Supplementary Material for "A Multimodal Multilevel Neuroimaging Model for Investigating Brain Connectome Development"

Yingtian Hu[a], Mahmoud Zeydabadinezhad[b],Longchuan Li[b] and Ying Guo[a] *

[a] Department of Biostatistics, Emory University, Atlanta, GA 30322

[b] Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 30322

*Correspondence: Ying Guo, yguo2@emory.edu. The authors gratefully acknowledge support from NIH under award number R01MH105561, R01MH118771,R01EB027147,R01MH119251

1 Details of EM Algorithm:

As stated in the main text, the complete data log-likelihood function can be represented as

$$l_C(\mathbf{\Theta}; \mathcal{D}, \mathcal{R}, \mathcal{A}, \mathcal{F}, \mathcal{T}) = \sum_{g=1}^{G} \sum_{j=1}^{Q-1} \sum_{k:j<k} \sum_{u=0}^{1} \sum_{v=-1}^{1} \mathbb{I}(A_{jk,g} = u, F_{jk,g} = v) \times log\{\Psi_{jk,g}^{u,v}(\mathbf{\Theta_1}) \prod_{i_g=1}^{I_g} \phi^{u,v}(D_{jk,g,i_g}, R_{jk,g,i_g}\tau_{jk,g,i_g,l}|\mathbf{\Theta_2}, \mathbf{\Theta_3})\},$$

where

$$\Psi_{jk,g}^{u,v}(\mathbf{\Theta}_1) = (\pi_{m(j)m(k),g}^{\alpha})^u (1 - \pi_{m(j)m(k),g}^{\alpha})^{1-u} p_{m(j)m(k),g}^{u,v},$$

$$\phi^{u,v}(D_{jk,g,i_g}, R_{jk,g,i_g}, \tau_{jk,g,i_g,l}|\mathbf{\Theta}_2, \mathbf{\Theta}_3) = [\rho^u \delta(p_{jk,g,i_g}) + (1-\rho^u) \times (\sum_{l=1}^{L} \gamma_l^u \tau_{jk,g,i_g,l} f(D_{jk,g,i_g}; \chi_l^u, (\xi_l^u)^2)] \times f(R_{jk,g,i_g}; \mu^{u,v}, (\sigma^{u,v})^2).$$

To simplify our notation in the derivation, we introduce a joint latent variable $Z_{jk,g}^{u,v} = \mathbb{I}(A_{jk,g} = u, F_{jk,g} = v)$ to take the place of the latent variables $A_{jk,g}$ and $F_{jk,g}$ with $\sum_{u=0}^{1} \sum_{v=-1}^{1} Z_{jk,g}^{u,v} = 1$. We then rewrite the complete data log-likelihood function as

$$l_C(\mathbf{\Theta}; \mathcal{D}, \mathcal{R}, \mathcal{Z}, \mathcal{T}) = \sum_{g=1}^{G} \sum_{j=1}^{Q-1} \sum_{k:j<k} \sum_{u=0}^{1} \sum_{v=-1}^{1} Z_{jk,g}^{u,v} log\{\Psi_{jk,g}^{u,v}(\mathbf{\Theta_1}) \prod_{i_g=1}^{I_g} \phi^{u,v}(D_{jk,g,i_g}, R_{jk,g,i_g}\tau_{jk,g,i_g}|\mathbf{\Theta_2}, \mathbf{\Theta_3})\}.$$

1.1 The Conditional Expectation Function in the E-step

The E-step of our EM algorithm which evaluates the conditional expectation of the complete data log-likelihood can be presented as

$$Q(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)}) = Q_1(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)}) + Q_2(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)}) + Q_3(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)}) + Q_4(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)})$$

where

$$Q_1(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}^{(t)}) = \sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:k<j}\left[(\sum_{v=-1}^{1}E(Z_{jk,g}^{1,v}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)}))\boldsymbol{x}_g^T\boldsymbol{\beta}_{m(j)m(k)} - log(1+exp(\boldsymbol{x}_g^T\boldsymbol{\beta}_{m(j)m(k)}))\right],$$

$$\begin{aligned}Q_2(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}^{(t)}) = &\sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:k<j}\sum_{u=0}^{1}(E(Z_{jk,g}^{u,-1}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})\boldsymbol{x}_g^T\boldsymbol{\alpha}_{m(j),m(k)}^{u,-1} + E(Z_{jk,g}^{u,1}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})\boldsymbol{x}_g^T\boldsymbol{\alpha}_{m(j),m(k)}^{u,1}\\ &- \sum_{v=-1}^{1}E(Z_{jk,g}^{u,v}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})log(1+exp(\boldsymbol{x}_g^T\boldsymbol{\alpha}_{m(j),m(k)}^{u,-1})) + exp(\mathbf{x}_g^T\boldsymbol{\alpha}_{m(j),m(k)}^{u,1}))),\end{aligned}$$

$$\begin{aligned}Q_3(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}^{(t)}) = &\sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:k<j}\sum_{u=0}^{1}\sum_{v=-1}^{1}E(Z_{jk,g}^{u,v}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})(\sum_{D_{jk,g,i_g}=0}log(\rho^u) + \sum_{D_{jk,g,i_g}\neq 0}log(1-\rho^u))\\ &+\sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:k<j}\sum_{u=0}^{1}\sum_{D_{jk,g,i_g}\neq 0}\sum_{l=1}^{L}(\sum_{v=-1}^{1}E(Z_{jk,g}^{u,v}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)}))E(\tau_{jk,g,i_g,l}^{u}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})\\ &\times(log(\gamma_l^u) - log(\sqrt{2\pi}\xi_l^u) - \frac{(D_{jk,g,i_g}-\chi_l^u)^2}{2(\xi_l^u)^2}),\end{aligned}$$

$$Q_4(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}^{(t)}) = \sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:k<j}\sum_{u=0}^{1}\sum_{v=-1}^{1}E(Z_{jk,g}^{u,v}|\mathcal{Y};\hat{\boldsymbol{\Theta}}^{(t)})\sum_{i_g=1}^{I_g}(-log(\sqrt{2\pi}\sigma^{u,v}) - \frac{(R_{jk,g,i_g}-\mu^{u,v})^2}{2(\sigma^{u,v})^2}),$$

where $\mathcal{Y} = \{\mathcal{D}, \mathcal{R}\}$. In order to evaluate the Q-functions, we need to evaluate conditional expectations $E(Z_{jk,g}^{u,v}|\mathcal{Y};\boldsymbol{\Theta})$ and $E(\tau_{jk,g,i_g,l}^{u}|\mathcal{Y};\boldsymbol{\Theta})$, which can be readily obtained via Bayes's theorem. Specifically, let $q_{jk,g}^{u,v} = E(Z_{jk,g}^{u,v}|\mathcal{Y};\boldsymbol{\Theta})$ and we have

$$\begin{aligned}q_{jk,g}^{u,v} &= P(A_{jk,g}=u, F_{jk,g}=v|D_{jk,g,i_g}, R_{jk,g,i_g};\boldsymbol{\Theta})\\ &= \frac{P_{jk,g}^{u,v}}{\sum_{u=0}^{1}\sum_{v=-1}^{1}P_{jk,g}^{u,v}},\end{aligned}$$

where $P^{u,v}_{jk,g} = P(A_{jk,g} = u, F_{jk,g} = v|\mathbf{\Theta})P(D_{jk,g,i_g}, R_{jk,g,i_g}|A_{jk,g} = u, F_{jk,g} = v; \mathbf{\Theta})$ with

$$P(A_{jk,g} = u, F_{jk,g} = v|\mathbf{\Theta}) = (\pi^{\alpha}_{m(j)m(k),g})^u(1 - \pi^{\alpha}_{m(j)m(k),g})^{1-u}p^{u,v}_{m(j)m(k),g}, \quad \text{and}$$

$$P(D_{jk,g,i_g}, R_{jk,g,i_g}|A_{jk,g} = u, F_{jk,g} = v; \mathbf{\Theta}) = [\rho^u\delta(D_{jk,g,i_g}) + (1 - \rho^u) \times (\sum_{l=1}^{L} \gamma^u_l f(D_{jk,g,i_g}; \chi^u_l, (\xi^u_l)^2)] \times f(R_{jk,g,i_g}; \mu^{u,v}, (\sigma^{u,v})^2).$$

Similarly, let $w^u_{jk,g,i_g,l} = E(\tau^u_{jk,g,i_g,l}|\mathcal{Y}; \mathbf{\Theta})$ and we have

$$\begin{aligned} w^u_{jk,g,i_g,l} &= E(\tau^u_{jk,g,i_g,l}|D_{jk,g,i_g}, R_{jk,g,i_g}, \mathbf{\Theta}) \\ &= \frac{\gamma^u_l f(D_{jk,g,i_g}; \chi^u_l, (\xi^u_l)^2)}{\sum \gamma^u_l f(D_{jk,g,i_g}; \chi^u_l, (\xi^u_l)^2)} \end{aligned}$$

We can obtain $(\hat{q}^{u,v}_{jk,g})^{(t)}$ and $(\hat{w}^u_{jk,g,i_g,l})^{(t)}$ by plugging in $\hat{\mathbf{\Theta}}^{(t)}$, and then evaluate the Q-function $\mathrm{Q}(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(t)})$.

1.2 Details of the M-step in the EM

For $\mathbf{\Theta_1} = \{\boldsymbol{\beta}, \boldsymbol{\alpha}^{u,-1}, \boldsymbol{\alpha}^{u,1}\}$, there is no explicit form for the maximum likelihood solution. We use an iterative algorithm, such as Newton-Raphson method or gradient descent method, to obtain the maximum likelihood estimator at each iteration. For other parameters in $\mathbf{\Theta_2} = \{\rho^u, \gamma^u_l, \chi^u_l, (\xi^u_l)^2\}$ and $\mathbf{\Theta_3} = \{\mu^{u,v}, (\sigma^{u,v})^2\}$, we have following explicit solutions for

updating the parameters,

$$(\hat{\rho}^u)^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{v=-1}^{1}(\hat{q}_{jk,g}^{u,v})^{(t)} I_{jk,g}^{0}}{\sum g^G \sum_{j=1}^{Q}\sum_{k:k\neq j}\sum_{v=-1}^{1}(\hat{q}_{jk,g}^{u,v})^{(t)} I_g},$$

$$(\hat{\gamma}_l^u)^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}\neq 0}(\hat{w}_{jk,g,i_g,l}^{u})^{(t)}}{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}}\sum_{l=1}^{L}\neq 0(\hat{w}_{jk,g,i_g,l}^{u})^{(t)}},$$

$$(\hat{\chi}_l^u)^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}\neq 0}(\hat{w}_{jk,g,i_g,l}^{u})^{(t)} D_{jk,g,i_g}}{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}\neq 0}(\hat{w}_{jk,g,i_g,l}^{u})^{(t)}},$$

$$((\hat{\xi}_l^u)^2)^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}\neq 0}(\hat{w}_{jk,g,i_g,l}^{u})^{(t)}(D_{jk,g,i_g} - (\hat{\chi}_l^u)^{(t)})^2}{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}\sum_{D_{jk,g,i_g}\neq 0}(\hat{w}_{jk,g,i_g,l}^{u})^{(t)}},$$

$$(\hat{\mu}^{u,v})^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}(\hat{q}_{jk,g}^{u,v})^{(t)}\sum_{i_g=1}^{I_g} R_{jk,g,i_g}}{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k\neq j}(\hat{q}_{jk,g}^{u,v})^{(t)} * I_g},$$

$$((\hat{\sigma}^{u,v})^2)^{(t+1)} = \frac{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k<j}(\hat{q}_{jk,g}^{u,v})^{(t)}\sum_{i_g=1}^{I_g}(R_{jk,g,i_g} - (\hat{\mu}^{u,v})^t)^2}{\sum_{g=1}^{G}\sum_{j=1}^{Q}\sum_{k:k\neq j}(\hat{q}_{jk,g}^{u,v})^{(t)} * I_g}.$$

2 Penalized Complete log-likelihood Function:

As mentioned in the main text, the penalized complete log-likelihood function is defined as,

$$l_C^*(\mathbf{\Theta}; \mathcal{D}, \mathcal{R}, \mathcal{A}, \mathcal{F}, \mathcal{T}) = \sum_{g=1}^{G}\sum_{j=1}^{Q-1}\sum_{k:j<k} l_{jk,g}(\mathbf{\Theta}; \mathcal{D}, \mathcal{R}, \mathcal{A}, \mathcal{F}, \mathcal{T}) + l_P(\mathbf{\Theta}; \mathcal{D}, \mathcal{R}),$$

where $l_P(\mathbf{\Theta}, \mathcal{D}, \mathcal{R})$ is the penalty term defined as,

$$l_P(\mathbf{\Theta}, \mathcal{D}, \mathcal{R}) = \sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\frac{1}{2}log(|\mathbf{S}_{m_1m_2}|) + \sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\sum_{u=0}^{1}\frac{1}{2}log(|\mathbf{T}_{m_1m_2}^{u}|).$$

Here, m_1, m_2 representing the module memberships and u represents the latent structural

state. $\mathbf{S}_{m_1m_2}$ and $\mathbf{T}^u_{m_1m_2}$ play the same role as the Fisher Information matrix does in Firth's method. $\mathbf{S}_{m_1m_2}$ serves as the correction term for latent structural states estimation and $\mathbf{T}^u_{m_1m_2}$ serves as the correction term for latent functional states estimation. To present the expressions of $\mathbf{S}_{m_1m_2}$ and $\mathbf{T}^u_{m_1m_2}$, we first introduce several variables. Let Q_{m_1} denote the number of nodes in module m_1 and $C_{m_1m_2}$ denote the total number of connections between module m_1 and m_2. Then $C_{m_1m_2} = Q_{m_1} * Q_{m_2}$ if $m_1 \neq m_2$ and $C_{m_1m_2} = Q_{m_1} * (Q_{m_1} - 1)$ if $m_1 = m_2$.

We define $\mathbf{S}_{m_1m_2} = \mathbf{X}^T\mathbf{W}_{m_1m_2}\mathbf{X}$ with $\mathbf{X}$ being the $G \times p$ covariate matrix with the gth row being $\boldsymbol{x}_g$ and $\mathbf{W}_{m_1m_2}$ being a G by G diagonal matrix with the gth diagonal element as $I_g * C_{m_1m_2} * \pi^{\alpha}_{m_1m_2,g} * (1 - \pi^{\alpha}_{m_1m_2,g})$. $\mathbf{T}^u_{m_1m_2} = \tilde{\mathbf{X}}^T\tilde{\mathbf{W}}^u_{m_1m_2}\tilde{\mathbf{X}}$ with $\tilde{\mathbf{X}}^T = \mathbf{X}^T \otimes \mathbf{I}_2$ and $\tilde{\mathbf{W}}^u_{m_1m_2}$ is a $2G \times 2G$ block diagonal matrix with G 2×2 blocks $\tilde{\mathbf{W}}^u_{m_1m_2,g} = \{w^{e_1e_2}_g\}$. The elements in $\tilde{\mathbf{W}}^u_{m_1m_2,g} = \{w^{e_1e_2}_g\}$ are $w^{11}_g = t^u_{m_1m_2,g} * p^{u,-1}_{m_1m_2,g}(1 - p^{u,-1}_{m_1m_2,g})$, $w^{12}_g = w^{21}_g = -t^u_{m_1m_2,g} * p^{u,-1}_{m_1m_2,g}p^{u,1}_{m_1m_2,g}$ and $w^{22}_g = t^u_{m_1m_2,g} * p^{u,1}_{m_1m_2,g}(1 - p^{u,1}_{m_1m_2,g})$, where

$$t^u_{m_1m_2,g} = \sum_{j:m(j)=m_1} \sum_{k:m(k)=m_2} \sum_{v=-1}^{1} E(Z^{u,v}_{jk,g}|\mathcal{D}, \mathcal{R}, \boldsymbol{\Theta}),$$

which is a function of parameter $\boldsymbol{\Theta}$.

3 Additional Simulation Results

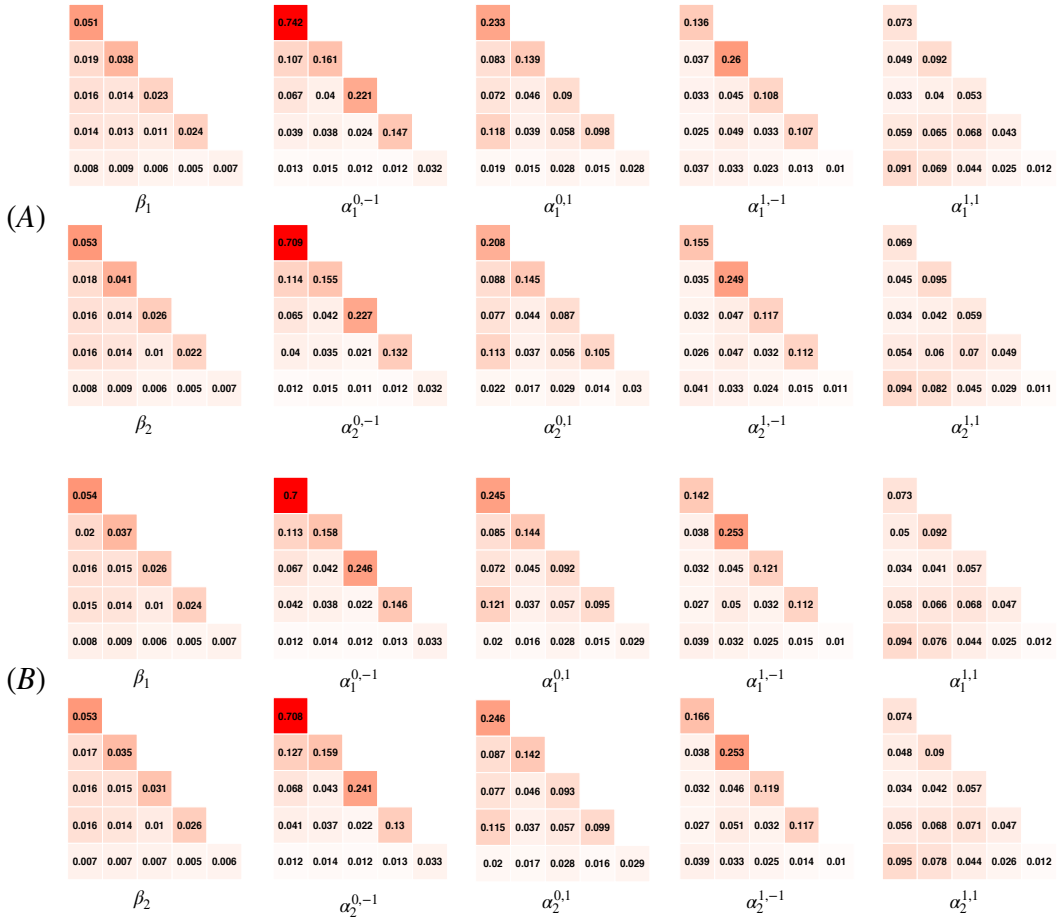

Figure S1: Variance Estimation Results of Parameters at Latent Network Modelling ($\boldsymbol{\Theta}_1$). Row(A) showed Monte Carlo variance of the parameter estimates across 1000 simulation replicates and Row(B) showed the proposed bootstrap variance estimates averaged across 200 bootstrap replicates. The bootstrap variance estimates were very close to the Monte Carlo variance of the parameters.

4 Power's Node System and Smith's Module System

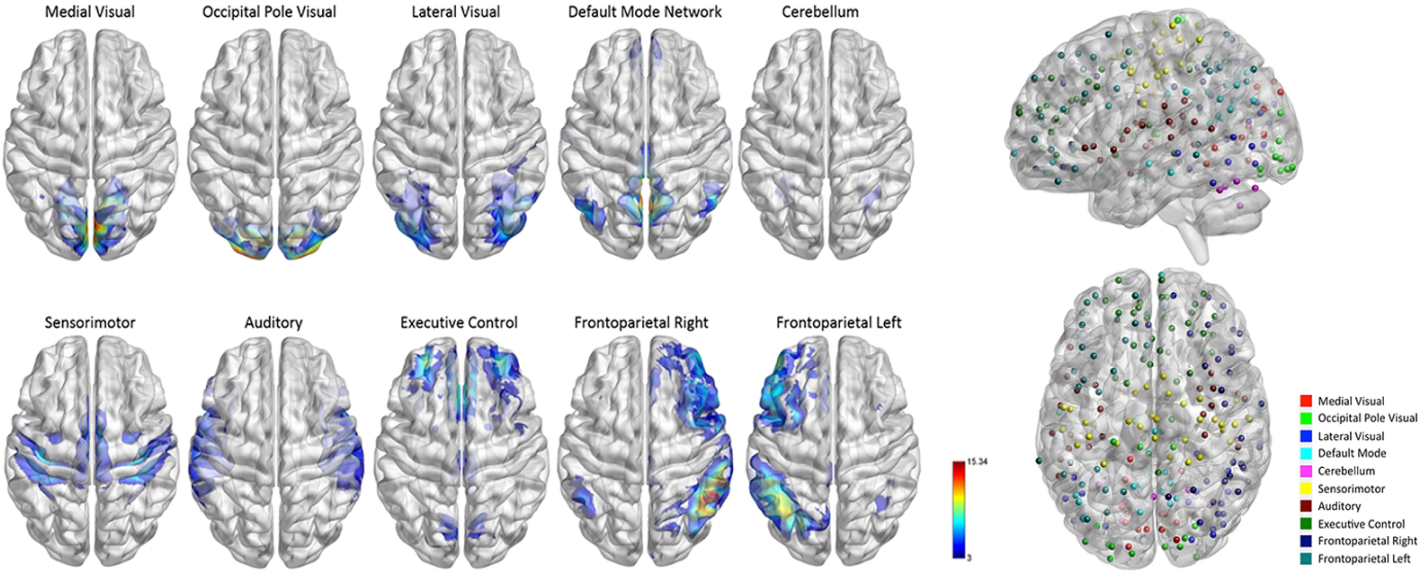

Figure S2: Power's node system and Smith's functional modules

5 Empirical Distribution of SC and FC

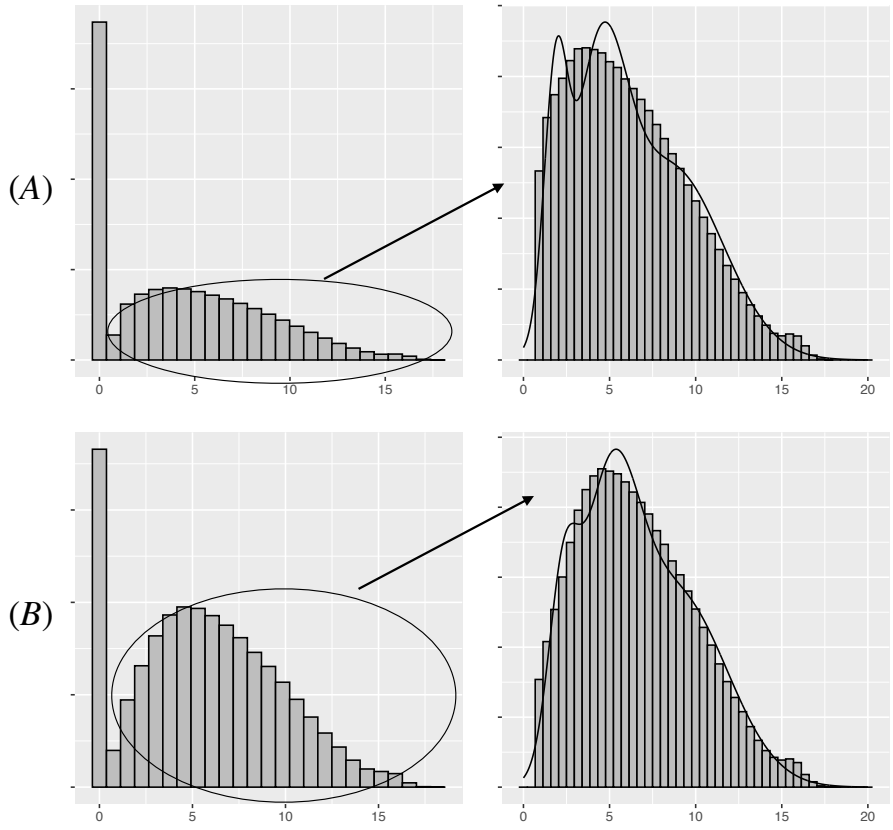

Figure S3: Empirical Distribution of Transformed SC in Empirically Estimated Latent Structural States. Empirical latent states are estimated by categorizing connections into 2 latent states based on transformed SC data. Row (A) shows the SC data with estimated latent structural states $\hat{A} = 0$ and row (B) shows the SC data with estimated latent structural states $\hat{A} = 1$. The black curve represents the fitted mixture of Gaussian distribution. The proposed distribution model provided a good fit to the empirical distribution of the SC measure.

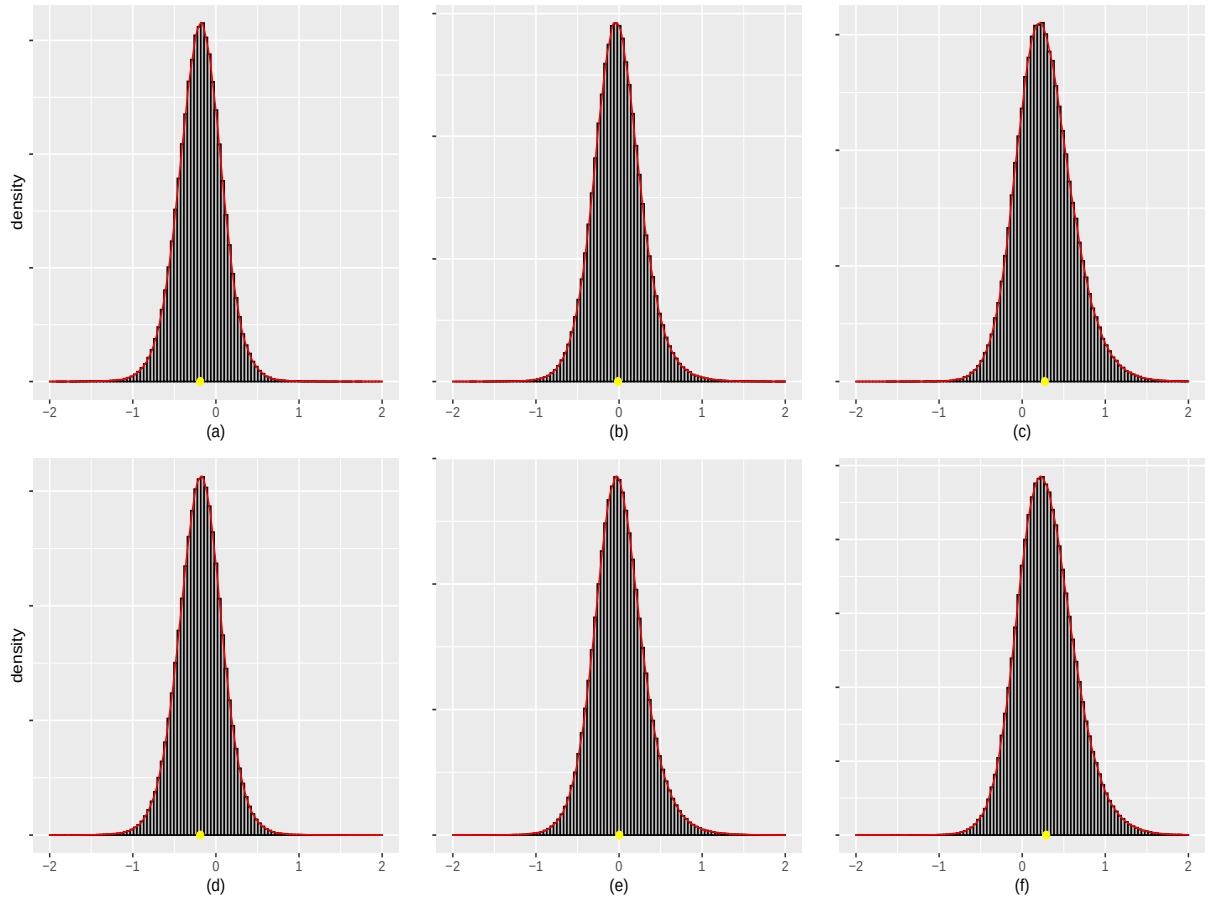

Figure S4: Empirical Distribution of FC in Empirically Estimated Latent States. Empirical latent states are estimated by categorizing connections into 6 latent states based on transformed SC and FC data. Gaussian distributions fitted to the data are overlaid in red. The yellow dots correspond to empirical means. (a) $\hat{A} = 0, \hat{F} = -1$; (b) $\hat{A} = 0, \hat{F} = 0$; (c) $\hat{A} = 0, \hat{F} = 1$; (d) $\hat{A} = 1, \hat{F} = -1$; (e) $\hat{A} = 1, \hat{F} = 0$; (f) $\hat{A} = 1, \hat{F} = 1$.

6 The estimated probabilities for negative functional connection state $F = -1$

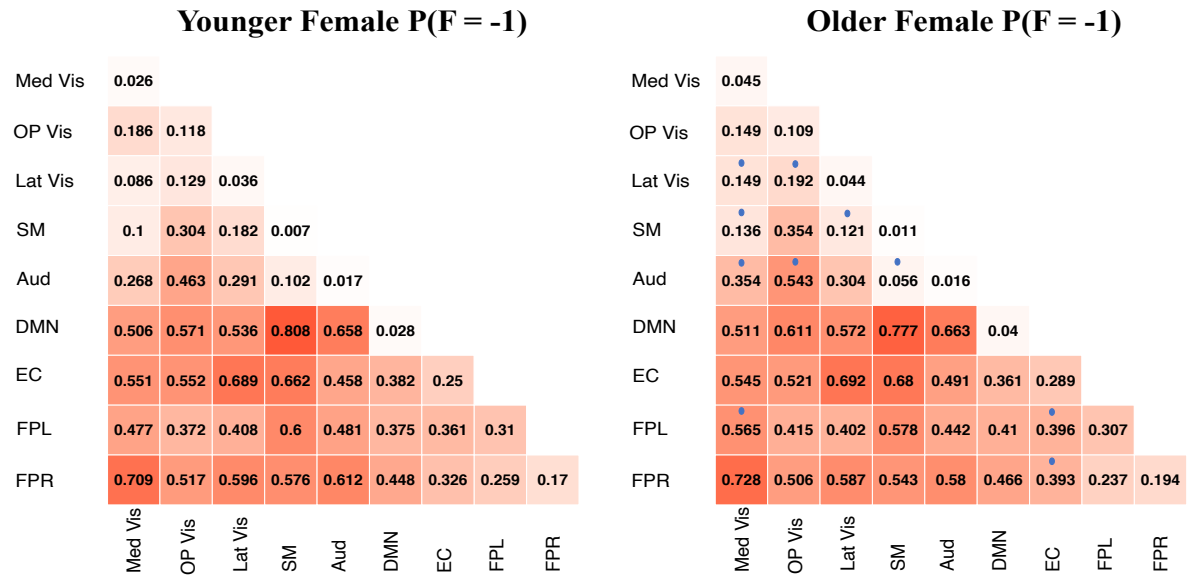

Figure S5: MMM estimated probabilities of latent functional connection equal to $F = -1$ in young female and old female groups. The numeric value in each module block indicates the estimated probability and the color shade indicates the magnitude of the probability. Module with blue dots in the older group are modules with significant difference between the two age groups at the significance level 0.01. The probability for negative functional connections tends to be low for within module connections and between modules with similar functionality such as the three visual modules.

7 MMM analysis of the PNC data with Power's Module System

We applied MMM to PNC data using an alternative module system to assess the reproducibility of the MMM analyses results with respect to different module systems. As a comparison with the Smith's module system used in the paper, we re-fit the MMM model using the module system which was used in the original paper of the Power's node atlas (Power et al. 2011). Based on Power's module system, the nodes are grouped into the function modules including visual module (Vis), sensory/somatomotor module (SM), auditory module (Aud), ventral attention module (Ventral), dorsal attention module (Dorsal), a cognitive module including default mode, executive control and frontoparietal network (DMN-EC-FP), cingulo-opercular task control module (CON), fronto-parietal task control module (FP-TC), a salience module including executive control (SN-EC), and a subcortical module including executive control (SubCort-EC).

Figure S6 presents the model estimated structural connection (SC) probabilities for the different subject groups based on Power's module system. Similar as the findings based on the Smith's module system, results in Figure S6 also showed that the within-module SC tend to be stronger than between-module SC and that the anatomical connection in the visual module is particularly strong. In Figure S7 presents the estimated difference in structural connection probabilities between the older and younger age group for female. The results for male were similar and hence omitted here and later. Figure S7 showed a general increase in white fiber structural connections across the brain with the increase in age. We observed that the increase in white matter fiber connection was especially noticeable and more statistically significant for higher-order cognitive networks such as DMN-EC-FP, SN-EC and SubCort-EC as compared to primary sensory and motor networks. These results are consistent with our findings presented based on Smith's module system.

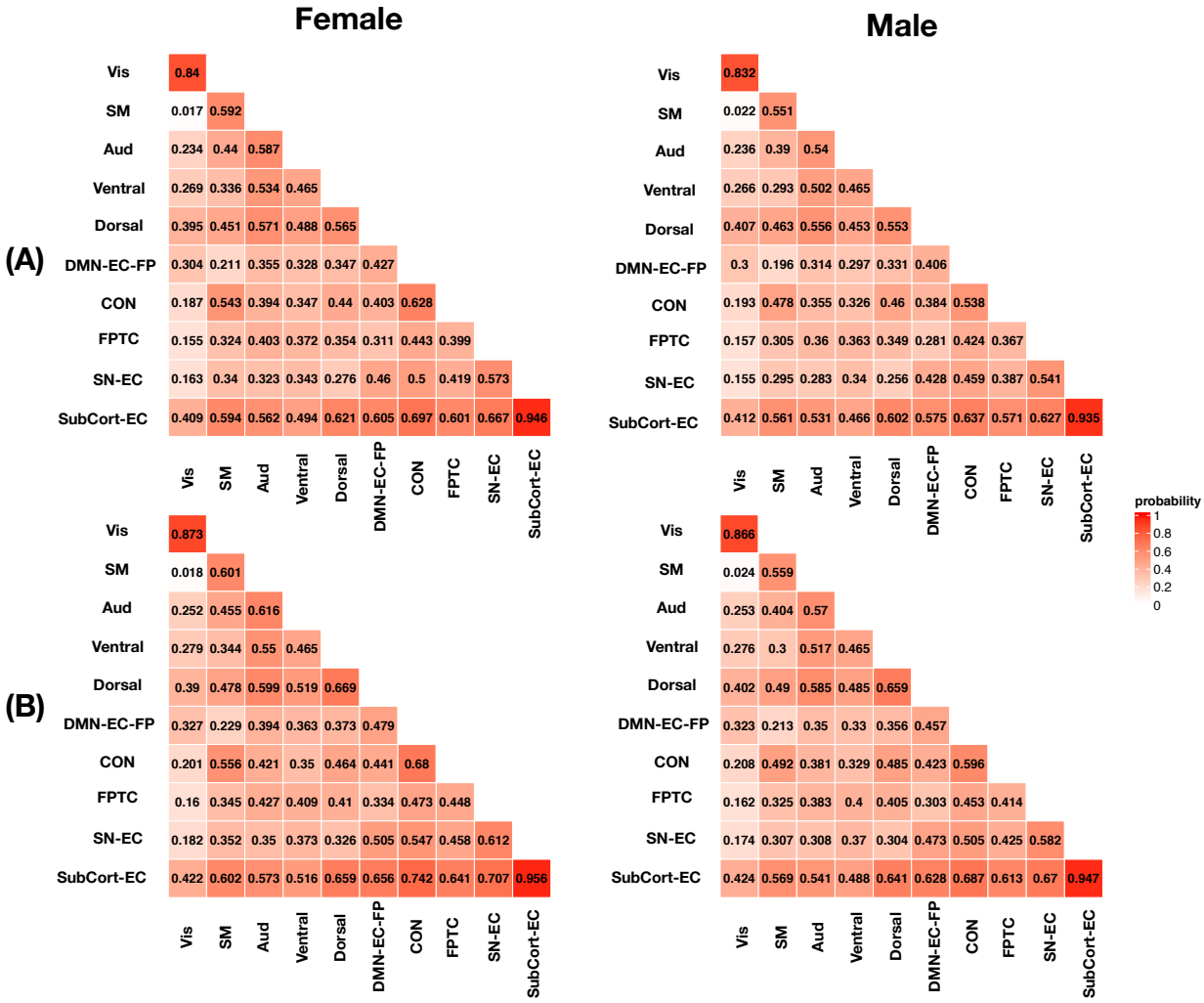

Figure S6: MMM estimated probabilities of latent structural connection (SC) in subject groups based on Power's module system. Row (A) represents younger group (Age 8-15) and Row (B) represents older group (Age 16 - 21). The numeric value in each module block indicates the estimated probability and the color shade indicates the magnitude of the probability.

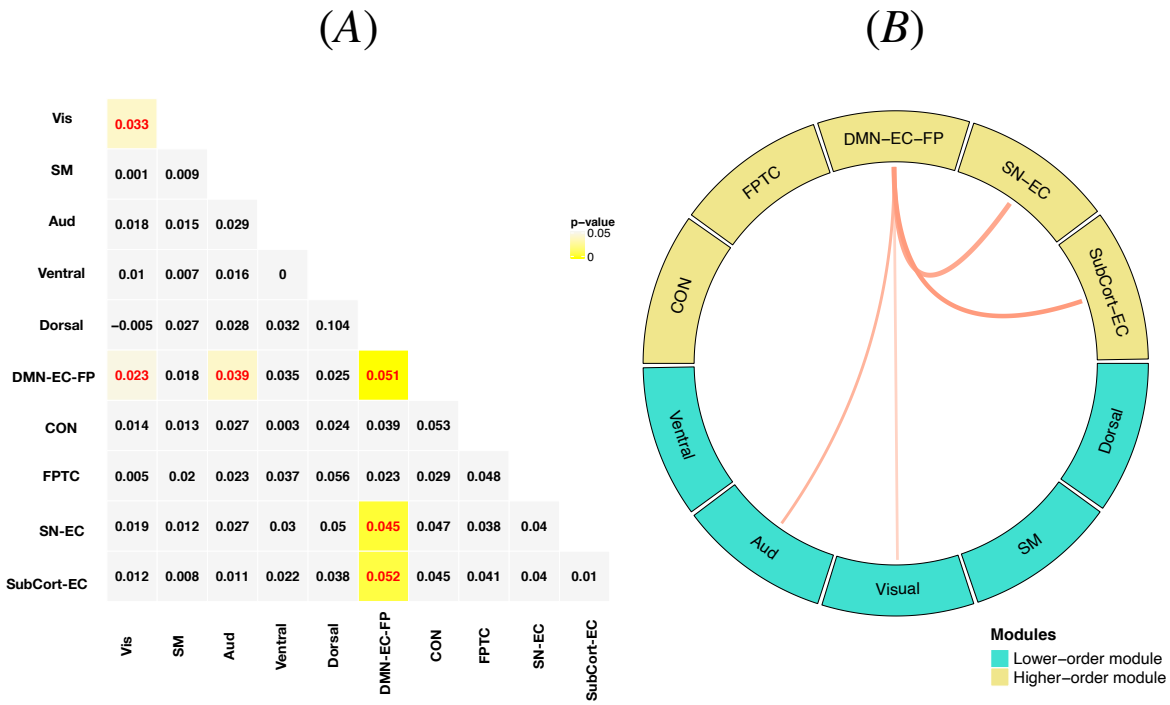

Figure S7: MMM estimated dierence in structural connection (SC) probabilities between the older (16-21) and younger (8-15) age group based on Power's module system. (A): the numeric values are the estimated age difference (older vs. younger) in the SC probabilities. We highlight in yellow the age differences that are are significant at the alpha=0.05 level, where the color of the numeric values indicates the direction of the difference (red=signiffcant positive difference; blue=significant negative difference). (B): A graphical illustration of the significant differences in SC across brain networks presented in (A). Turquoise modules represent high- order cognitive networks and yellow modules represent primary sensory and motor networks. The red lines show significantly increased SC with age, with the wider lines representing more significant age difference with smaller p-values. Results show a general increase in white fiber structural connections across the brain with neurodevelopment. Most of the SC increase are observed among connections involving higher-order cognitive networks.

For FC analyses, Figure S8 represents the estimated probability for the latent functional connectivity state F = 0 and F = 1 for two subject groups: younger female and older female, based on Power's module system. Similar as the findings based on Smith's system, results in Figure S8 show the highest probability of positive connections, i.e. F = 1, are found among within-module connections and that the probability of positive FC for between-module connections is generally higher between modules with similar type of functionality. Figure S9 presents the MMM estimated difference in functional connection probabilities between the older and younger age group for female based on the Power's module system. We obtain similar findings as with Smith's module system. As age increase from 8-15 to 16-21, the probability of the state of having no FC (F=0) generally decreases across networks while the probability of positive FC (F=1) mostly increase, indicating the brain becomes more functionally ordered or connected during adolescence.

We also evaluate the change in FC with increase of age conditional on latent structural connection states (Figure S10). Similar as the findings based on Smith's system, results from the Power's module system show that most significant changes in FC are observed at those edges with latent anatomical state of A = 0 while only limited age differences in FC are found for edges with A = 1, indicating again that functional connection development from late childhood and early adolescence to late adolescence mainly happen at edges that do not have strong direct structural connections.

In summary, the findings derived from MMM using Power's module system are mostly consistent with those from MMM using Smith's module system, showing the proposed MMM is able to yield consistent findings with different module systems.

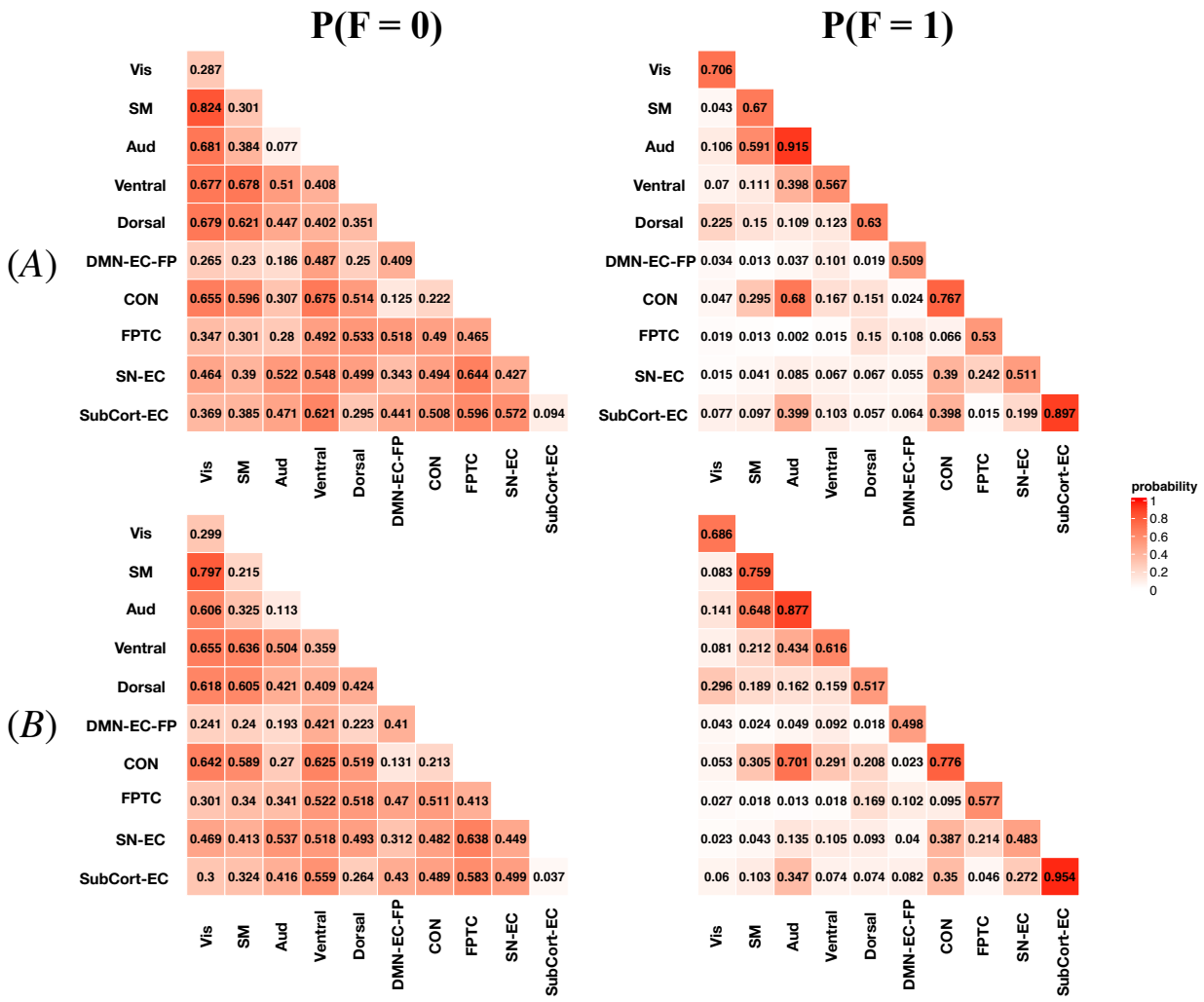

Figure S8: MMM estimated probabilities for the latent functional connectivity state F=0 and F=1 for two subject groups: (A) younger female and (B) older female, based on Power' module system. The numeric value in each module block indicates the estimated probability and the color shade indicates the magnitude. The highest probability of having positive connections, i.e. F= 1, are observed among within-module connections. For between-module connections, the probability of positive FCs are generally higher between modules with similar type of functionality.

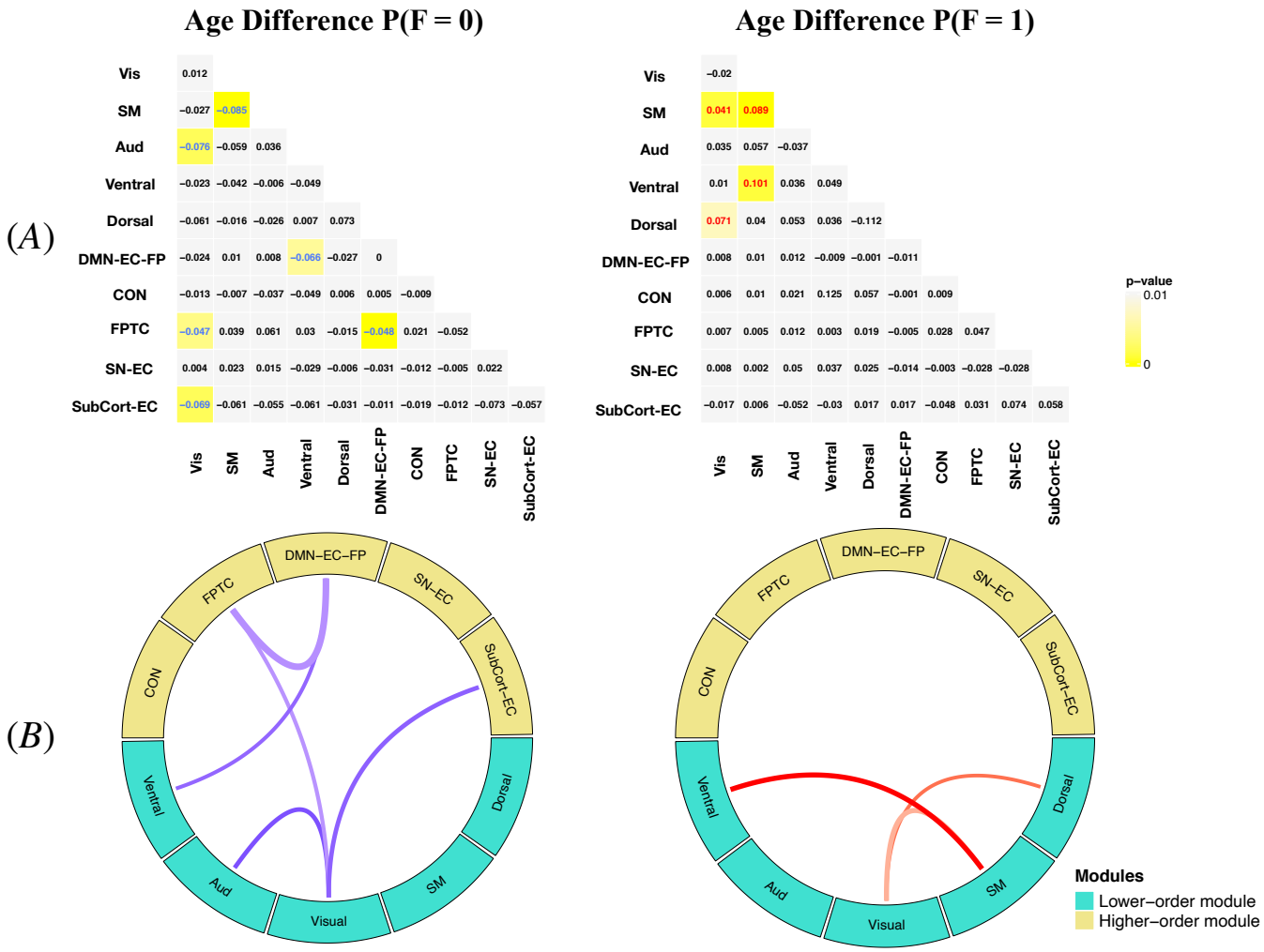

Figure S9: MMM estimated difference in probabilities of functional connection state between the older (16-21) and younger (8-15) age group based on Power's module system. (A): the estimated age difference (older vs. younger) for the probabilities of no FC (F=0) and positive FC (F=1). We highlight in yellow the age differences that are are significant at the alpha=0.05 level, where the color of the numerical values indicates the direction of the difference (red=significant positive difference; blue=significant negative difference). With age increases, the probability of no FC (F=0) generally decreases across the networks and the probability of positive FC (F=1) generally increases across the networks, indicating the brain gets more functionally organized with neurodevelopment. (B): A graphical illustration of the significant age differences in functional connections across brain networks presented in (A). Turquoise circles represent high-order cognitive networks and yellow circles represent primary sensory and motor networks. The blue lines show probabilities that significantly decreases with age increase and the red lines show probabilities that significantly increases. Wider lines represent more significant age differences with smaller p-value.

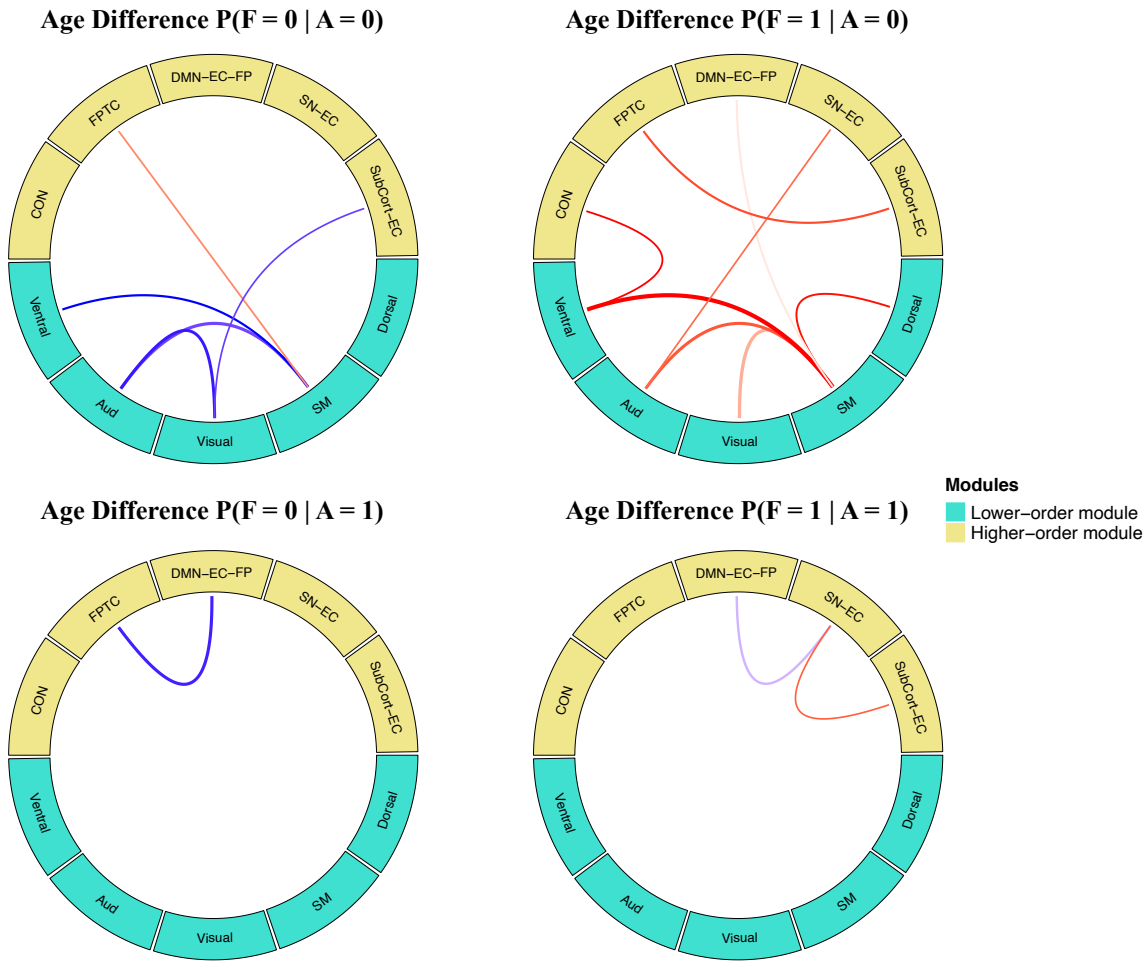

Figure S10: MMM estimated age difference in probabilities of latent functional connection state conditional on the latent structural state based on Power's module system. (A): A graphical illustration of significant age differences (older vs younger) for no FC (F=0) and positive FC (F=1) states, conditioning on no SC (A=0). (B): A graphical illustration of significant age differences (older vs younger) for no FC (F=0) and positive FC (F=1) states, conditioning on the presence of SC (A=1). Results show that most of the age-related changes in FC are found at those edges without direct SC (i.e. A = 0) while very limited age differences in FC are found for edges with SC (i.e. A = 1).

8 Results from a comparison modeling of the PNC data

We conducted additional analyses to compare the MMM results with those derived from a comparison model. The comparison method modeled SC and FC separately, used a simpler method to estimate SC and FC networks, and conducted edge-wise analysis to assess between-group differences.

Specifically, for the comparison model, we first merged information across nodes in the each module to generate a smaller network. Here, FC and SC measures between node pairs were averaged by modules to derive summarized FC and SC measures for modules and module pairs in the smaller network. Next, we modeled the SC and FC measures with mixture Gaussian distributions and then estimate the latent states for the binary SC network and the tri-state FC network based on the posterior probabilities of the fitted mixture distributions. Then we modeled the edge-wise SC and FC latent states in terms of age group and gender using the logistic regression model and the multinomial logistic model, respectively, to assess between group differences.

Figure S11 here presents the estimated difference in structural connection probabilities between the older and younger age group in female based on the comparison model (the results for male were similar and hence omitted here and later). As with the MMM model, the comparison model showed increase in white fiber structural connections with the increase in age. However, the comparison model only revealed the increase between a small number of modules, while MMM shows a general increase of structural connections across the brain. Furthermore, unlike MMM, the comparison model was not able to provide the important finding that there is more significant white fiber connectivity growth in cognitive networks as compared to sensory and motor networks.

For FC analyses, the results from MMM showed the probability of no FC (F=0) signifi-

cantly decreases between many modules with age increasing and the probability of positive FC (F=1) generally increases across the networks, indicating the brain gets more functionally organized with neurodevelopment. In Figure S12 below, we display the corresponding FC results for the age effects based on the comparison model. The comparison model was not able to find significant changes in the probability of no FC state (F=0) in the brain networks with age increases, and did not show a general increase in the positive FC (F=1) either. Therefore, unlike MMM, the comparison model was not able to yield results demonstrating the strengthening of brain functional organizations with neurodevelopment. Furthermore, since the comparison method modeled the SC and FC separately, it did not have the ability to investigate the FC changes conditional on the anatomical structural connection status. Hence, the comparison method was unable to provide the findings as the MMM to reveal that functional connection development during this period mainly happen at edges that do not have strong direct structural connections.

In summary, the results indicate that the proposed MMM is more powerful than the comparison model to yield neurobiologically meaningful new insights in neurodevelopment during adolescence.

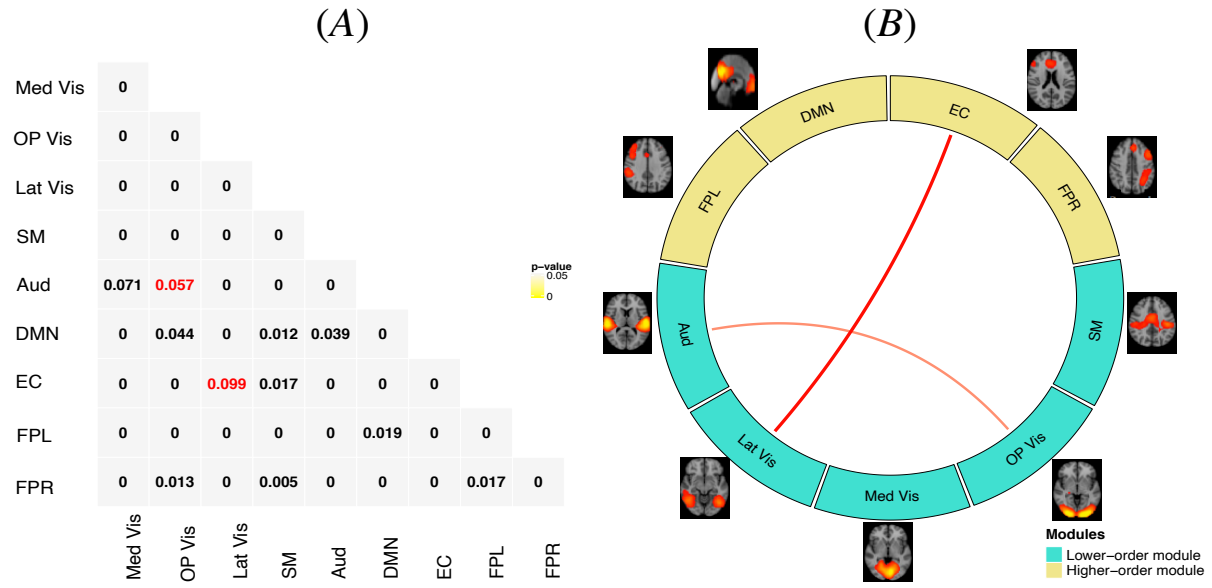

Figure S11: Estimated difference in structural connection (SC) probabilities between the older (16-21) and younger (8-15) age group based on the comparison model. (A): the numeric values are the estimated age difference (older vs. younger) in the SC probabilities. We highlight in yellow the age differences that are significant at the alpha=0.05 level, where the color of the numeric values indicates the direction of the difference (red=significant positive difference; blue=significant negative difference). (B): A graphical illustration of the significant differences in SC across brain networks presented in (A). Turquoise modules represent high-order cognitive networks and yellow modules represent primary sensory and motor networks. The red lines show significantly increased SC with age, with the wider lines representing more significant age difference with smaller p-values.The comparison model only revealed the increase in SC between a small number of modules.

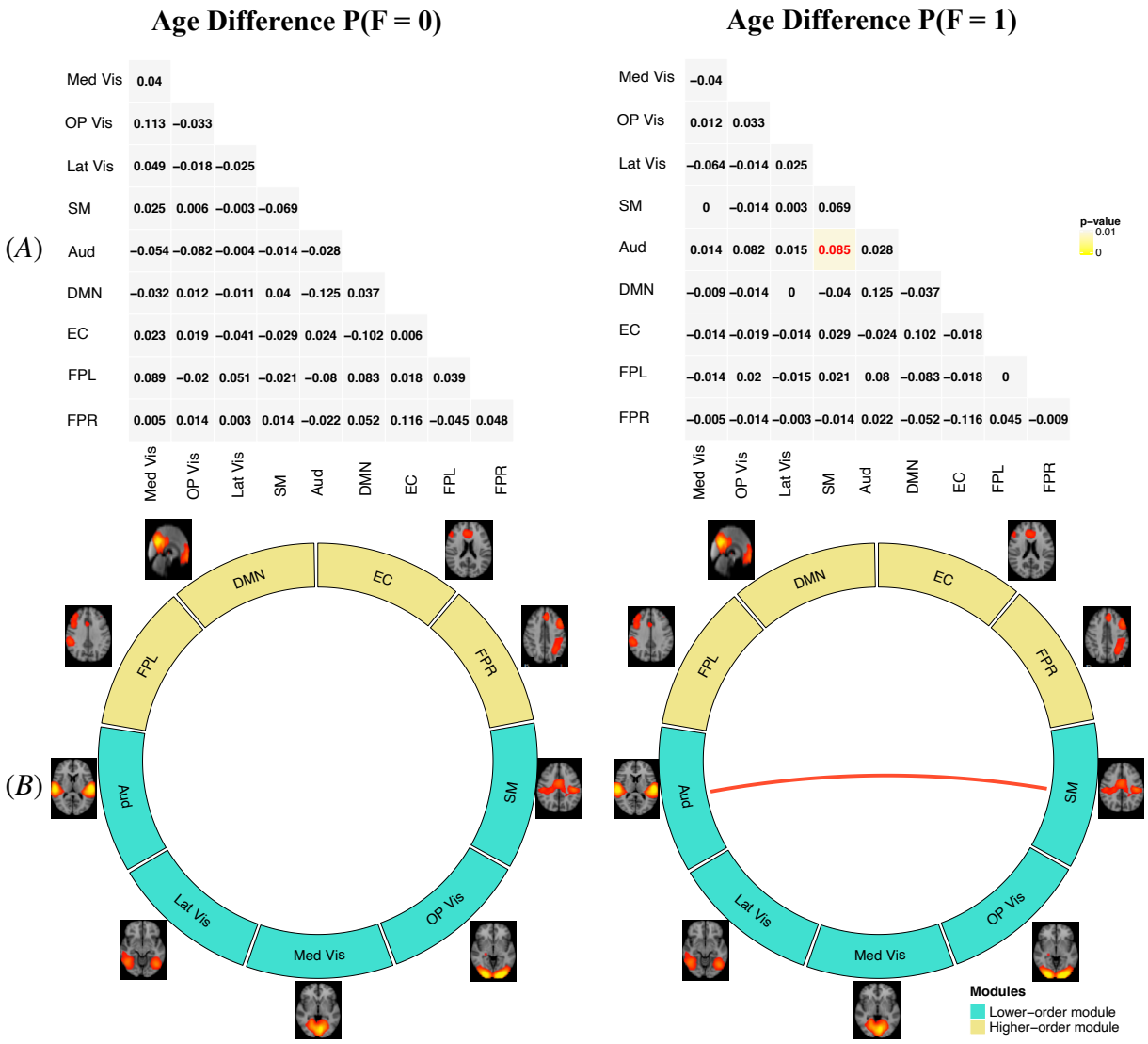

Figure S12: Estimated difference in probabilities of functional connection state between the older (16-21) and younger (8-15) age group based on the comparison model. (A): the estimated age difference (older vs. younger) for the probabilities of no FC (F=0) and positive FC (F=1). We highlight in yellow the age differences that are significant at the alpha=0.05 level, where the color of the numerical values indicates the direction of the difference (red=significant positive difference; blue=significant negative difference). (B): A graphical illustration of the significant age differences in functional connections across brain networks presented in (A). Turquoise circles represent high-order cognitive networks and yellow circles represent primary sensory and motor networks. The blue lines show probabilities that significantly decreases with age increase and the red lines show probabilities that significantly increases. Wider lines represent more significant age differences with smaller p-value. The comparison model was not able to find any significant changes in the probability of no FC state (F=0) in the brain networks with age increases, and didn't reveal a general increase in the positive FC (F=1).

References

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L. et al. (2011), 'Functional network organization of the human brain', *Neuron* **72**(4), 665–678.