Bayesian Joint Modeling of Multiple Brain Functional Networks - Supplementary Materials

1 Steps in the Posterior Computation

Let $\delta_{g,kl}$ denote the edge inclusion indicators, with $\delta_{g,kl} = 1$ if the edge (k, l) is included in $\mathcal{G}_g$ and $\delta_{g,kl} = 0$ otherwise. The posterior computation involves the following steps.
1. Update the scale parameters under the posterior distributions:

$$\begin{aligned}\pi\big(\tau_{g,kl}^{-1}|\delta_{g,kl}=1,a_\tau,b_\tau\big) &= Ga(a_\tau+\frac{1}{2},\frac{1}{b_\tau+0.5\omega_{g,kl}^2}),\\ \pi\big((\tau_{g,kl}^{*})^{-1}|\delta_{g,kl}=0,\lambda_0\big) &= Inv-Gaussian(\sqrt{\frac{\lambda_0^2}{\omega_{g,kl}^2}},\lambda_0^2).\end{aligned}$$

2. Update $\boldsymbol{\Omega}$ as follows. First, partition $\boldsymbol{\Omega}$ and $\boldsymbol{S}$ as:

$$\boldsymbol{\Omega}_g=\begin{pmatrix}\boldsymbol{\Omega}_{g,11} & \boldsymbol{\omega}_{g,12}\\ \boldsymbol{\omega}_{g,12}^T & \omega_{g,22}\end{pmatrix},\quad \boldsymbol{S}_g=\begin{pmatrix}\boldsymbol{S}_{g,11} & \boldsymbol{s}_{g,12}\\ \boldsymbol{s}_{g,12}' & s_{g,22}\end{pmatrix}.$$

Then, perform the change of variable $(\boldsymbol{\omega}_{g,12},\omega_{g,22})\rightarrow(\boldsymbol{\beta}=\boldsymbol{\omega}_{g,12},\gamma=\omega_{g,22}-\boldsymbol{\omega}_{g,12}^T\boldsymbol{\Omega}_{g,11}^{-1}\boldsymbol{\omega}_{g,12})$. Subsequently, sample $\pi(\gamma,\boldsymbol{\beta})\sim Ga(\frac{n}{2}+1,\frac{s_{g,22}+\alpha}{2})N(\boldsymbol{C}\boldsymbol{S}_{g,21},\boldsymbol{C})$, where $\boldsymbol{C}=[(s_{g,22}+\alpha)\boldsymbol{\Omega}_{11}^{-1}+\boldsymbol{D}_\tau^{-1}]^{-1}$.
3. Update the edge inclusion indicators as follows. First compute prior inclusion probabilities as $\theta_{g,kl}=\frac{\exp\{\eta_{0,kl}+\eta_{g,kl}\}}{1+\exp\{\eta_{0,kl}+\eta_{g,kl}\}}$. Then, for $l=1,2,\ldots,p-1$ and $k=l+1,l+2,\ldots,p$, sample the edge inclusion indicators from a Bernoulli distribution having posterior inclusion probabilities $P(\delta_{g,kl}=1|\cdot)=\frac{\pi_{g,kl1}^*\theta_{g,kl}}{\pi_{g,kl1}^*\theta_{g,kl}+\pi_{g,kl0}^*(1-\theta_{g,kl})}$, where

- $\pi_{g,kl1}^*=f_{Ga}(\tau_{g,kl}^{-1};a_\tau,b_\tau)(\tau_{g,kl})^{-1/2}\sqrt{\sigma_{21}^2}\exp\{0.5\frac{\mu_1^2}{\sigma_{21}^2}\}$
- $\pi_{g,kl0}^*=\frac{\lambda_0^2}{2}\exp\{-\frac{1}{2}\lambda_0^2\tau_{g,kl0}\}(\tau_{g,kl0})^{-1/2}\sqrt{\sigma_{20}^2}\exp\{\frac{1}{2}\frac{\mu_0^2}{\sigma_{20}^2}\}$
- $\sigma_{2m}^2=[(\boldsymbol{S}_g(l,l)+\alpha)\boldsymbol{\Omega}_{g,ll}^{-1}(k-1,k-1)+\tau_{g,klm}^{-1}]^{-1}$,

- $\mu_m = -(\sigma^2_{2m})[\boldsymbol{\Omega}^{-1}_{g,ll}(k-1,\cdot)\boldsymbol{\omega}_{gl}(-k,-l) + \boldsymbol{S}_g(l,k)]$, $m = 0, 1$,

where $f_{Ga}(\cdot)$ denotes the probability density function of a Gamma distribution with $\boldsymbol{\Omega}_{g,ll}$ being equal to $\boldsymbol{\Omega}_g$ after excluding the lth row and column, $\boldsymbol{\omega}_{gl}(-k,-l)$ denotes the lth row of $\boldsymbol{\Omega}_g$ after excluding the kth and lth elements, and $\boldsymbol{\Omega}^{-1}_{g,LL}(K-1,\cdot)$ denotes the $(k-1)$th row of $\boldsymbol{\Omega}^{-1}_{g,ll}(K-1,\cdot)$ with the diagonal term excluded.
4. Update shared and differential effects as follows. First, compute the prior inclusion probability for edge (k,l) as

$$\theta_{g,kl} = \frac{\exp\{-(u_{g,kl} - \theta^n_{g,kl})\}}{1+\exp\{-(u_{g,kl} - \theta^n_{g,kl})\}} = \int_0^\infty \frac{\exp\{-(u_{g,kl} - \theta^n_{g,kl})\}}{(1+\exp\{-(u_{g,kl} - \theta^n_{g,kl})\})^2} du_{g,kl}$$
$$\approx \int_0^\infty \int_0^\infty \mathbf{t}(u_{g,kl}; \theta^n_{g,kl}, \frac{\pi^2(\phi-2)}{3}\sigma^2_{\phi,g,kl})\pi(\sigma^2_{\phi,g,kl}, \frac{\phi}{2}, \frac{\phi}{2}) d\sigma^2_\phi du_{g,kl},$$

under the approximation to the logistic function where $\theta^n_{g,kl} = \eta_{0,kl} + \eta_{g,kl}$, and $\pi(\sigma^2_\phi)$ follows a inverse-Gamma distribution with parameters $\phi/2$ and $\phi/2$, with $\phi = 7.3$. In order to sample the common and differential effects, we perform the data augmentation described earlier by introducing latent variables $u_{g,kl} \sim N(\eta_{0,kl} + \eta_{g,kl}, \sigma^2_{\phi,g,kl}\frac{\pi^2(\phi-2)}{3})$. In particular the following sampling steps are performed sequentially:

- Sample $u_{g,kl} \sim N(\eta_{0,kl} + \sum_{g=1}^G \eta_{g,kl} 1_{g=m}, \sigma^2_{\phi,g,kl}\frac{\pi^2(\phi-2)}{3})$ with $u_{g,kl}$ truncated to 1 if $\delta_{g,kl} = 1$ or to zero if $\delta_{g,kl} = 0$.

- Denote the prior variance for the hth cluster by $\sigma^2_h, h = 1, \ldots, H$. Denote the cluster membership of the klth edge weight by c_{kl}, where $c_{kl} \in \{1, \ldots H\}$. The cluster memberships and regression coefficients are updated using a slice sampling technique (Walker, 2007). The update steps are as follows:

  1. Sample $u^*_{g,kl} \sim Uniform(0, \nu_{g,kl})$
  2. Sample the stick breaking weights as: $v_{g,h} \sim Beta(1 + n_{g,h}, M + \sum_{h'>h} n_{g,h'})$, followed by $\nu_{g,h} = v_{g,h}\prod_{h'<h}(1-v_{g,h'})$, where $n_{g,h}$ is the number of edges assigned to cluster h in platform g (Sethuraman, 1994).
  3. Sample the cluster membership indicators:
     (a) For each edge, calculate the atoms available to it as $A_{kl} = \{h : u^*_{g,kl} < \nu_{g,h}, h = 1, \ldots, H, g = 0, \ldots, G\}$
     (b) For each edge, calculate the vector of likelihood values $\pi^*_{kl} = [\pi^*_{1,kl} I_{1\in A_{kl}}, \ldots, \pi^*_{H,kl} I_{H\in A_{kl}}]$ where $\pi^*_{h,kl} = \prod_{g=1}^G N(u_{g,kl}; \eta_{g,h}, \sigma^2_{\phi,g,kl}\frac{\pi^2(\phi-2)}{3})$

(c) Calculate the normalized likelihoods $\pi_{h',kl} = \pi^*_{h',kl} / \sum_{h=1}^{H} \pi^*_{h,kl}$

(d) Draw $r \sim Uniform(0,1)$ and update $c_{kl} = h'$ where h' is the smallest h' s.t. $r < \sum_{h=1}^{h'} \pi_{h,kl}$

4. Draw the new values for the η terms conditioned on the updated cluster memberships, using a Gaussian posterior distribution.

- Update $\sigma_\eta^{-2} \sim \pi(\sigma_\eta^{-2}|\cdot) \equiv Ga(a_\sigma + \frac{0.5(G+1)p(p-1)}{2}, b_\sigma + 0.5\sum_{m=0}^{G}\sum_{l=1}^{p}\sum_{k<l}\eta^2_{m,kl})$.

- Similarly, draw $\sigma^{-2}_{\phi,g,kl} \sim \pi(\sigma^2_{\phi,g,kl} \mid \cdot) \equiv Ga(\frac{\phi+1}{2}, \frac{\phi+(u_{g,kl}-\theta^n_{g,kl})^2}{2(\pi^2(\phi-2)/3)}), g = 0, 1, \ldots, G, k \neq l, k, l = 1, \ldots, p$.

2 Sensitivity to Choice of the Precision Parameter

The hyperparameters a_m and b_m control the DP precision parameter, M. To examine the sensitivity of our performance to changes in the hyperparameter a_m, we ran an additional 25 simulation replicates using the data generated from Erdos-Renyi random graphs with 100 nodes as in Section 4 of the main manuscript. We varied a_m from 0.1 to 3, and tracked the resulting AUC for edge detection, L1 Error, and TPR/FPR for differential edge detection. Figure 1 displays the results of these simulations. We observe that though varying the hyperparameter has some effect on the number of clusters, the final results of the procedure are similar. A value of $a_m = 1$ does seem to yield slightly higher AUC values, and thus we recommend this as the default selection.

3 40 Node Simulation Results

Table 1 displays the simulation results for the 40 node simulations, and Figure 2 displays the ROC curves.

4 Simulation Boxplots

Figures 3 and 4 display the results of the small-world network and scale-free network simulations respectively.

Table 1: 40 node simulation results comparing BJNL, JGL, GL, and HS. Text in bold indicates a method was better than both competing methods as assessed through Wilcoxon signed rank tests at $\alpha = 0.05$.

	AUC				L_1 Error $\times$ 100			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.98 (0.01)**	0.91 (0.03)	0.91 (0.03)	0.84 (0.05)	**0.08 (0.02)**	1.02 (0.15)	1.91 (0.26)	4.66 (0.62)
med	**0.99 (0.01)**	0.91 (0.04)	0.91 (0.04)	0.78 (0.03)	**0.09 (0.02)**	1.02 (0.17)	1.90 (0.27)	11.15 (0.58)
high	**0.98 (0.02)**	0.91 (0.03)	0.91 (0.03)	0.83 (0.06)	**0.09 (0.02)**	1.02 (0.18)	1.90 (0.29)	11.35 (0.78)
Small World								
low	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.81 (0.02)	**0.50 (0.05)**	1.89 (0.37)	4.09 (0.27)	4.68 (0.62)
med	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.78 (0.03)	**0.50 (0.04)**	1.92 (0.41)	4.16 (0.33)	10.98 (0.53)
high	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.78 (0.03)	**0.49 (0.05)**	1.98 (0.41)	4.13 (0.33)	11.33 (0.67)
Scale Free								
low	**0.97 (0.01)**	0.93 (0.01)	0.79 (0.02)	0.84 (0.05)	**0.48 (0.03)**	1.95 (0.50)	4.13 (0.33)	4.59 (0.6)
med	**0.97 (0.01)**	0.92 (0.01)	0.79 (0.02)	0.81 (0.02)	**0.47 (0.02)**	2.01 (0.52)	4.18 (0.36)	11.05 (0.51)
high	**0.97 (0.01)**	0.93 (0.01)	0.79 (0.02)	0.79 (0.03)	**0.47 (0.03)**	1.97 (0.48)	4.13 (0.36)	11.21 (0.81)
	TPR				FPR			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.92 (0.09)**	0.76 (0.12)	0.78 (0.12)	0.70 (0.12)	**0.01 (0.003)**	0.23 (0.05)	0.10 (0.02)	0.03 (0.01)
med	**0.89 (0.07)**	0.75 (0.10)	0.76 (0.12)	0.69 (0.18)	**0.01 (0.003)**	0.23 (0.05)	0.10 (0.02)	0.32 (0.03)
high	**0.88 (0.06)**	0.75 (0.08)	0.77 (0.08)	0.57 (0.17)	**0.01 (0.003)**	0.24 (0.05)	0.10 (0.02)	0.28 (0.03)
Small World								
low	**0.89 (0.10)**	0.45 (0.16)	0.67 (0.16)	0.66 (0.11)	0.04 (0.01)	0.05 (0.01)	0.38 (0.02)	0.03 (0.01)
med	**0.88 (0.09)**	0.44 (0.11)	0.63 (0.12)	0.67 (0.11)	0.04 (0.01)	0.05 (0.01)	0.38 (0.02)	0.32 (0.03)
high	**0.85 (0.07)**	0.46 (0.09)	0.62 (0.10)	0.60 (0.13)	0.04 (0.01)	0.05 (0.01)	0.39 (0.02)	0.28 (0.03)
Scale Free								
low	**0.89 (0.09)**	0.44 (0.13)	0.66 (0.17)	0.68 (0.09)	0.04 (0.01)	0.04 (0.01)	0.35 (0.02)	0.02 (0.01)
med	**0.88 (0.07)**	0.43 (0.11)	0.65 (0.11)	0.64 (0.08)	0.04 (0.01)	0.05 (0.01)	0.35 (0.02)	0.33 (0.03)
high	**0.88 (0.07)**	0.42 (0.10)	0.62 (0.09)	0.58 (0.09)	0.04 (0.01)	0.04 (0.01)	0.35 (0.02)	0.28 (0.03)

5 fMRI Data Preprocessing

Image preprocessing was performed using the AFNI software package (Cox, 1996) and included slice-timing and motion correction, warping of brain volumes to standard Talairach space, Gaussian spatial blurring (6mm FWHM), and signal percent scaling.

In order to minimize the effect of spurious sources of temporal correlation in the fMRI time series, which could have negatively affected our network analysis, we performed the following additional pre-processing steps for each subject. First, a regression model was fit using a maximum likelihood approach with a baseline portion accounting for noise (a second-order Legendre polynomial, modeling slow signal drifts from biological or scanner-related confounds, plus 6 head-motion parameters) and a set of event-related regressors representing the expected BOLD response for the following classes of stimuli: (a) correctly-responded congruent trials, (b) correctly-responded incongruent trials, (c) commission error trials corresponding to incorrect responses, (d) omission error trials corresponding to omitted responses. Each event was modeled as a mini boxcar starting at the stimulus onset and ending at the time of the subject's response, followed by convolution with a gamma

function accounting for the BOLD hemodynamic properties; if a response was not issued within the allotted 2 second window, the trial was marked as an omission error and the duration of the corresponding boxcar was set at 2 seconds. The residual time-series from the regression analysis were collected, and task-related regressors, multiplied by their corresponding regression coefficients, were added back to them. This procedure aimed to remove the confounding components of the BOLD time series (slow scanner-related drifts, task-unrelated physiological variance, head motion) without also removing the component of interest, that is, task-related variance.

As a final preprocessing step prior to the analysis, we prewhitened the time series for each node. We fit an ARMA(1, 1) model at each node for each subject's data and used the residuals from these models as the corresponding node time series for the analysis. Boxplots of the autocorrelation of the resulting prewhitened data at lags $1-6$ are provided in Figure 5 for both the passive fixation and task conditions. The values at all lags are close to zero, which confirms that the temporal autocorrelation has been largely removed.

6 Graph Metric Description

Efficiency measures how effectively information is transmitted from node-to-node in a network. Global efficiency measures information transmission across the entire graph and is calculated by taking the average across all ROIs of the inverse shortest path lengths between ROIs. Thus, large values of global efficiency indicate that, on average, the number of steps required to transmit information from one node to another is small. Local efficiency measures information transmission between an ROI and its neighbors and is calculated for each ROI by taking the average of the inverse shortest path lengths between ROIs in the relevant neighborhood, where the relevant neighborhood is the collection of ROIs with a connection to the selected ROI. The clustering coefficient measures the interconnectedness of the graph and is calculated for each ROI by examining how many of its neighbors are also neighbors to each other. Finally, characteristic path length is the average across ROIs of the shortest path length in the networks, with smaller values indicating a more efficient network.

7 Stationarity of MCMC Sampler

We performed Dickey-Fuller tests of stationarity to assess the convergence of the MCMC sampler using the adf.test function in R (Dickey and Fuller, 1979). We tested the time series of MCMC samples for each of the off-diagonal elements of the precison matrix. Rejection of

the null hypothesis indicates that the time series of MCMC samples is stationary, suggesting that the mean and variance have stabilized. All obtained p-values were less than 0.01, indicating that the MCMC sampler converged. Figure 6 displays the corresponding test statistics.

8 Credible Interval Widths for the Stroop Task Data

One advantage of the Bayesian approach is that it provides us with a measure of posterior uncertainty. Figure 7 displays boxplots of the credible interval widths for the partial correlations in the analysis of exertion versus relaxed task performance. We note that there is a clear distinction between the credible interval widths for off-diagonal elements corresponding to present edges versus off-diagonal elements corresponding to absent edges. Specifically, the credible intervals corresponding to absent edges are quite narrow near zero. This is a desirable behavior, as it indicates that the BJNL is able to clearly differentiate between present and absent edges.

9 Stroop Task Data Goodness of Fit

To verify that the BJNL provides a good fit to the Stroop task data, we conduct a goodness of fit analysis. We test the validity of the distributional assumptions under our proposed approach using a chi-square goodness of fit test. We considered each observation from both conditions, $\boldsymbol{Y}_{irt}, t = 1, \ldots T,\ r = 1, 2;\ i = 1, \ldots, N$, and calculated the quadratic form $\boldsymbol{Y}'_{irt}\boldsymbol{\Omega}_r\boldsymbol{Y}_{irt}$ for each subject, which should follow a chi-squared distribution with 90 degrees of freedom if the normality assumption holds. Under the chi-squared test, only a small proportion of tests were significant for each condition (5.0% for exertion, and 6.6% for relaxed), suggesting that the fMRI data satisfies the inherent modeling assumptions.

10 Stroop Task Data Circle Plots

Figure 8 displays circle plots of the sum of the strengths of the significant connections between modules for the Stroop task data analysis. The first row of plots displays the circle plots for the analysis of TASK vs. REST, and the second row displays the plots for the analysis of EXR vs. RLX.

Table 2: P-values for testing differences between TASK and REST in local efficiency and clustering coefficient for the nodes identified in the Khachouf (2017) study. Our analysis revealed larger clustering coefficients for REST in all implicated regions, and larger local efficiencies for REST in 18 of the 20 regions

AAL Region	Metric Center (t-test)						Metric Distribution (ks-test)	
	Local Efficiency			Clustering Coefficient			Local Efficiency	Clustering Coefficient
	BJNL	JGL	GL	BJNL	JGL	GL	BJNL	BJNL
Cingulate_Ant_L	**0.01**	0.27	0.52	**<0.01**	0.40	0.53	**0.03**	**<0.01**
Cingulate_Ant_R	**<0.01**	0.75	0.61	**<0.01**	0.87	0.78	**<0.01**	**<0.01**
Cingulate_Mid_L	**0.01**	**<0.01**	0.94	**<0.01**	**<0.01**	0.70	**0.03**	**<0.01**
Cingulate_Mid_R	**<0.01**	0.50	0.58	**<0.01**	0.54	0.45	**<0.01**	**<0.01**
Cingulate_Post_L	**<0.01**	**0.01**	**0.04**	**<0.01**	**0.04**	**0.03**	**0.01**	**<0.01**
Cingulate_Post_R	**<0.01**	0.55	0.12	**<0.01**	0.26	0.08	**<0.01**	**<0.01**
Supp_Motor_Area_L	**<0.01**	**<0.01**	0.99	**<0.01**	**<0.01**	0.85	**<0.01**	**<0.01**
Supp_Motor_Area_R	**<0.01**	0.35	0.94	**<0.01**	0.73	0.86	**<0.01**	**<0.01**
Insula_L	**<0.01**	0.13	0.34	**<0.01**	**0.02**	0.34	0.07	**0.04**
Insula_R	**0.03**	0.31	0.06	**<0.01**	0.06	0.11	**0.02**	**<0.01**
Parietal_Sup_L	**<0.01**	**<0.01**	0.52	**<0.01**	**0.03**	0.71	**<0.01**	**<0.01**
Parietal_Sup_R	**<0.01**	**0.02**	0.64	**<0.01**	**0.05**	0.38	**<0.01**	**<0.01**
Parietal_Inf_L	**<0.01**	0.74	0.64	**<0.01**	0.64	0.69	**0.02**	**<0.01**
Parietal_Inf_R	**<0.01**	**<0.01**	0.14	**<0.01**	0.09	0.37	**<0.01**	**<0.01**
Thalamus_L	**<0.01**	0.70	0.68	**<0.01**	0.54	0.81	**<0.01**	**<0.01**
Thalamus_R	**<0.01**	0.64	**<0.01**	**<0.01**	0.69	**<0.01**	**<0.01**	**<0.01**
Hippocampus_L	**<0.01**	0.28	0.58	**<0.01**	0.71	0.61	**<0.01**	**<0.01**
Hippocampus_R	**0.01**	0.83	0.86	**<0.01**	0.63	0.72	0.13	**<0.01**
Paracentral_Lobule_L	**<0.01**	0.98	0.89	**<0.01**	0.36	0.57	**<0.01**	**<0.01**
Paracentral_Lobule_R	**<0.01**	0.38	0.60	**<0.01**	0.82	0.46	**<0.01**	**<0.01**

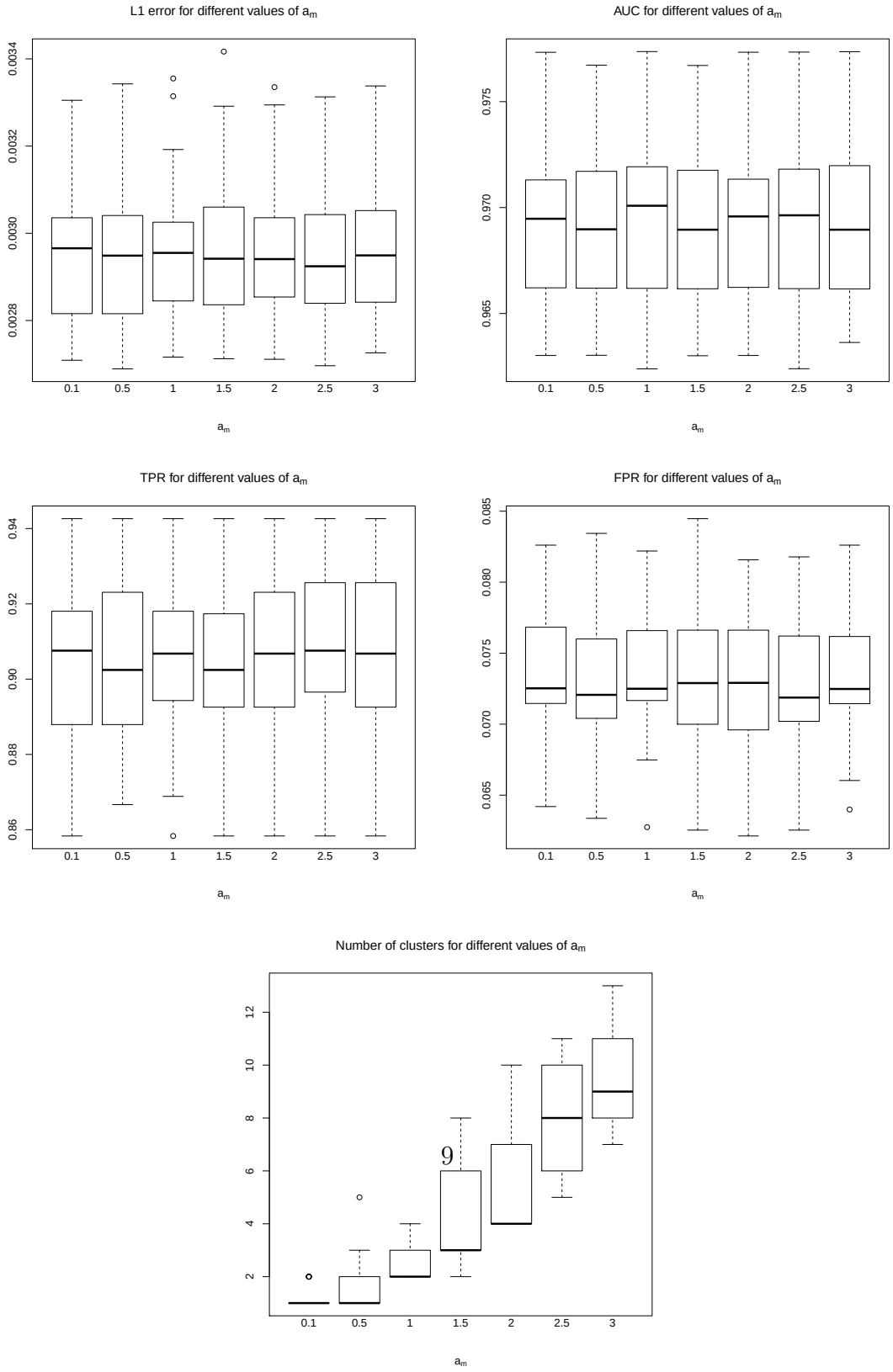
11 Stroop Task Additional Node Analyses

Tables 2 and 3 contain the p-values for testing differences in the local efficiency and the clustering coefficient for specific nodes in the analyses of TASK vs. REST and EXR vs. RLX, respectively.

Table 3: P-values for testing differences between EXR and RLX in local efficiency and clustering coefficient for the nodes identified in the Khachouf (2017) study. Significant differences were found in terms of mean local efficiency in the right inferior occipital node and the left caudate, and borderline significant results were found for the left superior parietal cortex and left insula. Similarly, we found significant differences in the mean and posterior distribution of the clustering coefficient for the right inferior occipital node. We identified borderline significant differences in the clustering coefficient distribution and center for the left anterior cingulate, the mean of the clustering coefficient in the left insula, and the mean and distribution of the clustering coefficient in the left caudate.

AAL Region	Metric Center (t-test)						Metric Distribution (ks-test)	
	Local Efficiency			Clustering Coefficient			Local Efficiency	Clustering Coefficient
	BJNL	JGL	GL	BJNL	JGL	GL	BJNL	BJNL
Parietal_Sup_L	0.10	0.37	0.97	0.33	0.36	0.70	0.36	0.42
Parietal_Sup_R	0.88	0.66	0.48	0.78	0.64	0.31	0.83	0.74
Supp_Motor_Area_L	0.88	0.92	0.14	0.53	0.90	0.21	0.80	0.36
Supp_Motor_Area_R	0.94	0.77	0.73	0.50	0.81	0.71	0.97	0.86
Cingulate_Ant_L	0.47	**0.03**	0.64	0.07	**0.03**	0.62	0.28	0.07
Cingulate_Ant_R	0.19	**<0.01**	0.13	0.15	**<0.01**	0.26	0.52	0.26
Cingulate_Mid_L	0.89	0.34	0.87	0.57	0.38	0.53	0.24	0.67
Cingulate_Mid_R	1.00	**0.02**	0.82	0.72	0.07	0.81	0.39	0.24
Insula_L	0.07	0.21	0.82	0.08	0.21	0.65	0.24	0.15
Insula_R	0.81	0.21	0.10	0.72	0.21	**0.04**	0.97	0.13
Caudate_L	**0.05**	0.35	0.84	0.09	0.34	0.39	0.24	0.07
Caudate_R	0.20	0.49	0.89	0.71	0.48	0.81	0.27	0.76
Occipital_Sup_L	0.31	**0.01**	0.66	0.43	**0.05**	0.63	0.57	0.69
Occipital_Sup_R	0.76	1.00	0.69	0.75	**<0.01**	0.77	0.71	0.93
Occipital_Mid_L	0.22	1.00	0.39	0.86	**<0.01**	0.56	0.51	0.80
Occipital_Mid_R	0.82	1.00	0.56	0.86	**<0.01**	0.46	0.91	0.93
Occipital_Inf_L	0.64	1.00	0.73	0.80	**<0.01**	0.63	0.57	0.56
Occipital_Inf_R	**0.01**	0.35	0.13	**0.03**	0.35	0.14	**0.01**	**0.04**
Precuneus_L	0.27	0.56	0.30	0.19	0.53	0.51	0.15	0.36
Precuneus_R	0.27	0.47	0.14	0.84	0.46	0.23	0.43	0.97
Paracentral_Lobule_L	0.70	0.54	0.93	0.68	0.56	0.98	0.54	0.21
Paracentral_Lobule_R	0.64	0.25	**0.02**	0.61	0.24	**0.03**	0.35	0.39
Fusiform_L	0.38	**0.01**	0.11	0.39	**0.01**	0.13	0.31	0.52

Figure 1: Performance on 25 replicates of the Erdos-Renyi simulations with 100 nodes for different values of a_m. The median number of clusters across all MCMC iterations is reported as the number of clusters.



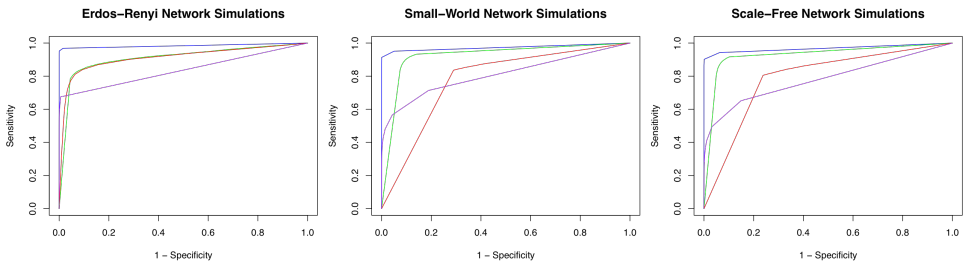


Figure 2: ROC curves for edge detection for the 40 node simulations. The blue, green, red, and purple solid lines correspond to BJNL, JGL, GL, and HS respectively.

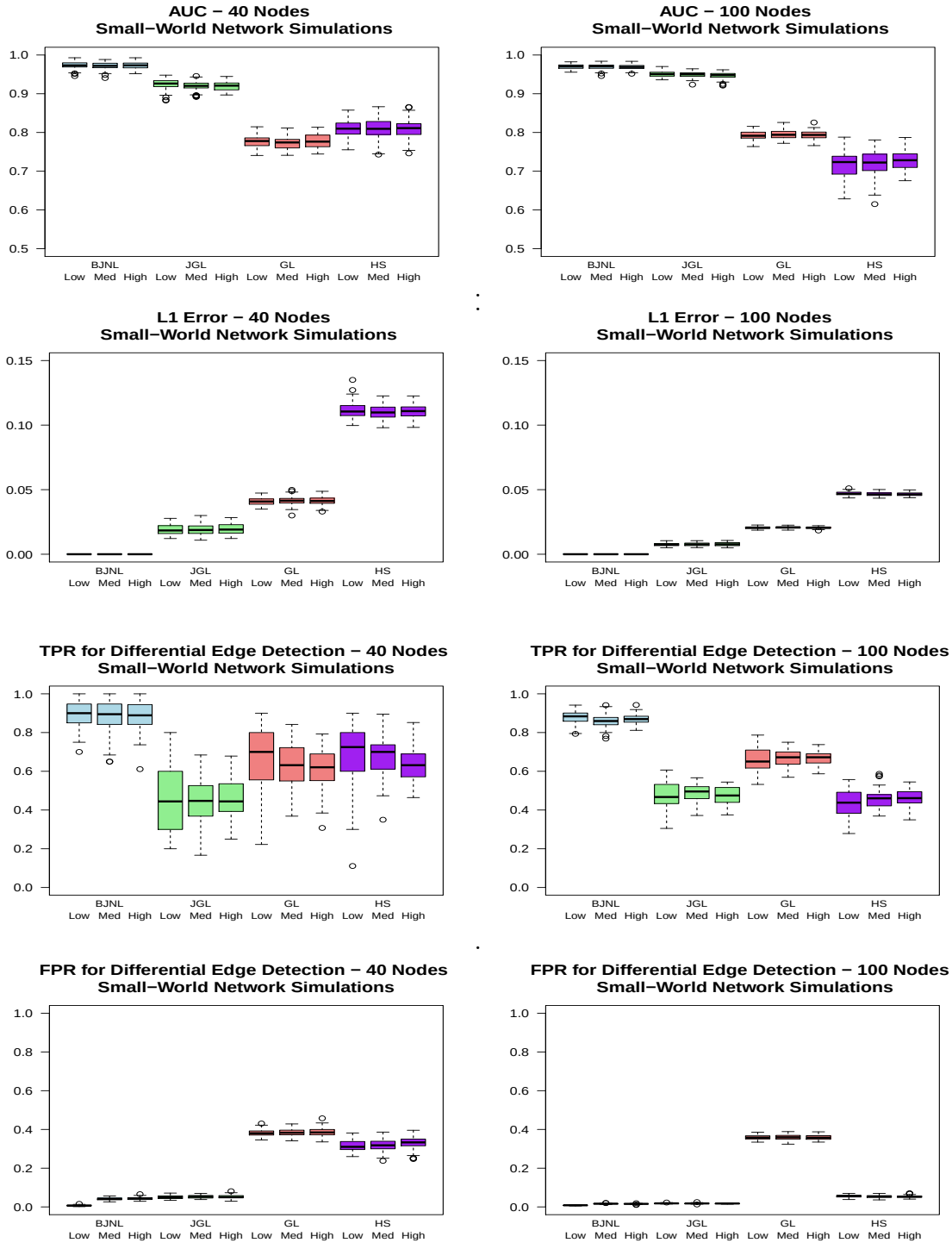


Figure 3: Box plots of the AUC, L_1 Error, and TPR/FPR for differential edge detection for the Small-World network simulations. Within each color, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

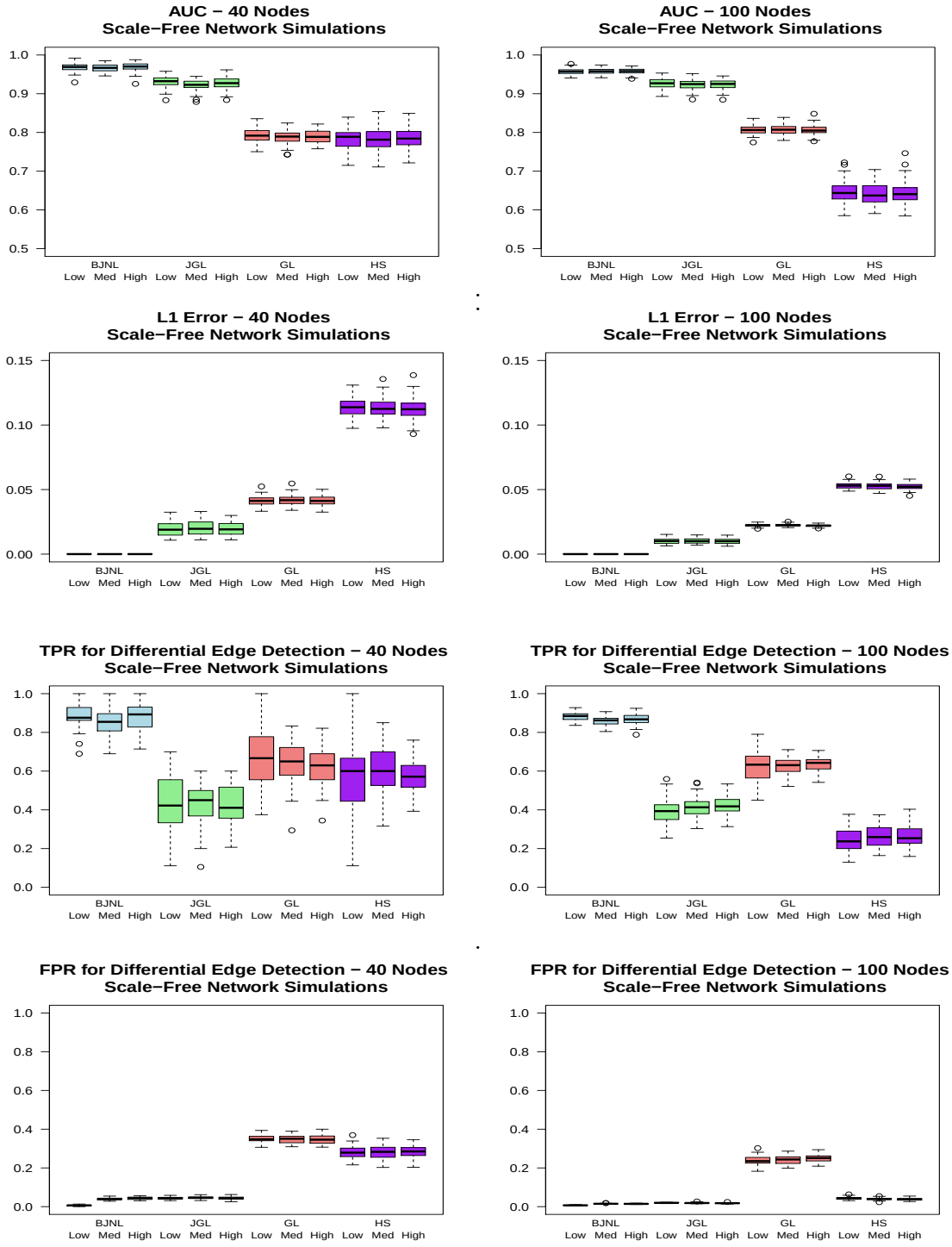


Figure 4: Box plots of the AUC, L_1 Error, and TPR/FPR for differential edge detection for the Scale-Free network simulations. Within each color, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

Figure 5: Boxplots of the mean ACF across subjects for the passive fixation condition (left) and the task condition (right).

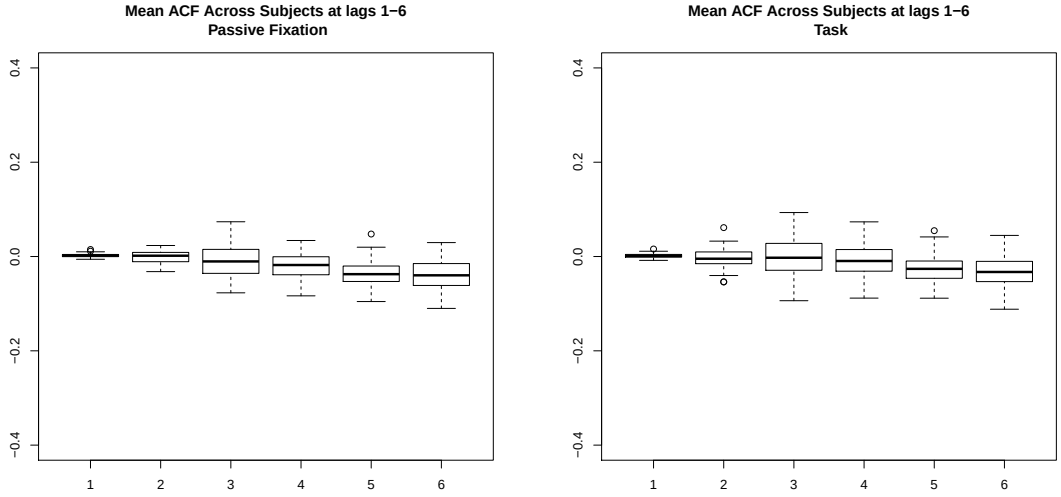


Figure 6: Histogram of the Dickey-Fuller test statistic for the off-diagonal precision elements (task state).

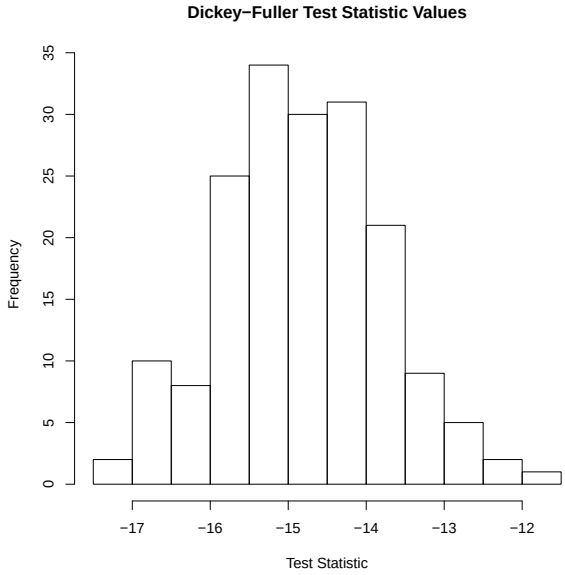
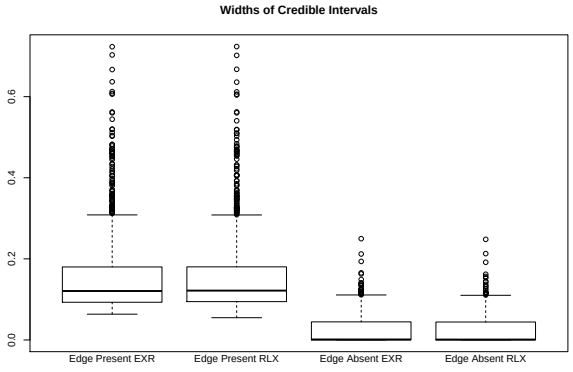


Figure 7: Widths of credible intervals for partial correlations from the analysis of exertion (EXR) vs. relaxed (RLX) task performance.



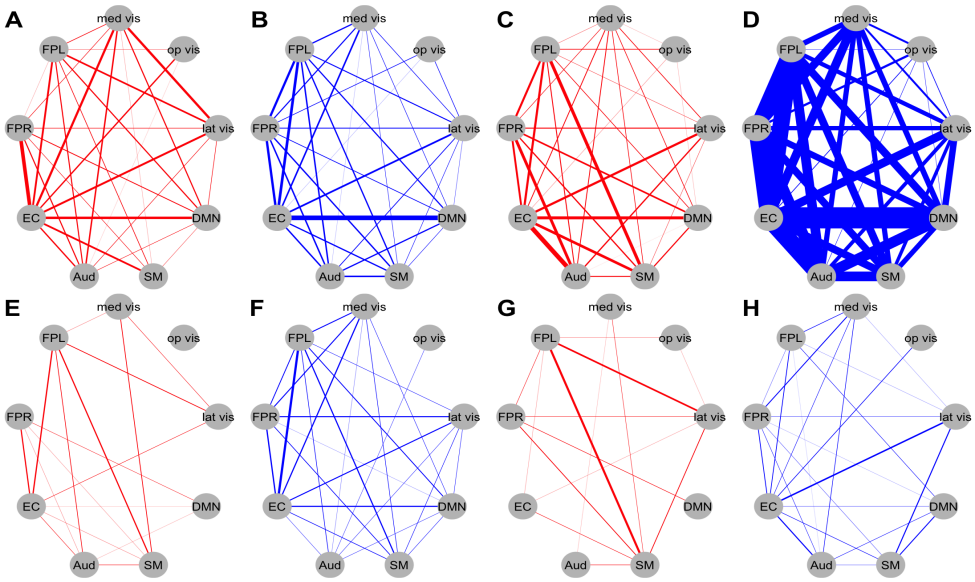


Figure 8: Circle plots of the difference in partial correlations for the analyses of TASK vs. REST and maximum exertion (EXR) vs. relaxed task performance (RLX). Line width indicates the strength of the sum of the significant partial correlations across all network-to-network connections obtained by adding the strengths of all significant edges. For the graphs with red (blue) edges, this is the sum of the significant positive (negative) partial correlations between two functional modules. Plots A and B correspond to TASK edges stronger than REST, C and D to REST edges stronger than TASK, E and F to EXR edges stronger than RLX, and G and H to RLX edges stronger than EXR.