first simulated for each subject a set of  $C$  IFNs (e.g.,  $C = 3$  in Figure 2), which were called the *ground-truth* IFNs. These IFNs were generated according to  $C$  templates constructed in SimTB (Figure 2a) Then, the IFNs were randomly

<sup>1</sup>Matlab implementation is available at <http://sibis.sri.com>



**FIGURE 2** (a) Each longitudinal rs-fMRI (subject) was simulated from  $C$  synthetic networks, each related to four ROIs randomly selected from the SimTB template (Erhardt et al., 2012). (b) The  $C$  synthetic IFNs of a “subject” contained random translation, rotation, and linear longitudinal effects across the three visits. (c) The corresponding subject-specific longitudinal effects represented as the slope across the three visits. The arrow points to the ROI with the largest slope. (d) Group-level estimation of the  $C$  IFNs based on the baseline rs-fMRI images of 50 longitudinal rs-fMRIs [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

moved and rotated separately for each subject. To simulate developmental effects, the IFNs were linearly altered across the three visits, such as in the example shown in Figure 2b,c. Next, the rs-fMRI images were generated by SimTB based on the  $C$  ground-truth subject-specific IFNs. Finally, different approaches were applied to the simulated longitudinal rs-fMRI to estimate subject-specific IFNs, which were compared to the ground-truth for measuring the accuracy of the approaches. Appendix B provides additional details on the simulation procedure of the 22 synthetic datasets while the remainder of this section describes the procedure for measuring the accuracy of an approach (DUAL, GIG-ICA, RL-ICA, or VL-ICA) in estimating subject-specific IFNs on each synthetic dataset.

Measuring the accuracy of each approach on a dataset first required finding the optimal parameters for the three GIG-ICA-based approaches (GIG-ICA, RL-ICA, VL-ICA). The search space for  $\lambda$  (relevant to the three approaches) and  $\gamma$  (only relevant to VL-ICA) was  $[0, 1,000]$ . This parameter exploration aimed to achieve the most accurate subject-specific IFN estimation independent from the hypothesis of group-level analysis. Therefore, the optimal parameter setting of an approach was defined as the one that achieved the optimal group-mean square error (MSE) score on the dataset. For each dataset, the group-MSE score for each approach (including DUAL) with respect to a specific setting was determined by the following procedure:

1. Group-level IFNs (Figure 2b) were estimated by applying probabilistic ICA of MELODIC FSL V5.0.10 (Beckmann et al., 2009) to the collection of 50 “baseline” rs-fMRI images.
2. Based on the group-level IFNs, subject-specific IFNs ( $50 \text{ subjects} \times C \text{ networks} \times 3 \text{ visits}$ ) were estimated by the approach (e.g., VL-ICA). For GIG-ICA-based approaches, the time-course dimension was reduced to 29 as the rs-fMRI image was a linear mixture of time courses associated with the 29 ROIs defined according to the masks provided by the SimTB software (Appendix B). These ROIs were also used for constructing the ROI-wise longitudinal constraints of RL-ICA.

3. The accuracy in estimating the IFNs associated with one subject was captured by the subject-specific MSE. Specifically, each of the estimated and ground-truth subject-specific IFNs was first normalized by computing their z-scores. For each IFN and visit, the similarity between the estimated and ground-truth z-scores was then measured via MSE. Finally, the average over the MSE scores of all IFNs and visits defined the subject-specific MSE.
4. The group-MSE of the approach was then defined as the average of the 50 subject-specific MSE scores. Low group-MSE scores were preferable.

## 2.3 | Experiments on the NCANDA data

This section first describes the NCANDA cohort (Section 2.3.1) and the preprocessing of the corresponding longitudinal rs-fMRI data (Section 2.3.2). Section 2.3.3 presents the group-level IFNs generated from this dataset. Briefly described in Section 2.3.4 is the experimental setup for parameter exploration of the GIG-ICA-based approaches on the NCANDA dataset, which is now based on simulated rs-fMRI with similar characteristics to the real NCANDA data. The last subsection (Section 2.3.5) reviews the statistical approach for identifying group-level developmental effects on the IFNs extracted by each approach from the NCANDA dataset.

### 2.3.1 | Participants

The NCANDA cohort is composed of 831 subjects that were age from 12 to 21 years at baseline and recruited across five collection sites: University of California at San Diego (UCSD), SRI International, Duke University Medical Center, University of Pittsburgh Medical Center, and Oregon Health & Science University (OHSU). The Institutional Review Boards (IRB) of each site approved this study (Brown et al., 2015). At baseline, 692 adolescents met the minimum drinking criteria of NCANDA, 593 of whom had usable structural and rs-fMRI scans (Müller-Oehring et al., 2018). By Year 3 of the study, 246 youth

**TABLE 1** Demographic characteristics of the 593 NCANDA subjects that were no-to-low drinking adolescents and had usable baseline scans. By Year 3 of the study, 246 subjects remained no-to-low drinkers with three usable longitudinal rs-fMRI scans

	Baseline	Longitudinal
Participants (N)	593	246
Age (years)	15.9 ± 2.2 (12–21.9)	15.3 ± 2.2 (12–21.3)
Sex (male/female)	184/309	117/129
Scanner (Siemens/GE)	200/393	85/161

remained no-to-low drinkers and had usable scans (structural and rs-fMRI) that were acquired annually. The demographics of this cohort relevant to the current analysis (Müller-Oehring et al., 2018) are shown in Table 1. In compliance with NIH policy, the data release NCANDA\_PUBLIC\_Y2\_RESTINGSTATE\_V01 is released to the public according to the NCANDA Data Distribution agreement (see <https://www.niaaa.nih.gov/research/major-initiatives/national-consortium-alcohol-and-neurodevelopment-adolescence>) (Nichols & Pohl, 2015; Rohlfing, Cummins, Henthorn, Chu, & Nichols, 2013).

### 2.3.2 | MRI acquisition and preprocessing

MRI data were acquired using 3T GE Discovery MR750 at three sites (67 longitudinal/155 baseline from University of California-San Diego, 35/109 from SRI International, and 59/129 from Duke University) and 3T Siemens TIM TRIO scanners at 2 sites (28/73 from University of Pittsburgh and 57/127 from Oregon Health & Sciences University). Details on the acquisition protocols of the two scanners are summarized in Table 2. Image quality was ensured with fBIRN phantom scans collected within the week of participant scanning. Additionally, three “human phantoms” were repeatedly scanned at all sites. Subject motion was minimized by following best practices for head fixation such as using a head-strap and providing additional padding of the head under the neck and at the sides. Structural image series were inspected for residual motion.

**TABLE 2** MRI acquisition protocols of the GE/Siemens scanner used in the NCANDA study

	GE (8-channel)				Siemens (12-channel)			
	IR-SPGR	T2	Field map	Rs-fMRI	MPRAGE	T2	Field map	Rs-fMRI
TR	5.912 ms	2,500 ms	1,000 ms	2,200 ms	1900 ms	3,200 ms	700 ms	2,200 ms
TE	1.932 ms	99.646 ms	5/7 ms	30 ms	900 ms	404 ms	4.92/7.38 ms	30 ms
Matrix	256 × 256	512 × 512	128 × 128	64 × 64	256 × 256	512 × 512	96 × 96	64 × 64
FOV	24 cm	24 cm	24 cm	24 cm	24 cm	24 cm	24 cm	24 cm
Vox dimension (mm)	0.9375×	0.4688×	1.875×	3.75×	0.9375×	0.46875×	2.5×	3.75×
	0.9375×	0.4688×	1.875×	3.75×	0.9375×	0.46875×	2.5×	3.75×
	1.2	1.2	2.5	5	1.2	1.2	2.5	5
# of slices	146	146	64	32	160	160	64	32
Acquisition time	7m14s	3m26s	3m24s	10m3s (275 TRs)	8m8s	4m18s	2m16s	10m12s (275 TRs)
Other	TI = 400 ms Flip angle = 11° NEX = 1 ASSET = 1	ETL = 100 Fat sat = on ASSET = 4	Skip = 0 mm	Phase = A/P Flip angle = 79° Freq DIR = R/L Fat sat = on	TI = 900 ms Flip angle = 9° NEX = 1	GRAPPA Acceleration = 2		Phase = A/P Flip angle = 79° Freq DIR = R/L

Resting-state data were preprocessed using the publicly-available NCANDA pipeline (Müller-Oehring et al., 2018), which consisted of motion correction, outlier-detection, detrending, physiological noise removal as well as temporal (low pass frequency: 0.1, high pass frequency: 0.01) and spatial smoothing. The approach was based on the following software packages: FSL V5.06 (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012), Numpy V1.8.2 (<http://www.numpy.org/>), Nipype V0.10 (Gorgolewski et al., 2011), and CMTK V3.2.3 (Rohlfing, Maurer Jr, Bluemke, & Jacobs, 2003). Furthermore, the pipeline labeled frames in individual rs-fMRI image as outliers with displacement >0.3 mm/TR. After outlier detection, rs-fMRI images that passed quality control and with a scan time ≥7.8 min were marked as usable for further analysis (Müller-Oehring et al., 2018). Each rs-fMRI image was registered to the SRI24 atlas (Rohlfing, Zahr, Sullivan, & Pfefferbaum, 2010). While some recent studies suggest to directly register the mean BOLD image to the atlas (Adhikara, Jahanshad, Shukla, 2017; Calhoun et al., 2017; Dohmatob et al., 2018), the processing here complied with the baseline publication of NCANDA (Müller-Oehring et al., 2018), which first aligned the mean BOLD image to the subject-specific T1-weighted MRI and then nonrigidly registered T1-weighted MRI to the SRI atlas. Visual inspection insured that the quality of the registration between the mean BOLD image and the atlas space was high.

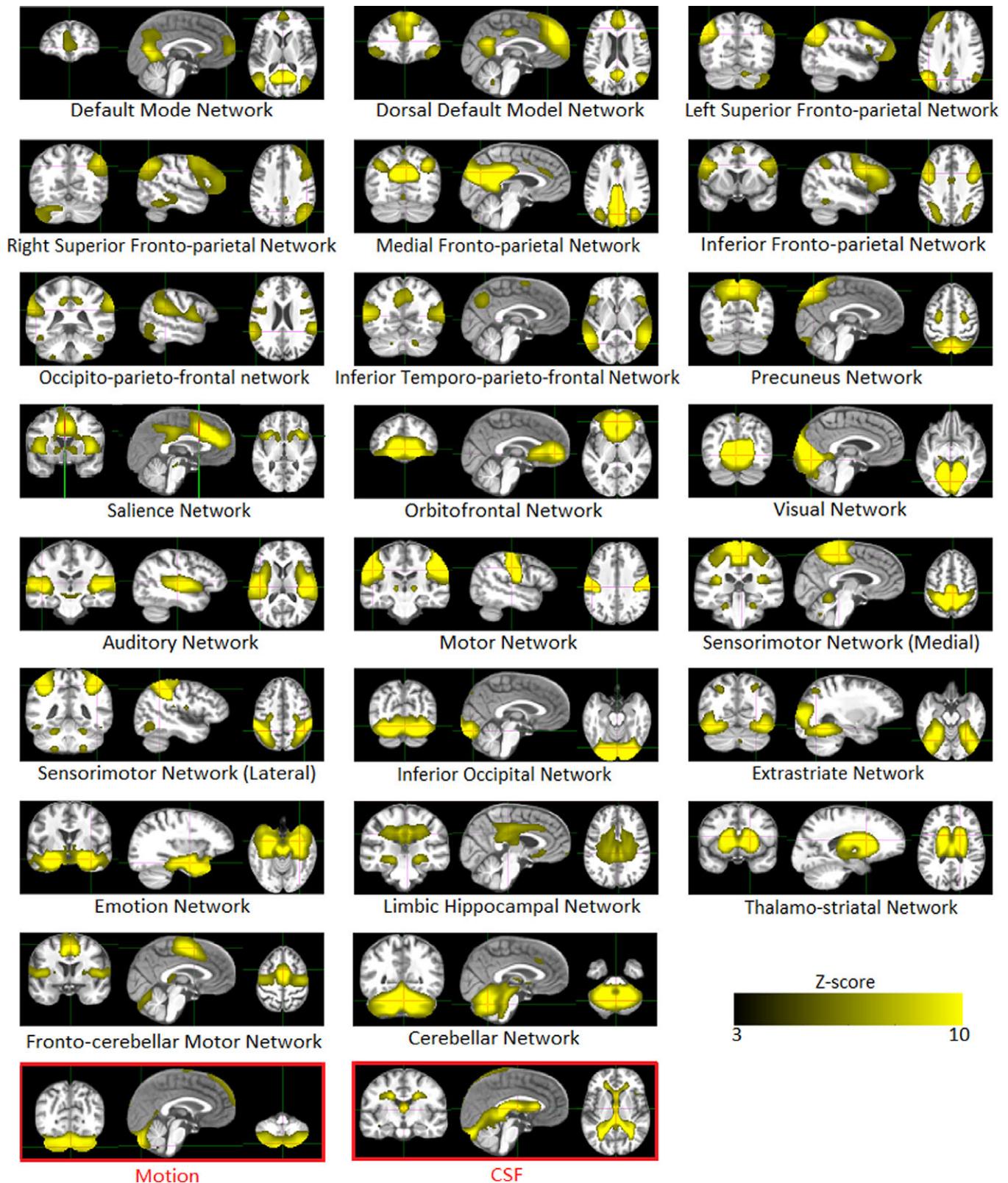
### 2.3.3 | Group-level IFNs of the NCANDA dataset

As suggest by Damoiseaux, Prater, Miller, and Greicius (2012); Dipasquale et al. (2015); Odish et al. (2015), 25 high-quality, group-level independent components were generated by applying Probabilistic ICA of MELODIC FSL V5.0.10 (Beckmann et al., 2009) to all 593 preprocessed rs-fMRI images at baseline. Here, computing relatively few group-level independent components (typically <30) is known as the low-dimensional approach (Dipasquale et al. (2015), in which each component may consist of several spatially distinct regions and represents a functional network on a whole-brain scale. This is contrary to the high-dimensional approach (>70 components), which produces single-region components and therefore is often used for brain parcellation rather than within-component analysis, the focus of this study.

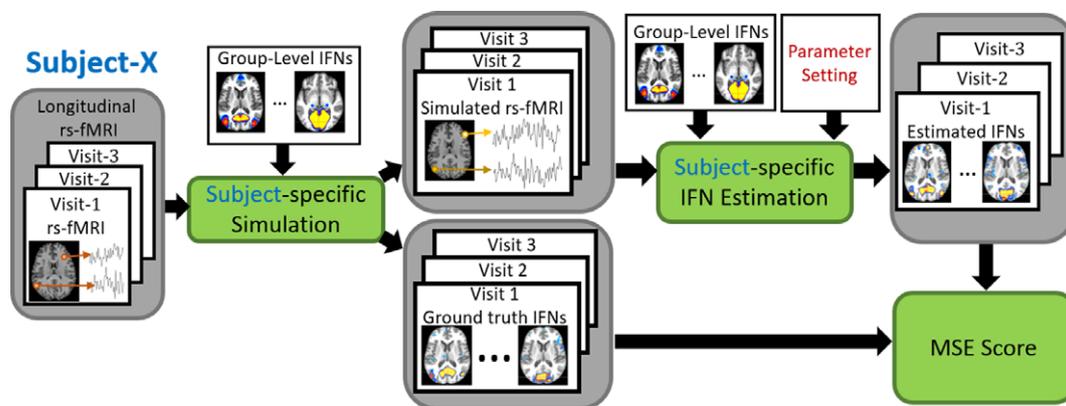
Finally, expert review (E.M., A.L.) identified 23 independent components out of the 25 as true IFNs and labeled 2 independent components as artifacts related to cerebrospinal fluid and motion (Griffanti et al., 2017; Figure 3).

### 2.3.4 | Parameter exploration of GIG-ICA-based approaches

Parameter exploration of GIG-ICA, VL-ICA, and RL-ICA relied on measuring group-MSE on data generated by a novel simulation. The



**FIGURE 3** The 23 group-level IFNs and the two artifacts (motion and CSF) overlaid with the SRI24 atlas (Rohlfing et al., 2010). The IFNs were generated from the baseline rs-fMRIs of 593 minimal drinkers of the NCANDA cohort. For each network, only regions with high connectivity (z-scores  $\geq 3$ ), also called hubs, are visualized [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** The simulation produced semi-realistic data for measuring the accuracy of an approach (e.g., VL-ICA) with a specific parameter [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

simulation transformed each of the 246 longitudinal rs-fMRI of the NCANDA dataset into a synthetic, realistic-looking one with ground-truth IFNs (Figure 4). Then for each approach and setting, the group-MSE on the synthetic dataset was computed as in the SimTB-based experiments with the following exceptions: time course dimension was reduced to 80 (suggested to be between 50 and 100 (Du & Fan, 2013)), which retained in average over 95% of the energy of the PCA model; the 100 ROIs determined by the whole-brain functional parcellation tool *pyClusterROI* (Craddock, James, Holtzheimer, Hu, & Mayberg, 2012) defined the parcellation for RL-ICA. For readers interested in technical details, the remainder of this section describes the novel simulation in further detail.

For each of the 246 adolescents, the simulation applied dual regression (Beckmann et al., 2009) to the longitudinal rs-fMRI. For the rs-fMRI of each visit, dual regression produced 25 independent components with corresponding time courses and a residual rs-fMRI image, which was computed by subtracting the independent-components-induced image from the raw rs-fMRI. Corresponding to the group-level IFNs, the 23 subject-specific IFNs at baseline were viewed as ground-truth. For each of these IFNs, the simulator then computed the ground-truth for the remaining two visits: (a) a linear slope was computed at each voxel across the three subject-specific IFNs estimated by dual regression; (b) AFNI 3dClustSim V17.3.09 (Forman et al., 1995) estimated the mixed model auto-correlation function (the “heavy-tail ACF” (Cox, Chen, Glen, Reynolds, & Taylor, 2017; Eklund, Nichols, & Knutsson, 2016)) of the resulting slope map to generate a random field with zero mean (i.e., no overall development effects across subjects); (c) this random field was viewed as the ground-truth slope, which was successively added to the baseline IFN to generate the ground-truth IFNs for the remaining two visits.

For each subject and visit, a synthetic rs-fMRI image was generated by remixing the 25 independent components, that is, 23 the ground-truth IFNs and 2 artifacts, with their corresponding time courses. The 2 artifact-related components were re-introduced to the simulation to create realistic looking rs-fMRI. Afterwards, another random-field was generated by AFNI 3dClustSim based on the auto-correlation function estimated from the residual rs-fMRI image. This random field represented the additional noise not captured by the

independent components and was added to the current synthetic rs-fMRI.

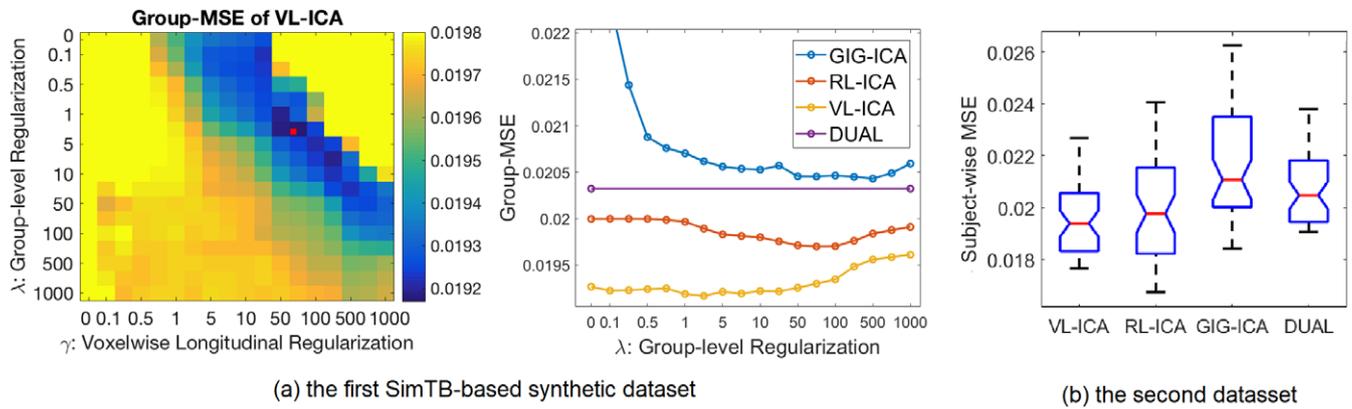
By construction, all simulated rs-fMRI closely resembled real NCANDA data with the underlying subject-specific developmental effects being randomly sampled from a distribution with zero mean. Thus, optimal parameter exploration on this dataset only maximized for the accuracy of IFN estimation while being impartial toward any group-level effects supporting the hypothesis of the NCANDA study.

### 2.3.5 | Detecting group-level developmental effects of IFNs

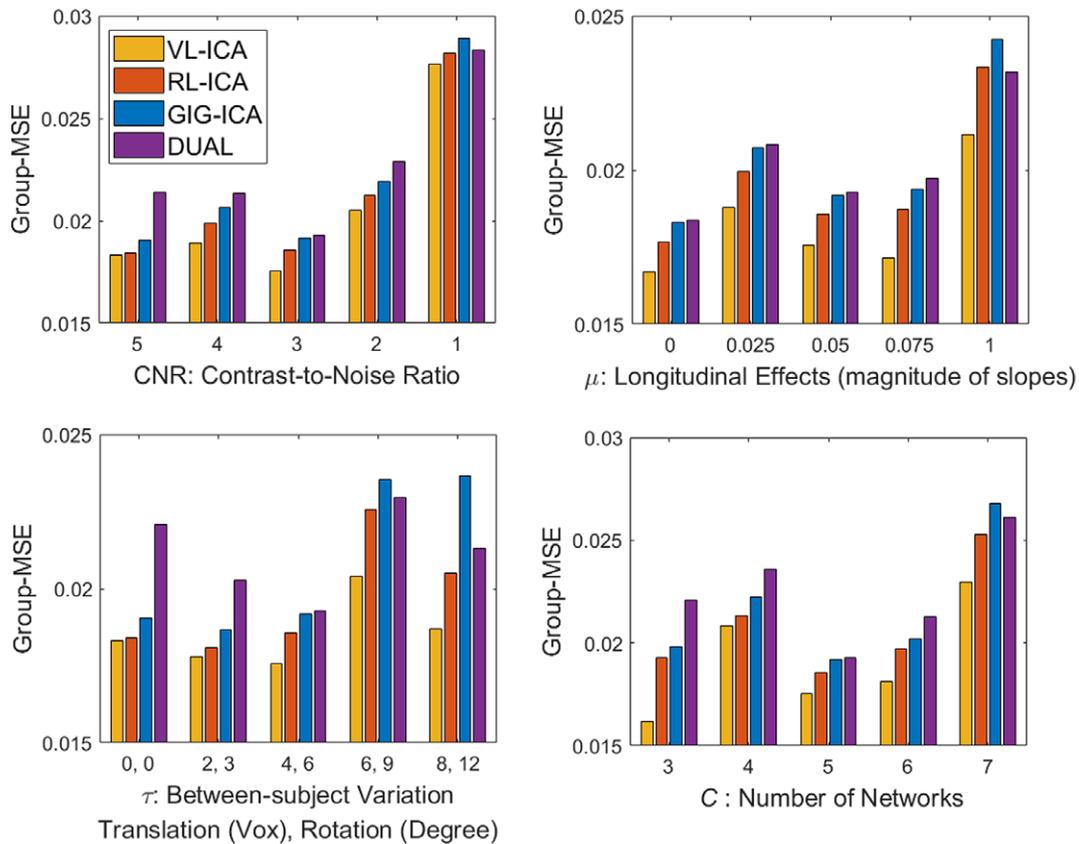
Each approach (DUAL, GIG-ICA, RL-ICA, and VL-ICA) with its optimal setting was applied to the 246 real longitudinal rs-fMRI of NCANDA. For each resulting subject-specific IFN at baseline, a slope map (or longitudinal trajectory map) was generated by computing at each voxel the slope of the IFN estimates across the three visits. For each IFN and approach, group-level developmental effects were then tested via a whole-brain voxel-wise permutation test on all 246 subject-specific slope maps associated with that IFN. Voxels with significantly positive slopes indicated increased functional connectivity within that IFN, and vice versa. Appendix C describes in further detail the test procedure, which includes removing confounding demographics, the nonparametric permutation test and multiple comparison correction.

Any cluster of voxels that reached the significance level was further validated by linear mixed effect (LME) analysis and motion analysis. LME analysis examined whether a cluster detected by one specific approach would also exhibit similar group-level developmental effects when the IFN estimates of other approaches were used. Based on the average connectivity inside the cluster, the LME model considered a fixed effect of linear aging as the group-level developmental effect and a random effect of intercept.

The resulting fixed effect was considered significant if one-sided  $p \leq .001$ . With respect to motion analysis, each detected cluster was analyzed by three different analyses, which examined whether the detected group-level developmental effects were confounded by motion artifacts (see Appendix C for details).



**FIGURE 5** (a) Group-MSEs recorded on the first SimTB-based synthetic dataset based on different regularization (parameter) settings (left: VL-ICA; right: DUAL, GIG-ICA, RL-ICA, and VL-ICA). Note, DUAL is independent of  $\lambda$ , and the 1D curve of VL-ICA was derived from the 2D plot on the left by determining the lowest group-MSE recorded for each setting of  $\lambda$ . (b) Based on the parameter settings determined in the first SimTB dataset, the 50 subject-wise MSEs of the four approaches recorded on the second SimTB-based synthetic dataset [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 6** The optimal group-MSEs achieved across 20 SimTB-based synthetic datasets with different levels of noise, longitudinal effects, between-subject IFN variation and number of networks [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3 | RESULTS

#### 3.1 | Synthetic data highlight robustness of longitudinal ICA

With respect to the first SimTB-based synthetic dataset, Figure 5a shows the group-MSE scores associated with the parameter exploration of different approaches. For all settings, the two variants of longitudinal ICA (RL-ICA and VL-ICA) achieved lower group-MSEs than the

two cross-sectional approaches. The optimal setting of VL-ICA (red dots in Figure 5a) led to a group-MSE that was 2.2% lower than the minimum group-MSE of RL-ICA, 5.6% lower than GIG-ICA, and 6.1% lower than DUAL. The differences in group-MSE between cross-sectional and longitudinal approaches were statistically significant ( $p < 0.001$  paired two-sample  $t$ -tests on subject-wise MSEs).

Fixed to their optimal settings, the approaches were then applied to the second SimTB dataset (without further parameter exploration) to investigate the repeatability of the findings. VL-ICA and RL-ICA

achieved significant lower MSE scores compared to the cross-sectional approaches ( $p < 0.001$  paired two-sample  $t$ -tests, Figure 5b). This indicated that optimal parameter settings of longitudinal ICA can be used across datasets with similar characteristics as the first two datasets shared the same simulation parameter (see Appendix B).

In the last set of experiments, parameter exploration of each approach was repeated on each of the remaining 20 SimTB-based synthetic datasets. These datasets differed from the previous two datasets in terms of noise level, between-subject IFN variation, level of longitudinal effects, and number of networks per subject (see Appendix B). The resulting optimal group-MSE of each approach for each dataset is given in Figure 6.

Based on these plots, GIG-ICA generally achieved lower group-MSE (performed better) than DUAL. Similar to the previous experiments, these two cross-sectional approaches mostly performed worse than the two longitudinal ICA implementations, and VL-ICA always achieved the best MSE scores. This result also held for the dataset with no longitudinal effects ( $\mu = 0$ ).

Unlike the first experiment, the optimal setting of each approach now varied across datasets (Tables 3–6). Furthermore, when applying the optimal parameter setting for all 20 datasets (1,000 longitudinal rs-fMRIs) to each individual dataset, the group-MSE score increased up to 7.1% for VL-ICA and 6.1% for RL-ICA. While a “universal” setting for  $\lambda$  was suggested by Du et al. (Du &

Fan, 2013), our findings on datasets with different characteristics suggested that the accuracy in estimating subject-specific IFN improved when GIG-ICA and our longitudinal approaches were run using “dataset-specific” parameter settings.

### 3.2 | Parameter exploration of GIG-ICA-based approaches on the NCANDA-based simulated data

To determine the optimal parameter settings for the NCANDA dataset, the three GIG-ICA-based approaches were first applied to the 246 simulated longitudinal rs-fMRIs (Figure 4). The initial parameter exploration in  $\lambda \in [0, 1000]$ ,  $\gamma \in [0, 1000]$  suggested that the optimal value for  $\lambda$  was close to 0. To find the optimal setting, the final exploration of  $\lambda$  was performed on a finer scale in the interval  $[0, 2]$ . As shown in example of Figure 7, the optimal estimates by GIG-ICA were strongly affected by the noise and artifacts at each visit and contained large variation across visits. Controlling for intra-subject dependencies, RL-ICA and VL-ICA yielded longitudinally consistent estimates that accurately reflected the ground-truth developmental effects across visits. This qualitative assessment was further supported by the plots of Figure 8, which showed that the group-MSE scores of the longitudinal approaches were significantly lower than GIG-ICA ( $p < 0.001$  paired two-sample  $t$ -tests). Furthermore, VL-ICA achieved significantly lower group-MSE scores than RL-ICA and GIG-ICA

**TABLE 3** Optimal settings for SimTB experiments with respect to contrast-to-noise ratio (CNR)

$\mu = 0.05, \tau_1 = 4, \tau_2 = 6, C = 5$					
CNR	5	4	3	2	1
GIG-ICA ( $\lambda$ )	100	100	500	200	500
RL-ICA ( $\lambda$ )	50	75	100	100	50
VL-ICA ( $\lambda, \gamma$ )	0.2, 5	0.5, 7.5	2, 50	10, 100	10, 100

**TABLE 4** Optimal settings for SimTB experiments with respect to level of longitudinal effects ( $\mu$ )

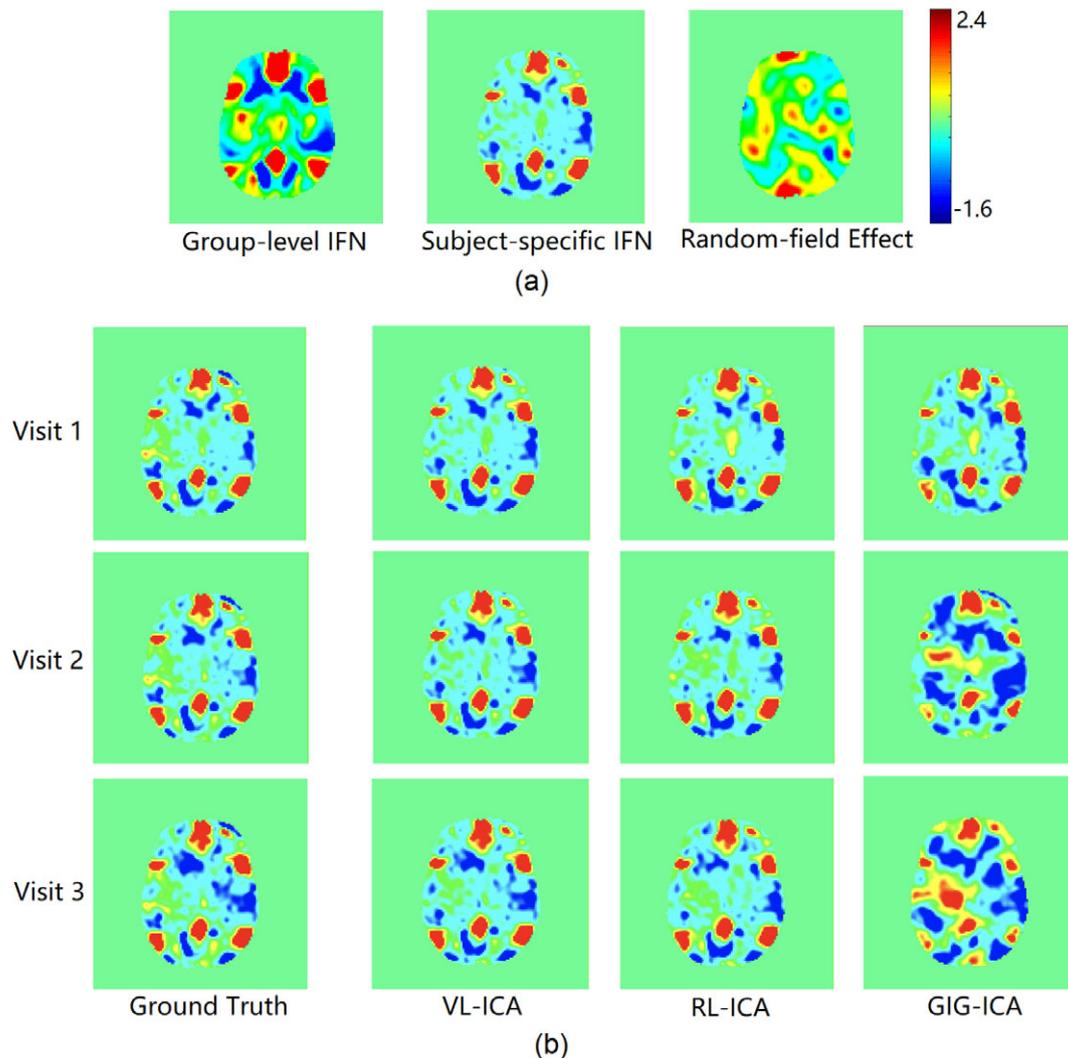
CNR = 3, $\tau_1 = 4, \tau_2 = 6, C = 5$					
$\mu$	0	0.025	0.05	0.075	0.1
GIG-ICA ( $\lambda$ )	20	20	200	100	200
RL-ICA ( $\lambda$ )	100	75	100	500	200
VL-ICA ( $\lambda, \gamma$ )	1, 5	1, 10	2, 50	20, 100	20, 500

**TABLE 5** Optimal settings for SimTB experiments with respect to between-subject variation ( $\tau$ )

CNR = 3, $\mu = 0.05, C = 5$					
Translation $\tau_1$ (vox)	0	2	4	6	8
Rotation $\tau_2$ (degree)	0	3	6	9	12
GIG-ICA ( $\lambda$ )	500	200	200	75	50
RL-ICA ( $\lambda$ )	500	100	100	20	5
VL-ICA ( $\lambda, \gamma$ )	20,500	20, 200	2, 50	0.5,10	0.1,1

**TABLE 6** Optimal settings for SimTB experiments with respect to number of networks (C)

CNR = 3, $\mu = 0.05, \tau_1 = 4, \tau_2 = 6$					
C	3	4	5	6	7
GIG-ICA ( $\lambda$ )	500	500	200	200	100
RL-ICA ( $\lambda$ )	200	200	100	20	20
VL-ICA ( $\gamma, \lambda$ )	5, 100	1, 100	2, 50	0.75, 20	1, 20



**FIGURE 7** (a) Left: The spatial map of the group-level estimation of the dorsal default mode network; Middle: A simulated subject-specific dorsal default mode network at baseline; Right: A random-field-based longitudinal effect, which was successively added to the baseline network to simulate networks for the following two visits. (b) Left: Simulated ground truth of dorsal default mode network at the three visits of a longitudinal rs-fMRI. Right: The three estimated networks derived by VL-ICA, RL-ICA, and GIG-ICA, using their respective optimal parameters as indicated in Figure 8 [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

( $p < 0.001$  paired two-sample  $t$ -tests). Note that for both NCANDA-based and SimTB-based experiments, the MSE scores of GIG-ICA listed here were based on our implementation (Appendix A), and they were slightly better (i.e., lower) than those based on the implementation suggested by Du and Fan (2013).

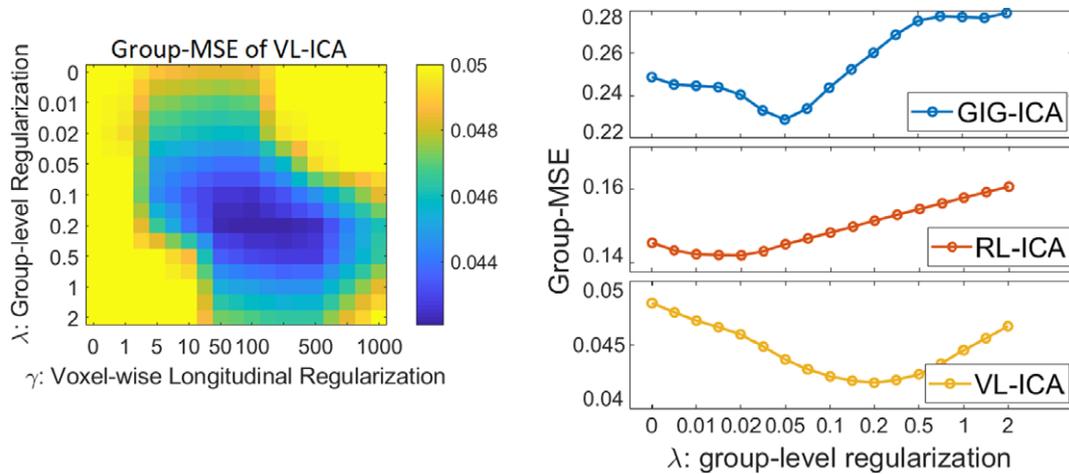
In addition, the different search ranges of  $\lambda$  between the SimTB-based and NCANDA-based datasets further highlighted the need for data-driven parameterization of GIG-ICA-based approaches. Since all simulated data in this experiment were “semi-realistic” with similar characteristics to the real NCANDA data, optimal parameter settings for this dataset would translate to the real NCANDA dataset for analyzing developmental effects of IFNs.

### 3.3 | Group-analysis of slopes identifies developmental effects in the NCANDA data

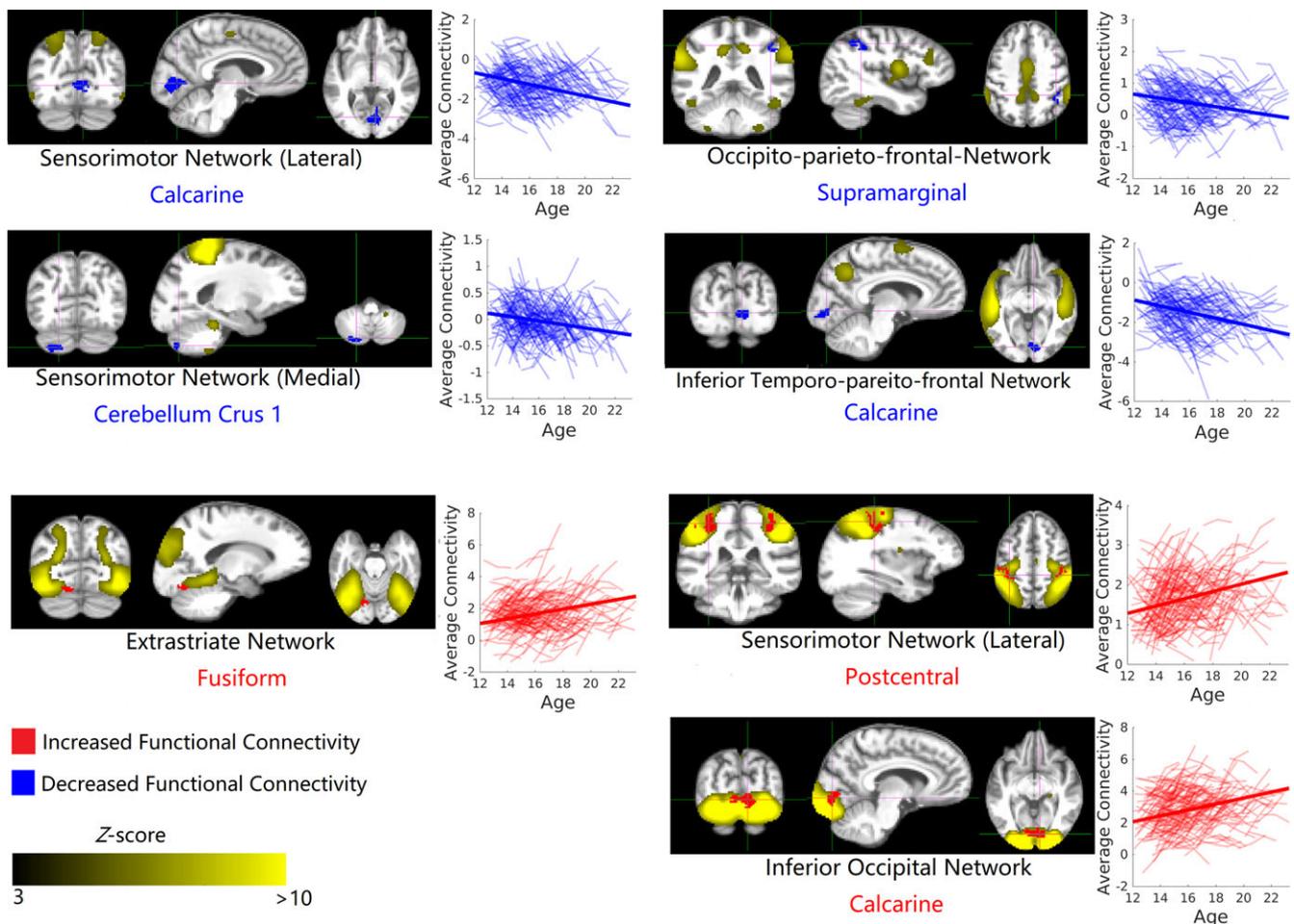
Group-level developmental effects of IFNs on the NCANDA data were examined based on subject-specific IFNs computed by each of the four approaches using their optimal setting. Among the statistical

group-analyses, only the analysis based on VL-ICA-derived slope maps detected significant developmental effects. These effects were confined to 6 IFNs (see Figure 9 and Table 7). Specifically, increased functional connectivity was revealed within the fusiform region of the extrastriate network, the calcarine region of the inferior occipital network, and the postcentral region of the lateral sensorimotor network. Decreased functional connectivity was found within the cerebellum crus 1 region of the medial sensorimotor network, the supramarginal region of the occipito-parieto-frontal network, and the calcarine region of the lateral sensorimotor network and of the inferior temporo-parietal-frontal network. Regions were defined according to the SRI24 atlas (Rohlfing, Zahr, Sullivan, & Pfefferbaum, 2010).

Moreover, when performing LME analysis confined to the seven detected clusters, similar group-level effects could also be observed in the IFNs produced by all other approaches (RL-ICA, GIG-ICA, and DUAL) despite the fact that they were not detectable in the whole-brain voxel-wise analysis (Table 8). This agreement among the four approaches confirmed the reliability of the findings and again



**FIGURE 8** Group-MSE scores of the three GIG-ICA-based approaches on the 246 simulated longitudinal rs-fMRIs from NCANDA with different parameter settings. The optimal parameter setting of each approach was used for estimating subject-specific IFNs on real NCANDA data [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 9** Based on VL-ICA estimates, linear developmental effects were found in 6 IFNs. Yellow regions visualize the hub (high-connectivity) regions of the networks. Clusters in red (with the anatomical names of the corresponding regions printed in red) consist of voxels, whose slopes in functional connectivity across visits are significantly positive. Correspondingly, clusters in blue (with anatomical names printed in blue) consist of voxels with “significantly negative” slopes. For each cluster, the average functional connectivity after sex and scanner correction was further examined by linear mixed effects (LMEs) analysis. Bold lines correspond to the fixed aging (group-level) effects. Notice, all red regions reside inside the hubs whereas all blue regions reside outside the hubs, which suggests both within-network cohesiveness and between-network segregation increase throughout adolescence [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 7** List of clusters (cluster size  $\geq 50$ ) identified by voxel-wise one-sample t-tests ( $p < 0.05/23 = 0.0022$ , TFCE-corrected) on the 246 slope maps. The slope maps were determined from subject-specific IFN estimates computed by applying VL-ICA to the 246 longitudinal rs-fMRIs of the NCANDA cohort. In each cluster, the minimum voxel-wise  $p$ -value is displayed in the right column. Region names were specified according to the SRI24 atlas (Rohlfing et al., 2010)

Effects	Network (region)	Size (voxel)	Min $p$ -value
Linear increasing	Lateral sensorimotor (postcentral)	557	0.0002
	Inferior occipital (Calcarine)	248	0.0004
	Extrastriate (fusiform)	53	0.001
Linear decreasing	Lateral sensorimotor (Calcarine)	283	0.0002
	Medial sensorimotor (cerebellum Crus1)	83	0.0002
	Inferior Temporo-parietal-frontal (Calcarine)	197	0.0004
	Occipito-parieto-frontal (Supramarginal)	89	0.0004

demonstrated that VL-ICA was the most accurate approach with respect to the NCANDA dataset.

To ensure the specificity of the detected clusters, the above experiment of detecting group-level developmental effects was also repeated on the 246 simulated longitudinal rs-fMRIs (generated by the simulation for parameter exploration; Figure 4). As expected, none of the group analyses on this NCANDA-based simulated dataset revealed any significant results as the underlying ground-truth subject-specific slopes were random fields simulated with zero mean. This further showed the parameter exploration of longitudinal ICA improved IFN estimation but was unlikely to yield false-positive group-level findings. Finally, the motion analysis (Appendix C) indicated that the detected developmental effects were not confounded by motion artifacts.

## 4 | DISCUSSION

The results support the central hypothesis that longitudinal ICA determined more precise estimates of subject-specific IFNs than estimates generated separately for each visit. In both synthetic-data experiments (SimTB- and NCANDA-based), the two implementations of longitudinal ICA were more accurate in estimating subject-specific IFNs across visits than cross-sectional approaches. On the NCANDA

dataset, only the implementation with voxel-level longitudinal modeling (i.e., VL-ICA) was able to generate subject-specific IFNs that revealed significant group-level developmental effects. These effects in the corresponding six IFNs were impartial to motion and also observed via post hoc LME analysis based on the IFN-estimates of the other cross-sectional and longitudinal approaches. Note, RL-ICA might have also discovered significant developmental effects using a different parcellation, which again would have supported our central hypothesis.

In the present study, all findings were based on the linear modeling of within-subject developmental effects, that is, slopes. This assumption was reasonable for the NCANDA study as the longitudinal rs-fMRI of each adolescent covered a relatively small time span (e.g., 2 years; Vetter et al., 2017). Based on this assumption, longitudinal ICA can help in disentangling biologically plausible developmental effects from complex artifacts (such as physiology noise and motion) as these artifacts do not adhere to the linear intra-subject constraint. The longitudinal constraint substantially improved IFN estimates and was essential for the detection of group-level effects. Our study thus further supports that errors in cross-sectional estimates can overpower the developmental effects to be detected by post hoc statistical analysis methods (e.g., LME; Gao et al., 2015; Stoffers et al., 2015). Even when there was no developmental effect, the constraint

**TABLE 8**  $p$ -values of the fixed (group-level) effect in the cluster-wise LME analysis. The LME was repeatedly performed inside the VL-ICA-derived clusters based on the IFNs estimated by different approaches (\*denotes  $p < 0.001$ , **bold** corresponds to  $p > 0.05$ )

Effects	Network (region)	VL-ICA	RL-ICA	GIG-ICA	DUAL
Linear increasing	Lateral sensorimotor (postcentral)	*	*	*	*
	Inferior occipital (Calcarine)	*	0.0016	*	0.0023
	Extrastriate (fusiform)	*	0.0012	*	0.001
Linear decreasing	Lateral sensorimotor (Calcarine)	*	0.0038	*	0.042
	Medial sensorimotor (cerebellum Crus1)	*	<b>0.15</b>	0.045	0.042
	Inferior Temporo-parietal-frontal (Calcarine)	*	*	*	<b>0.068</b>
	Occipito-parieto-frontal (Supramarginal)	*	0.033	0.007	<b>0.082</b>

improved estimation as indicated by the findings on the SimTB-based experiments with  $\mu = 0$ . In this case the constraint essentially penalized the difference among IFNs across visits.

In principle, modeling linear within-subject developmental effects does not prevent from examining nonlinear group-level effects, a topic tangential to the present study.

Moreover, longitudinal ICA itself can be readily extended to examine more complex within-subject effects by regularizing consistency with higher-order longitudinal constraints. However, robustly fitting high-order models would require longitudinal rs-fMRI with more than three time points, which currently limits analysis to a few longitudinal studies. As for the NCANDA study, the ongoing collection of longitudinal rs-fMRI data of up to five visits would then permit retesting the detected developmental effects by applying both linear and nonlinear models with more than three scans for each subject.

As presented here, longitudinal ICA requires longitudinal rs-fMRI of at least three visits for each subject because IFNs of only two visits naturally satisfy the linear constraint. A potential extension of the method to datasets with two-visits data is to incorporate a random effects model. The random effects model would then regularize individual developmental effects between two visits to the average and variance of the developmental effects in a population. In general, such statistical models are inefficient due to the iterative estimation of population-level and subject-specific effects (Wang & Guo, 2018) and thus prohibitive for large neuroimaging datasets.

Another insight from the experiments is that the impact of longitudinal constraints on IFN estimation had to be carefully balanced with that of the group-level IFNs. This balance was determined through parameter exploration based on synthetic data, a popular data-driven approach (Du & Fan, 2013). The synthetic experiments also revealed that the outcome of parameter exploration only generalized well (i.e., preserve the accuracy of cross-sectional and longitudinal approaches) to datasets with similar characteristics (Figure 5b). This observation motivated the creation of a novel simulator that generated semi-realistic data closely resembling the target rs-fMRI to be analyzed. Our simulation was based on dual regression, which was well recognized for generating accurate cross-sectional IFN estimates but lacked longitudinal modeling. This motivated us to use dual regression only for simulating ground-truth IFNs of the first visit whereas longitudinal effects were simulated using the mixed model random-field theory. This novel simulation required little a priori knowledge to run and was fully data-driven.

Being data-driven was essential to the presented analysis as it identified developmental effects within NCANDA's longitudinal rs-fMRI without expert-domain knowledge.

Specifically, the data-driven analysis identified 23 meaningful group-specific IFNs based on the baseline rs-fMRI of 593 participants of the NCANDA study. The 23 IFNs comprised distinct networks tagging self-referential (default mode network, dorsal default mode network; Greicius & Menon, 2004) and integrative functions (left and right superior, medial, and inferior fronto-parietal networks, occipito-parieto-frontal, inferior temporo-parieto-frontal, precuneus, anterior insular-cingulate salience, and orbitofrontal networks; Chahine, Richter, Wolter, Goya-Maldonado, & Gruber, 2017; Gaudio et al., 2015). Also reflected were primary input (visual, auditory, motor, somato-

sensori networks; Braga, Wilson, Sharp, Wise, & Leech, 2013) and secondary visual association functions (inferior occipital, extrastriate networks; Pamilo et al., 2012). The remaining networks were associated with subcortical emotion, memory, reward (emotion, limbic hippocampal, and thalamo-striatal networks; Killgore et al., 2017), and integrative cerebellar functions (cerebellar and fronto-cerebellar motor networks; Dobromyslin et al., 2012). These IFNs were also identified when the group-level ICA was performed only on the 246 adolescents' baseline scans, but the resulting IFNs were (spatially) less smooth. This motivated us to use all available 593 baseline scans resulting in more robust group-level IFN estimates.

Of those 23 IFNs, 6 IFNs contained linear developmental effects: the secondary visual association (inferior occipital, extrastriate) and primary lateral somatosensori networks exhibited an "in-network" strengthening of functional connectivity with older adolescent age, while at the same time lateral and medial somatosensori, occipito-parieto-frontal, and inferior temporo-parieto-frontal network showed decreased "out-network" functional connectivity (Figure 9). The pattern of these effects complied with recent findings in the maturation of functional organization during adolescence. All detected regions with increased connectivity resided inside the hubs, that is, regions with high connectivity ( $z$ -scores  $>3$ ) within an IFN. This pattern was also observed in the default mode network but failed to meet the stringent cluster-size threshold of this study. The strengthening of hub regions suggests that the IFNs become more cohesive and specialized toward specific cognitive functioning during adolescence. This pattern was previously reported in the default mode network (Fair et al., 2008; Sherman et al., 2014) and in the seed-based analysis of the NCANDA baseline data with respect to the saliency network and the executive control network (Müller-Oehring et al., 2018). This developmental pattern was also frequently suggested to exist during childhood (Betzel et al., 2014; Cao, Huang, Peng, Dong, & He, 2016; Cao et al., 2014). Regions with decreased functional connectivity resided outside the IFN's key hubs ( $3 \geq z > 0$ ) suggesting that an IFN becomes more segregated from others to enable more efficient, faster, and functionally specific processing with maturation. The increased segregation during adolescence was previously suggested from both longitudinal (Sherman et al., 2014) and cross-sectional studies (Gu et al., 2015) in the executive control network (Sherman et al., 2014), sensorimotor systems and higher order cognitive systems (saliency, memory, and attention systems; Gu et al., 2015).

## 5 | CONCLUSION

In summary, we have shown that longitudinal ICA was able to produce more accurate subject-specific IFN estimates than cross-sectional approaches. This finding was supported by the results on the SimTB-based synthetic datasets, which measured the accuracy of each approach with respect to various data characteristics. Finally, only the statistical group analysis on the IFNs estimated by a longitudinal ICA implementation identified significant developmental effects within the NCANDA cohort. When confining analysis to regions exhibiting significant effects, the findings were confirmed by the cross-sectional approaches. The findings also agreed with existing adolescent

literature, which mostly has reported on cross-sectional rs-fMRIs analysis using seed-based (Fair et al., 2008; Sherman et al., 2014) or ROI-based (Betzel et al., 2014; Cao et al., 2014; Gu et al., 2015) approaches.

## AUTHOR CONTRIBUTIONS

Q.Z. developed methods, designed the studies, performed analysis, and wrote the manuscript. D.J.K. contributed to image preprocessing and the study design. E.M.M.O, A.P.L.B., A.P., and E.V.S interpreted the results and helped in writing. K.M.P contributed to the methodological development, analyzing results, and writing the manuscript.

## ORCID

Qingyu Zhao  <https://orcid.org/0000-0002-6368-0889>

Edith V. Sullivan  <https://orcid.org/0000-0001-6739-3716>

## REFERENCES

- Adhikara, B. M., Jahanshad, N., Shukla, D., et al. (2017). Heritability estimates on resting-state fMRI data using ENIGMA analysis pipeline. *Pacific Symposium on Biocomputing*, 23, 307–318.
- Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., et al. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data*, 4, 170181. <https://doi.org/10.1038/sdata.2017.181>
- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3), 663–676. <https://doi.org/10.1093/cercor/bhs352>
- Beckmann, C. F., Mackay, C. E., Filippini, N., & Smith, S. M. (2009). Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. *NeuroImage*, 47, S148. [https://doi.org/10.1016/S1053-8119\(09\)71511-3](https://doi.org/10.1016/S1053-8119(09)71511-3)
- Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X. N., & Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage*, 102, 345–357. <https://doi.org/10.1016/j.neuroimage.2014.07.067>
- Biswal, B. B. (2012). Resting state fMRI: A personal history. *NeuroImage*, 62(2), 938–944. <https://doi.org/10.1016/j.neuroimage.2012.01.090>
- Braga, R. M., Wilson, L. R., Sharp, D. J., Wise, R. J., & Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. *NeuroImage*, 74(100), 77–86. <https://doi.org/10.1016/j.neuroimage.2013.02.023>
- Bright, M. G., & Murphy, K. (2013). Removing motion and physiological artifacts from intrinsic bold fluctuations using short echo data. *NeuroImage*, 64, 526–537. <https://doi.org/10.1016/j.neuroimage.2012.09.043>
- Bright, M. G., & Murphy, K. (2015). Is fMRI “noise” really noise? Resting state nuisance regressors remove variance with network structure. *NeuroImage*, 114, 158–169. <https://doi.org/10.1016/j.neuroimage.2015.03.070>
- Brown, S. A., Brumback, T., Tomlinson, K., Cummins, K., Thompson, W. K., Nagel, B. J., et al. (2015). The national consortium on alcohol and neurodevelopment in adolescence (NCANDA): A multisite study of adolescent development and substance use. *Journal of Studies on Alcohol and Drugs*, 76(6), 895–908. doi:10.15288/jsad.2015.76.895
- Buckner, R. L., Krienen, F. M., & Yeo, B. T. (2013). Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience*, 16(7), 832–837. <https://doi.org/10.1038/nn.3423>
- Buxton, R., Uludağ, K., Dubowitz, D., & Liu, T. (2004). Modeling the haemodynamic response to brain activation. *NeuroImage*, 23, S220–S233. <https://doi.org/10.1016/j.neuroimage.2004.07.013>
- Calhoun, V. D., Adali, T., McGinty, V. B., Pekar, J. J., Watson, T. D., & Pearlson, G. D. (2001). fMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis. *NeuroImage*, 14(5), 1080–1088. <https://doi.org/10.1006/nimg.2001.0921>
- Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001a). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3), 140–151. <https://doi.org/10.1002/hbm.1048>
- Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001b). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping*, 13(1), 43–53. <https://doi.org/10.1002/hbm.1024>
- Calhoun, V. D., Wager, T. D., Krishnan, A., Rosch, K. S., Seymour, K. E., Nebel, M. B., ... Kiehl, K. (2017). The impact of the T1 versus EPI spatial normalization templates for fMRI data analyses. *Human Brain Mapping*, 38(11), 5331–5342. <https://doi.org/10.1002/hbm.23737>
- Cao, M., Huang, H., Peng, Y., Dong, Q., & He, Y. (2016). Toward developmental connectomics of the human brain. *Frontiers in Neuroanatomy*, 10(25). <https://doi.org/10.3389/fnana.2016.00025>
- Cao, M., Wang, J., Dai, Z., Cao, X., Jiang, L., Fan, F., et al. (2014). Topological organization of the human brain functional connectome across the lifespan. *Developmental Cognitive Neuroscience*, 7, 76–93. <https://doi.org/10.1016/j.dcn.2013.11.004>
- Casey, B., Jones, R. M., & Hareb, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, 1124(1), 111–126. <https://doi.org/10.1196/annals.1440.010>
- Chahine, G., Richter, A., Wolter, S., Goya-Maldonado, R., & Gruber, O. (2017). Disruptions in the left frontoparietal network underlie resting state endophenotypic markers in schizophrenia. *Human Brain Mapping*, 38(4), 1741–1750. <https://doi.org/10.1002/hbm.23477>
- Chase, A. (2014). Alzheimer disease: Altered functional connectivity in pre-clinical dementia. *Nature Reviews. Neurology*, 10, 609. <https://doi.org/10.1038/nrnneurol.2014.195>
- Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). Fmri clustering in afni: False-positive rates redux. *Brain Connectivity*, 7(3), 152–171. <https://doi.org/10.1089/brain.2016.0475>
- Craddock, R. C., James, G. A., Holtzheimer, P. E. r., Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8), 1914–1928. <https://doi.org/10.1002/hbm.21333>
- Damaraju, E., Caprihan, A., Lowe, J. R., Allen, E. A., Calhoun, V. D., & Phillips, J. P. (2014). Functional connectivity in the developing brain: A longitudinal study from 4 to 9 months of age. *NeuroImage*, 84, 169–180. <https://doi.org/10.1016/j.neuroimage.2013.08.038>
- Damoiseaux, J. S., Prater, K. E., Miller, B. L., & Greicius, M. D. (2012). Functional connectivity tracks clinical deterioration in Alzheimer's disease. *Neurobiology of Aging*, 33(4), 828. <https://doi.org/10.1089/brain.2012.0073>
- Dipasquale, O., Griffanti, L., Clerici, M., Nemni, R., Baselli, G., & Baglio, F. (2015). High-dimensional ICA analysis detects within-network functional connectivity damage of default-mode and sensory-motor networks in Alzheimer's disease. *Frontiers in Human Neuroscience*, 9(43), 1–7. <https://doi.org/10.3389/fnhum.2015.00043>
- Dobromylin, V. I., Salat, D. H., Fortier, C. B., Leritz, E. C., Beckmann, C. F., Milberg, W. P., & McGlinchey, R. E. (2012). Distinct functional networks within the cerebellum and their relation to cortical systems assessed with independent component analysis. *NeuroImage*, 60(4), 2073–2085. <https://doi.org/10.1371/journal.pone.0044428>
- Dohmatob, E., Varoquaux, G., & Thirion, B. (2018). Inter-subject registration of functional images: Do we need anatomical images? *The Frontiers in Neuroscience*, 12(64), 1–11. <https://doi.org/10.3389/fnins.2018.00064>
- Du, Y., & Fan, Y. (2013). Group information guided ICA for fMRI data analysis. *NeuroImage*, 69, 157–197. <https://doi.org/10.1016/j.neuroimage.2012.11.008>
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *PNAS*, 113(28), 7900–7905. <https://doi.org/10.1073/pnas.1602413113>
- Erhardt, E. B., Allen, E. A., Wei, Y., Eichele, T., & Calhoun, V. D. (2012). SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. *NeuroImage*, 59(4), 4160–4167. <https://doi.org/10.1016/j.neuroimage.2011.11.088>
- Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., et al. (2008). The maturing architecture of the brain's

- default network. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10), 4028–4032. <https://doi.org/10.1073/pnas.0800376105>
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI). *Magnetic Resonance in Medicine*, 33(5), 636–647. <https://doi.org/10.1002/mrm.1910330508>
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9673–9678. <https://doi.org/10.1073/pnas.0504136102>
- Gao, W., Alcauter, S., Smith, J. K., Gilmore, J. H., & Lin, W. (2015). Development of human brain cortical network architecture during infancy. *Brain Structure & Function*, 220(2), 1173–1186. <https://doi.org/10.1007/s00429-014-0710-3>
- Gaudio, S., Piervincenzi, C., Zobel, B. B., Montecchi, F. R., Riva, G., Carducci, F., & Quattrocchi, C. C. (2015). Altered resting state functional connectivity of anterior cingulate cortex in drug naïve adolescents at the earliest stages of anorexia nervosa. *Scientific Reports*, 5(10818), 1–11. <https://doi.org/10.1038/srep10818>
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5(13), 1–15. <https://doi.org/10.3389/fninf.2011.00013>
- Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, 16(9), 1484–1492. <https://doi.org/10.1162/0898929042568532>
- Griffanti, L., Douaud, G., Bijstervosch, J., Evangelisti, S., Alfaro-Almagro, F., Glasser, M. F., ... Smith, S. M. (2017). Hand classification of fMRI ICA noise components. *NeuroImage*, 154, 188–205. <https://doi.org/10.1016/j.neuroimage.2016.12.036>
- Gu, S., Satterthwaite, T. D., Medaglia, J. D., Yang, M., Gur, R. E., Gur, R. C., & Bassett, D. S. (2015). Emergence of system roles in normative neurodevelopment. *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), 13681–13686. <https://doi.org/10.1073/pnas.1502829112>
- Harrison, S. J., Woolrich, M. W., Robinson, E. C., Glasser, M. F., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2015). Large-scale probabilistic functional modes from resting state fMRI. *NeuroImage*, 109, 217–231. <https://doi.org/10.1016/j.neuroimage.2015.01.013>
- Hyvarinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4), 411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5)
- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62(9), 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
- Killgore, W., Smith, R., Olson, E. A., Weber, M., Rauch, S. L., & Nickerson, L. D. (2017). Emotional intelligence is associated with connectivity within and between resting state networks. *Social Cognitive and Affective Neuroscience*, 12(10), 1624–1636. <https://doi.org/10.1093/scan/nsx088>
- Koch, W., Teipel, S., Mueller, S., Buerger, K., Bokde, A. L., Hampel, H., et al. (2010). Effects of aging on default mode network activity in resting state fMRI: Does the method of analysis matter. *NeuroImage*, 51(1), 280–287. <https://doi.org/10.1016/j.neuroimage.2009.12.008>
- Kooijman, M. N., Kruithof, J., van Duijn, C. M., Duijts, L., Franco, O. H., van Ijzendoorn, M. H., et al. (2016). The generation R study: Design and cohort update 2017. *European Journal of Epidemiology*, 31(12), 1243–1264. <https://doi.org/10.1007/s10654-016-0224-9>
- Lai, R., & Osher, S. (2014). A splitting method for orthogonality constrained problems. *Journal of Scientific Computing*, 58(2), 431–449. <https://doi.org/10.1007/s10915-013-9740-x>
- Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting state fMRI: A review of methods and clinical applications. *AJNR. American Journal of Neuroradiology*, 34(10), 1866–1872. <https://doi.org/10.3174/ajnr.A3263>
- Liu, D. C., & Nocedal, J. (1989). On the limited memory bfgs method for large scale optimization. *Journal Mathematical Programming: Series A and B*, 45(3), 503–528. <https://doi.org/10.1007/BF01589116>
- Madhyastha, T., Peverill, M., Koh, N., McCabe, C., Flournoy, J., Mills, K., ... McLaughlin, K. A. (2017). Current methods and limitations for longitudinal fmri analysis across development. *Developmental Cognitive Neuroscience*, 33, 118–128. <https://doi.org/10.1016/j.dcn.2017.11.006>
- Müller-Dehring, E. M., Kwon, D., Nagel, B. J., Sullivan, E. V., Chu, W., Rohlfing, T., ... Pohl, K. M. (2018). Influences of age, sex, and moderate alcohol drinking on the intrinsic functional architecture of adolescent brains. *Cerebral Cortex*, 28(3), 1049–1063. <https://doi.org/10.1093/cercor/bhx014>
- Nichols, B. N., & Pohl, K. M. (2015). Neuroinformatics software applications supporting electronic data capture, management, and sharing for the neuroimaging community. *Neuropsychology Review*, 25(3), 356–368. <https://doi.org/10.1007/s11065-015-9293-x>
- Nora, D. V., George, F. K., Robert, T. C., Diana, W. B., Gordon, J. A., Koroshetz, W. J., et al. (2018). The conception of the abcd study: From substance use to a broad nih collaboration. *Developmental Cognitive Neuroscience*, 32, 4–7. <https://doi.org/10.1016/j.dcn.2017.10.002>
- Odish, O. F., van den Berg-Huysmans, A. A., van den Bogaard, S. J., Dumas, E. M., Hart, E., Rombouts, S. A., et al. (2015). Longitudinal resting state fMRI analysis in healthy controls and premanifest huntington's disease gene carriers: A three-year follow-up study. *Human Brain Mapping*, 36(1), 110–119. <https://doi.org/10.1002/hbm.22616>
- Oron Semper, J. V., Murillo, J. I., & Bernacer, J. (2016). Adolescent emotional maturation through divergent models of brain organization. *Frontiers in Psychology*, 7, 1263. <https://doi.org/10.3389/fpsyg.2016.01263>
- Pamilo, S., Malinen, S., Hlushchuk, Y., Seppä, M., Tikka, P., & Hari, R. (2012). Functional subdivision of group-ICA results of fMRI data collected during cinema viewing. *PLoS One*, 7(7), e42000. <https://doi.org/10.1371/journal.pone.0042000>
- Park, H., Pae, C., Friston, K., Jang, C., Razi, A., Zeidman, P., et al. (2017). Hierarchical dynamic causal modeling of resting-state fMRI reveals longitudinal changes in effective connectivity in the motor system after thalamotomy for essential tremor. *Frontiers in Neurology*, 8(346), 1–9. <https://doi.org/10.3389/fneur.2017.00346>
- Pfefferbaum, A., Rohlfing, T., Pohl, K. M., Lane, B., Chu, W., Kwon, D., ... Sullivan, E. V. (2016). Adolescent development of cortical and white matter structure in the ncanda sample: Role of sex, ethnicity, puberty, and alcohol drinking. *Cerebral Cortex*, 26(10), 4101–4121. <https://doi.org/10.1093/cercor/bhv205>
- Power, J. D., Fair, D. A., Schlaggar, B. L., & Petersen, S. E. (2010). The development of human functional brain networks. *Neuron*, 67(5), 735–748. <https://doi.org/10.1016/j.neuron.2010.08.017>
- Rohlfing, T., Cummins, K., Henthorn, T., Chu, W., & Nichols, B. N. (2013). N-CANDA data integration: Anatomy of an asynchronous infrastructure for multi-site, multi-instrument longitudinal data capture. *Journal of the American Medical Informatics Association*, 21(4), 758–762. <https://doi.org/10.1136/amiajnl-2013-002367>
- Rohlfing, T., Maurer, C. R., Jr., Bluemke, D. A., & Jacobs, M. (2003). Volume-preserving nonrigid registration of mr breast images using free-form deformation with an incompressibility constraint. *IEEE TMI*, 22(6), 730–741. <https://doi.org/10.1109/TMI.2003.814791>
- Rohlfing, T., Zahr, N. M., Sullivan, E. V., & Pfefferbaum, A. (2010). The SRI24 multi-channel atlas of normal adult human brain structure. *Human Brain Mapping*, 31(5), 798–819. <https://doi.org/10.1002/hbm.20906>
- Rosazza, C., & Minati, L. (2011). Resting-state brain networks: Literature review and clinical applications. *Neurological Sciences*, 32(5), 773–785. <https://doi.org/10.1007/s10072-011-0636-y>
- Schmithorst, V. J., & Holland, S. K. (2004). Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data. *Journal of Magnetic Resonance Imaging*, 19(3), 365–368. <https://doi.org/10.1002/jmri.20009>
- Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, 46, 561–584. <https://doi.org/10.1146/annurev.ps.46.020195.003021>
- Shen, H. H. (2015). Core concept: Resting-state connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 112(46), 14115–14116. <https://doi.org/10.1073/pnas.1518785112>
- Sherman, L. E., Rudie, J. D., Pfeifer, J. H., Masten, C. L., McNealy, K., & Dapretto, M. (2014). Development of the default mode and central executive networks across early adolescence: A longitudinal study. *Developmental Cognitive Neuroscience*, 10, 148–159. <https://doi.org/10.1016/j.dcn.2014.08.002>

- Smith, S., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, 44(1), 83–98. <https://doi.org/10.1016/j.neuroimage.2008.03.061>
- Stoffers, D., Diaz, B. A., Chen, G., den Braber, A., van 't Ent, D., Boomsma, D. I., et al. (2015). Resting-state fMRI functional connectivity is associated with sleepiness, imagery, and discontinuity of mind. *PLoS One*, 10(11), e0142014. <https://doi.org/10.1371/journal.pone.0142014>
- Sullivan, E. V., Brumback, T., Tapert, S. F., Fama, R., Prouty, D., Brown, S., et al. (2016). Cognitive, emotion control, and motor performance of adolescents in the NCANDA study: Contributions from alcohol consumption, age, sex, ethnicity, and family history of addiction. *Neuropsychology*, 30(4), 449–473. <https://doi.org/10.1037/neu0000259>
- Tai, X., & Wu, C. (2009). Augmented lagrangian method, dual methods and split bregman iteration for rof model. *SSVM 09 Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision*. (doi: [https://doi.org/10.1007/978-3-642-02256-2\\_42](https://doi.org/10.1007/978-3-642-02256-2_42))
- van der Horn, H. J., Scheenen, M. E., de Koning, M. E., Liemburg, E. J., Spikman, J. M., & van der Naalt, J. (2017). The default mode network as a biomarker of persistent complaints after mild traumatic brain injury: A longitudinal functional magnetic resonance imaging study. *Journal of Neurotrauma*, 34(23), 3262–3269. <https://doi.org/10.1089/neu.2017.5185>
- van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, 17(12), 683–696. <https://doi.org/10.1016/j.tics.2013.09.012>
- Vetter, N. C., Steding, J., Jurk, S., Ripke, S., Mennigen, E., & Smolka, M. N. (2017). Reliability in adolescent fmri within two years – A comparison of three tasks. *Scientific Reports*, 7(2287), 2287. <https://doi.org/10.1038/s41598-017-02334-7>
- Wang, Y., Guo, Y. (2018). A hierarchical independent component analysis model for longitudinal neuroimaging studies. *Preprint arxiv.org/abs/1808.01557*
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92, 381–397. <https://doi.org/10.1016/j.neuroimage.2014.01.060>
- Xu, Y., & Yin, W. (2013). A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM Journal on Imaging Sciences*, 6(3), 1758–1789. <https://doi.org/10.1137/120887795>

**How to cite this article:** Zhao Q, Kwon D, Müller-Oehring EM, et al. Longitudinally consistent estimates of intrinsic functional networks. *Hum Brain Mapp*. 2019;40: 2511–2528. <https://doi.org/10.1002/hbm.24541>

## APPENDIX A: MULTI-SPLITTING METHOD FOR LONGITUDINAL ICA

This section introduces a numerical algorithm that efficiently solves the objective function of RL-ICA (Equation 13). The algorithm is based on the splitting method introduced in (Lai & Osher, 2014), which was originally designed for objective functions with a single orthogonality constraint. Here the splitting method is extended to handle multiple orthogonality constraints and homogeneous linear constraints as in the form of Equation 13. As we show at the end of the section, the algorithm for solving Equation 9. (VL-ICA) directly follows as a specific case of Equation 13 (RL-ICA).

To derive the algorithm, let  $\hat{N}$  be the orthogonal basis for the null space of  $B_{ROI}$  and let  $\hat{W}$  be reparameterized via  $\hat{W} = \hat{N}Y$ . By construction, the homogeneous linear system in Equation 12 is strictly satisfied

under this reparameterization, that is,  $B_{ROI}\hat{N}Y = 0$  for any  $Y$ . Now, Equation 13 can be rewritten with respect to  $Y$  as

$$\begin{aligned} & \operatorname{argmin}_Y E(\hat{N}Y), \\ & \text{s.t. } \hat{W} = \hat{N}Y, \text{ and } W^j T W^j = I, \text{ for } j = 1, \dots, V \end{aligned} \quad (14)$$

To solve this minimization problem, the Bregman method (Tai & Wu, 2009) is applied with an auxiliary variable  $\hat{D}$ , which leads to the iterative procedure of Algorithm 1.

### Algorithm 1 The Bregman Iteration

1. Optimize  $Y, \hat{W}$  with the following augmented objective function.

$$\begin{aligned} \{Y^{k+1}, \hat{W}^{k+1}\} &= \operatorname{argmin}_{Y, \hat{W}} E(\hat{N}Y) + \rho \|\hat{N}Y - \hat{W} + \hat{D}^k\|_F^2, \\ & \text{s.t. } W^j T W^j = I, \text{ for } j = 1, \dots, V \end{aligned} \quad (15)$$

2. Update  $\hat{D}$

$$\hat{D}^{k+1} = \hat{D}^k + \hat{N}Y^{k+1} - \hat{W}^{k+1}$$

To determine the solution to Step 1 of in each Bregman iteration, Block Coordinate Descent (Xu & Yin, 2013) further splits the minimization problem of Equation (15) into an unconstrained minimization problem with respect to  $Y$  and a set of constrained minimization problems with respect to each  $W^j$  in  $\hat{W}$ . Now let  $\hat{N} := [N^1 T, \dots, N^V T]^T$  such that  $W^j = N^j Y$ , and let  $\hat{D} := [D^1 T, \dots, D^V T]^T$ . In so doing, Algorithm 1 is turned into the multi-splitting method (Algorithm 2)

### Algorithm 2 Multi-Splitting Method

1: Optimize the unconstrained problem with respect to  $Y$

$$Y^{k+1} = \operatorname{argmin}_Y E(\hat{N}Y) + \rho \|\hat{N}Y - \hat{W}^k + \hat{D}^k\|_F^2$$

2: For each  $j \in \{1, \dots, V\}$ , optimize

$$W^{jk+1} = \operatorname{argmin}_{W^j} \rho \|N^j Y^{k+1} - W^j + D^j\|_F^2 \text{ s.t. } W^j T W^j = I \quad (16)$$

3: Update  $\hat{D}$

$$\hat{D}^{k+1} = \hat{D}^k + \hat{N}Y^{k+1} - \hat{W}^{k+1}$$

The efficiency of the multi-splitting method comes from the fact that (a) Step 1 of can be optimized with any gradient-based method (e.g., here L-BFGS [Liu & Nocedal, 1989] was used) as the objective is unconstrained; and (b) Step 2 of has the following closed-form solution:

Let the SVD factorization of  $N^j Y^{k+1} + D^j$  be

$$N^j Y^{k+1} + D^j = U^T \Sigma V$$

then the solution to Equation (16) is

$$w_j^{k+1} = U^T IV.$$

The derivation of this closed-form solution can be found in (Lai & Osher, 2014).

*The Application to GIG-ICA and VL-ICA.* The derivation of this closed-form solution can be found in (Lai & Osher, 2014). Replacing the orthogonal basis  $\hat{N}$  in Equation 14 with an identity matrix, Equation 14 is then numerically equivalent to the objective function of GIG-ICA (Equation 4) and VL-ICA (Equation 9), and the multi-splitting method can be applied in the same way.

## APPENDIX B: SIMTB-BASED SIMULATION OF SYNTHETIC DATASETS

In the SimTB-based experiments, a longitudinal rs-fMRI or a “subject” contained three synthetic “visit-specific” rs-fMRI images, which were generated based on  $C$  subject-specific IFNs. Each of the  $C$  networks was related to four ROIs randomly selected from the SimTB template (Figure 2a). To create between-subject variation, the  $C$  subject-specific IFNs were randomly translated by sampling from a uniform distribution

$U(-\tau_1, \tau_1)$  (voxel) and randomly rotated by sampling from  $U(-\tau_2, \tau_2)$  (degree) to create the ground-truth, subject-specific IFNs at baseline. To model intra-subject variability of an IFN, the functional connectivity at baseline was altered within each region according to the “slope” sampled from a normal distribution  $N(0, \mu)$ , resulting in the ground-truth subject-specific IFN of the second visit. Using the same slope, the process was repeated to generate the ground-truth for the third visit (Figure 2c).

Based on the ground-truth IFNs at each visit, a synthetic rs-fMRI image consisting of 150 2D images was generated by coupling each IFN with time courses reflecting activity of 29 ROIs (provided by the SimTB template Figure 2a). Specifically for each time point (of the time course), ROIs belonging to the same IFN activated simultaneously with probability  $1/C$  (event activation), and each individual ROI was set to activate with probability  $0.1/C$  chance (unique activation). Based on those activity patterns, the time course associated with each ROI was then generated via the hemodynamic response function model (Buxton, Uludağ, Dubowitz, & Liu, 2004).

From those time courses, SimTB generated a synthetic rs-fMRI image. Finally, Rician noise of a certain contrast-noise-ratio (CNR) was added to the synthetic rs-fMRI image. This procedure was repeated for each visit resulting in a subject-specific longitudinal rs-fMRI.

The 50 longitudinal rs-fMRIs in each synthetic dataset were produced using the same simulation setting. For the first two datasets the setting was  $C = 5$ ,  $\mu = 0.05$ ,  $\tau_1 = 4$ ,  $\tau_2 = 6$  and  $CNR = 3$ . The simulation settings of the remaining 20 datasets are summarized in Tables 3–6.

## APPENDIX C: STATISTICAL ANALYSIS OF LONGITUDINAL TRAJECTORIES

This section describes the procedure of performing statistical group analysis for detecting group-level developmental effects on the 246 longitudinal rs-fMRI of the NCANDA dataset.

*Voxel-wise analysis.* After all subject-specific IFNs (246 subjects  $\times$  23 IFNs  $\times$  3 visits) were estimated by a specific approach, group-level developmental effects were detected by performing statistical group analysis on subject-specific “slope maps.”

Specifically, for each three subject-specific IFNs associated with the three visits, a slope map (or longitudinal trajectory map) was generated by computing a slope at each voxel via linear regression. Afterwards, the slope maps of all 246 subjects were used to test for group-level developmental effects associated with that IFN. As revealed in the previous cross-sectional analysis of NCANDA baseline data (Müller-Oehring et al., 2018), modeling scanner difference would account for the site difference because the acquisition protocol was exactly the same for the 3 GE sites and for the 2 Siemens sites. Therefore, the effects of scanner and sex were considered as confounding factors. The previous analysis also revealed that as GLM analysis regressed them out from the slopes. Group-level developmental effects with respect to the residuals were then tested within a gray-matter mask using one-sample  $t$ -tests as implemented in FSL PALM (Winkler, Ridgway, Webster, Smith, & Nichols, 2014; Permutation Analysis of Linear Models) with 5,000 permutations. TFCE (Smith & Nichols, 2009) was used for family-wise error correction. The above procedure was repeated to analyze each IFN. The number of IFNs was then further corrected by the Bonferroni procedure (Shaffer, 1995), which led to a significance-level threshold of one-sided  $p \leq 0.05/23 = 0.0022$  (after TFCE family-wise error correction). Additionally, clusters with more than 50 significant voxels were considered as true positives.

*Cluster analysis.* Any cluster of voxels that reached the significance level was further analyzed by LME analysis. For each cluster detected based on the IFN estimates of one specific approach, the LME analysis was applied repeatedly based on the IFN estimates produced by each of the four approaches. Based on the IFNs estimated by a specific approach, the functional connectivity in the cluster was first averaged. The effects of sex and scanner were then removed from the averaged values by regression analysis. LME analysis was applied to the residual values with a fixed effect of linear aging and a random effect of intercept. The resulting fixed effect was considered significant if one-sided  $p \leq 0.001$ .

*Motion analysis.* To ensure the detected developmental effects were not confounded by subject-specific motion, three statistical tests measured the correlation of motion with either age or functional connectivity. First, a slope with respect to the number of outlier-frames (according to the preprocessing) over the three visits was computed for each subject. A one-sample  $t$ -test was applied to the slopes of all subjects resulting in an insignificant finding ( $p > 0.2$ ). Second, the linear fixed effect of age on the number of outlier-frames was investigated via LME, which was also insignificant ( $p > 0.2$ ). These two experiments indicated there was no subject-specific or group-level change in motion with respect to age. Third, the motion-related slope was correlated with the slope of functional connectivity in each of the detected clusters, all resulting in insignificant  $p$ -values ( $>0.05$ ). This suggests that the slope of functional connectivity fully reflected age-related developmental effects and was not confounded by the change in motion.