Biometrics 0, 1–7 DOI: 000
0000 0000

Supporting Information for "Bayesian interaction selection model for multi-modal neuroimaging data analysis" by Yize Zhao, Ben Wu and Jian Kang

Yize Zhao[1*], Ben Wu[2] and Jian Kang[3]

[1]Department of Biostatistics, Yale School of Public Health, New Haven, CT, US.

[2] Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, China.

[3]Department of Biostatistics, University of Michigan, Ann Arbor, MI, US.

email: yize.zhao@yale.edu

This paper has been submitted for consideration for publication in *Biometrics*

Web Appendix A. Other types of outcome

Besides continuous outcomes, we can extend the model to fit more general outcome types in biomedical practice like binary or categorical outcomes. For instance, in the ABCD study, it is also of interest to explore brain functional activities associated with the risk of attention deficit hyperactivity disorder (ADHD). When $y_i \in \{0, 1\}$, we use a logistic regression

$$\text{logit}P(y_i = 1 \mid \mathbf{s}_i, \mathbf{x}_i) \;=\; \mathbf{s}_i^T\boldsymbol{\beta}_0 + \sum_{m=1}^{M} \mathbf{x}_{im}^T\boldsymbol{\beta}_m + \sum_{m<m'} \sum_{m'=1}^{M} (\mathbf{x}_{im} \circ \mathbf{x}_{im'})^T\boldsymbol{\beta}_{<m,m'>}, \quad \text{(A.1)}$$

where the partition construction and prior settings can be directly inserted into the linear term on the right hand side. To perform the posterior inference, following the data augmentation approach by Polson et al. (2013), we introduce latent variable $\omega_i, i = 1, \ldots, n$, and define $\tilde{\mathbf{y}} = (\frac{y_1-1/2}{\omega_1}, \ldots, \frac{y_n-1/2}{\omega_n})$. The posterior inference for each unknown parameter within $\boldsymbol{\Phi}$ will keep consistent with the main MCMC algorithm by replacing $\mathbf{y}$ with $\tilde{\mathbf{y}}$, and posterior inference for latent variable ω_i will sample from a Pólya-Gamma (PG) distribution

$$[\omega_i \mid \boldsymbol{\beta}_0, \boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{s}_i, \mathbf{x}_i, \mathbf{x}_{i,inter}] \;\sim\; \text{PG}\{1, \mathbf{s}_i^T\boldsymbol{\beta}_0 + \mathbf{x}_i^T(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) + \mathbf{x}_{i,inter}^T(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter})\}. \quad \text{(A.2)}$$

Based on the posterior samples, we can directly identify and prioritize risk imaging features and quantify their contribution with respect to the odds ratio. Compared with the alternative probit regression (Albert and Chib, 1993) which involves sampling from a truncated normal distribution, the above choice is more straightforward in computation and explanation.

Web Appendix B. MCMC algorithm

We provide the detailed MCMC algorithm for the BSSS model based on Gibbs sampler. Under random initials, we iteratively update each parameter according to the following steps:

Sampling scheme for $\boldsymbol{\beta}_0$. Draw

$$[\boldsymbol{\beta}_0 \mid \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{inter}, \boldsymbol{\delta}, \boldsymbol{\delta}_{inter}, \sigma_\epsilon^2] \sim \mathrm{N}(\boldsymbol{\mu}_{\beta_0}, \Sigma_{\beta_0}),$$

with $\Sigma_{\beta_0} = (\mathbf{S}^T\mathbf{S}/\sigma_\epsilon^2 + \sigma_0^{-2}I)^{-1}$, and $\boldsymbol{\mu}_{\beta_0} = \Sigma_{\beta_0}(\mathbf{y} - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter}))\mathbf{S}/\sigma_\epsilon^2$.

Sampling scheme for $\boldsymbol{\alpha}$. Based on $\boldsymbol{\delta}$, divide $\boldsymbol{\alpha}$ into the unselected segment $\boldsymbol{\alpha}_0$ and selected one $\boldsymbol{\alpha}_1$. Then we sample $\boldsymbol{\alpha}_0 \sim \mathrm{N}(\mathbf{0}, \sigma_1^2\mathrm{I})$ and

$$[\boldsymbol{\alpha}_1 \mid \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}_{inter}, \boldsymbol{\delta}_{inter}, \sigma_1^2] \sim \mathrm{N}(\boldsymbol{\mu}_\alpha, \Sigma_\alpha),$$

with $\Sigma_\alpha = (\mathbf{X}_\delta^T\mathbf{X}_\delta/\sigma_\epsilon^2 + \sigma_1^{-2}I)^{-1}$, and $\boldsymbol{\mu}_\alpha = \Sigma_\alpha\mathbf{X}_\delta^T(\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter}))/\sigma_\epsilon^2$, where $\mathbf{X}_\delta$ includes the columns of $\mathbf{X}$ corresponding to $\delta = 1$.

Sampling scheme for $\boldsymbol{\alpha}_{inter}$. Based on $\boldsymbol{\delta}_{inter}$, divide $\boldsymbol{\alpha}_{inter}$ into the unselected segment $\boldsymbol{\alpha}_{inter,0}$ and selected one $\boldsymbol{\alpha}_{inter,1}$. Then we sample $\boldsymbol{\alpha}_{inter,0} \sim \mathrm{N}(\mathbf{0}, \sigma_2^2\mathrm{I})$ and

$$[\boldsymbol{\alpha}_{inter,1} \mid \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}_{inter}, \boldsymbol{\delta}, \sigma_1^2] \sim \mathrm{N}(\boldsymbol{\mu}_{\alpha_{inter}}, \Sigma_{\alpha_{inter}}),$$

with $\Sigma_{\alpha_{inter}} = (\mathbf{X}_{\delta_{inter}}^T\mathbf{X}_{\delta_{inter}}/\sigma_\epsilon^2 + \sigma_2^{-2}I)^{-1}$, and $\boldsymbol{\mu}_{\alpha_{inter}} = \Sigma_{\alpha_{inter}}\mathbf{X}_{\delta_{inter}}^T(\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}))/\sigma_\epsilon^2$, where $\mathbf{X}_{\delta_{inter}}$ includes the columns of $\mathbf{X}_{inter}$ corresponding to $\delta_{inter} = 1$.

Sampling scheme for $\gamma_l^{(1)}, l = 1, \ldots, m^{(1)}$. The full conditional of $\gamma_l^{(1)}$ is given by

$$\begin{aligned}&[\gamma_l^{(1)} \mid \boldsymbol{\delta}, \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{inter}] \\ &\quad \propto p_1^{\gamma_l^{(1)}}(1 - p_1)^{1-\gamma_l^{(1)}} \exp\{\frac{1}{2\sigma_\epsilon^2}||\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter})||^2\}\end{aligned}$$

Sampling scheme for $\gamma_l^{(k)}, k = 2, \ldots, K-1$. Divide $\boldsymbol{\delta}^{(k-1)}$ into two pieces by their values and update the index sets as $\Theta_t^{(k-1)} = \{l : \delta_l^{(k-1)} = t\}$, $t = 0, 1$. Correspondingly, $\Theta_t^{(k)} = \{l : a_{l,l'}^{(k)}, l' \in \Theta_t^{(k-1)}\}$, $t = 0, 1$. For $l \in \Theta_0^{(k)}$, sample $\gamma_l^{(k)} \sim \mathrm{Bern}(p_k)$. For $l \in \Theta_1^{(k)}$, sample

$$\begin{aligned}&[\gamma_l^{(k)} \mid \boldsymbol{\delta}, \boldsymbol{\delta}_{inter}, \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{inter}] \\ &\quad \propto p_k^{\gamma_l^{(k)}}(1 - p_k)^{1-\gamma_l^{(k)}} \exp\{\frac{1}{2\sigma_\epsilon^2}||\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter})||^2\}.\end{aligned}$$

Sampling scheme for $\gamma_l^{(K)}$. Update index set $\Theta_t^{(K)} = \{l : \mathcal{I}[A^{(K-2)}\boldsymbol{\delta}^{(K-1)}]_l = t\}$, $t = 0, 1$. For $l \in \Theta_0^{(K)}$, sample $\gamma_l^{(K)} \sim \text{Bern}(p_K)$. For $l \in \Theta_1^{(K)}$

$$[\gamma_l^{(K)} \mid \boldsymbol{\delta}, \boldsymbol{\delta}_{inter}, \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{inter}]$$
$$\propto p_K^{\gamma_l^{(K)}} (1 - p_K)^{1-\gamma_l^{(K)}} \exp\{\frac{1}{2\sigma_\epsilon^2} ||\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter})||^2\}.$$

Sampling scheme for σ_1^2. Draw

$$[\sigma_1^2 \mid \boldsymbol{\alpha}] \sim \text{IG}\left(a_\sigma + PM/2, b_\sigma + \frac{\boldsymbol{\alpha}^T\boldsymbol{\alpha}}{2}\right).$$

Sampling scheme for σ_2^2. Draw

$$[\sigma_2^2 \mid \boldsymbol{\alpha}_{inter}] \sim \text{IG}\left(a_\sigma + PM(M-1)/4, b_\sigma + \frac{\boldsymbol{\alpha}_{inter}^T\boldsymbol{\alpha}_{inter}}{2}\right).$$

Sampling scheme for σ_ϵ^2. Draw

$$[\sigma_\epsilon^2 \mid \boldsymbol{\delta}, \boldsymbol{\delta}_{inter}, \mathbf{y}, \mathbf{X}, \mathbf{X}_{inter}, \mathbf{S}, \boldsymbol{\beta}_0, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{inter}]$$
$$\sim \text{IG}(a_\epsilon + n/2, b_\epsilon + \frac{1}{2\sigma_\epsilon^2} ||\mathbf{y} - \mathbf{S}\beta_0 - \mathbf{X}(\boldsymbol{\delta} \circ \boldsymbol{\alpha}) - \mathbf{X}_{inter}(\boldsymbol{\delta}_{inter} \circ \boldsymbol{\alpha}_{inter})||^2).$$

Web Appendix C. Additional simulations and assessment

C.1 *Posterior Convergence*

To show the posterior convergence for the simulation over 5,000 iterations, we randomly select two main effect selection indicators at level $K-1$ and two interaction effect selection indicators at level K, and summarize their marginal inclusion probability for every 50 iterations. Figure 1 shows the trace plots for simulated 2D image data with $n = 300$ and low noise, where we can clearly identify convergence after the burn-in period. We have also conducted convergence diagnostics using the GR method (Gelman and Rubin, 1992). The potential scale reduction factors (PSRF) for the log-likelihood with multiple runs started from random initials is smaller than 1.2, confirming convergence.

C.2 *Simulation 2*

To assess the robustness of the BSSS to the partition construction, besides BSSS-P1 and BSSS-P2 from simulation 1, we consider three additional partition designs (BSSS-P3, BSSS-P4 and BSSS-P5) with four levels of partition for the 2D case as shown in Figure 2. The only difference among the three designs lies on the partition level 1. For BSSS-P3, the partition level 1 can perfectly distinguish activated and unactivated regions. For BSSS-P4 and BSSS-P5, the partition level 1 shifts, considering as "wrong" construction. We apply all the five BSSS versions to the simulated data following the same implementation procedure as that for the first simulation, and as shown in Table 1, the five designs yield quite consistent results and all of them outperform the completing approaches. This finding implies the relatively robustness of our model to the partition design.

C.3 *Simulation 3*

We perform a real data scale simulation by directly using the 2-back versus baseline and 0-back versus baseline whole brain imaging data from the ABCD study as our multi-modality main effects and consider their interaction in prediction. To generate $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_{<1,2>}$, we set the locations of the main and interaction effect signals closely related to actual imaging signals we identified in the data analysis in Section 4; and generate the nonzero elements in $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ from $\mathrm{N}(2,1)$; and those in $\boldsymbol{\beta}_{<1,2>}$ from $\mathrm{N}(-2,1)$. We also consider different signal-to-noise-ratios with R^2 around 0.76 (high noise) and 0.99 (low noise). We simulate 100 MC datasets, and to assess the out-of-sample performance, for each dataset, we randomly split the data by 5-fold, with 4-fold to be training set and 1-fold to be testing set. The partition construction for BSSS stays the same as that in the data application, and the implementations for BSSS, LASSO and horseshoe directly follow those in the first simulation. Eventually, the selection accuracy and prediction performance are presented in Table 2. Based on the result, the proposed BSSS maintains its superior performance in both feature selection

and prediction compared with competing alternatives. This reassures the use of BSSS in the ABCD application. Particularly, when comparing between different noise levels, we conclude that BSSS displays satisfactory robustness to high noises without a significant decrease of AUCs and the out-of-sample R^2. The LASSO and horseshoe, however, show different degrees of deterioration in predictive power compared to low noise scenarios.

Web Appendix D. Additional analysis results for the ABCD study

D.1 *Posterior Convergence*

We have checked the posterior convergence of our BSSS implementation. We randomly select two main effect selection indicators at level $K-1$ and two interaction effect selection indicators at level K, and summarize their marginal inclusion probability for every 50 iterations. As it is shown in Figure 3, the chain is clearly converged after the burn-in period (marked by the red dashed line). To further confirm the posterior convergence based on the GR method (Gelman and Rubin, 1992), under multiple runs started from random initials, we have also calculated the potential scale reduction factors (PSRF) for the log-likelihood over iterations after burn-in, and the value is smaller than 1.2, confirming convergence.

D.2 *Additional tables*

Based on the analysis results for the ABCD data, we summarize the posterior median proportion of voxels with a posterior inclusion probability higher than 0.95 for each AAL ROI. The full list of brain regions and the associated signal proportion under both main and interaction effects are provided in Table 3.

D.3 *Additional Figures*

The imaging protocol of the ABCD study follows a well-developed harmonization process to remove site effect, i.e. the measurement of brain structure and function have been harmonized to be compatible across three 3 tesla (T) scanner platforms: Siemens Prisma, General

Electric 750 and Phillips across 21 sites (Casey et al., 2018). In addition, with respect to our analysis, we also summarize the residuals across 21 sites as showed in Figure 4. Overall, the residuals are balanced across different sites with none of them displaying an obvious difference compared to the rest. This further confirms site information does not bring significant impact on our prediction.

D.4 *Sensitivity analysis*

We perform a sensitivity analysis to assess how sensitive the proposed BSSS method is to the pre-specified parameter settings. Given all of our hyper-priors are non-informative, we mainly focus on the partition design to evaluate the result changes under different partition constructions. Besides the partition design under six levels presented in the paper, we also implement BSSS under additional four partition designs with each of them also established under six levels. The difference between each additional partition design to the main one is we separately modify the partition construction at level 2 for two of them, and level 3 for the rest two. Once certain level of partitions are modified, all the subsequent levels will be changed accordingly. To assess the result consistency, we summarize the posterior median proportion of voxels with a posterior inclusion probability higher than 0.95 for each brain region and the corresponding brain sub-network for BSSS under each partition design, and the top 5 regions and top 3 sub-networks under each implementation are provided in Tables 4 and 5. Based on the results, we see a general consistency with mild differences, which confirms that partitions design though play a vital role in our method for signal detection, does not dramatically impact the final conclusion.

References

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* **88,** 669–679.

Casey, B., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., et al. (2018). The adolescent brain cognitive development (abcd) study: imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience* **32,** 43–54.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7,** 457–472.

Polson, N. G., Scott, J. G., and Windle, J. (2013). Bayesian inference for logistic models using pólya–gamma latent variables. *Journal of the American statistical Association* **108,** 1339–1349.

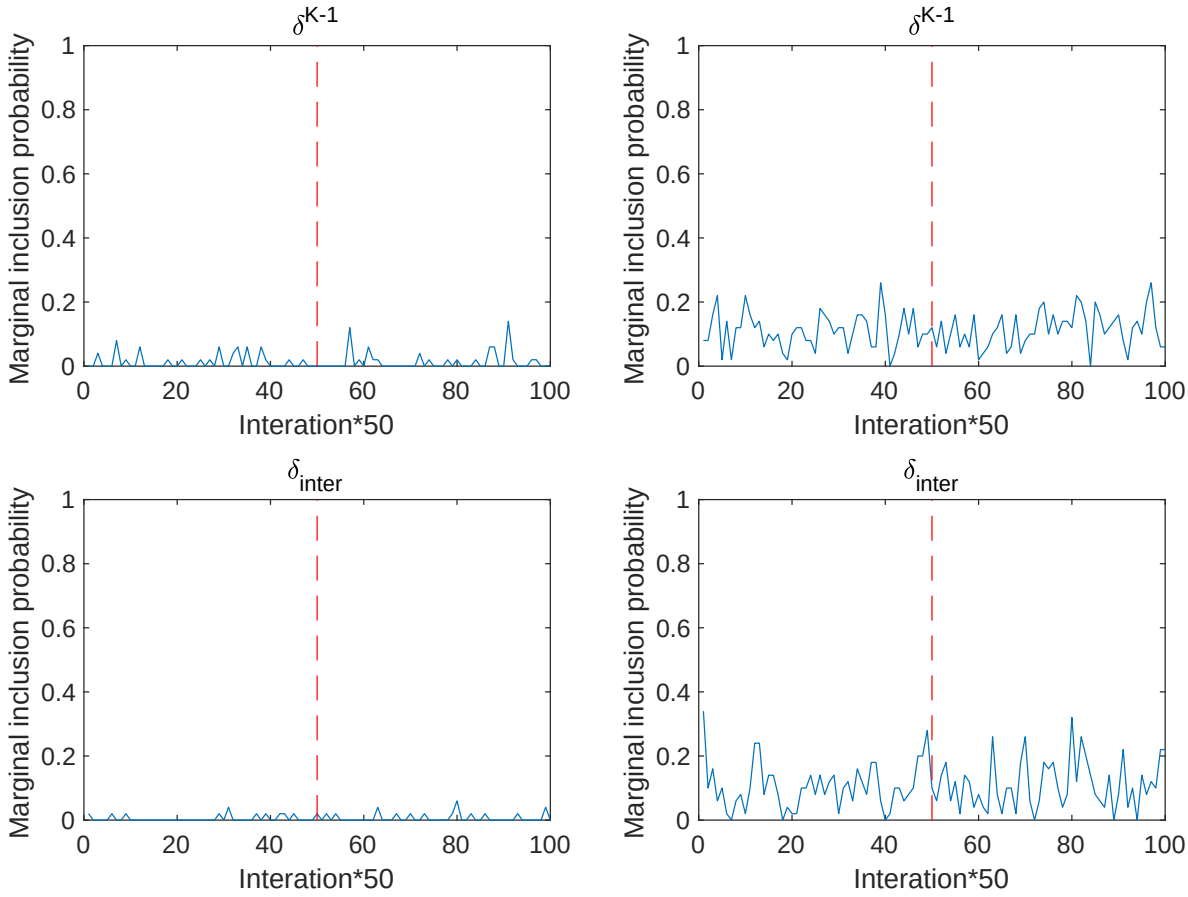

Figure 1 The trace plots over 5,000 iterations for two randomly selected main effect indicator and two randomly selected interaction indicator for their marginal inclusion probability over every 50 iterations on simulated 2D image data with $n = 300$ and high noise. The burn-in period is marked by the red dashed line.

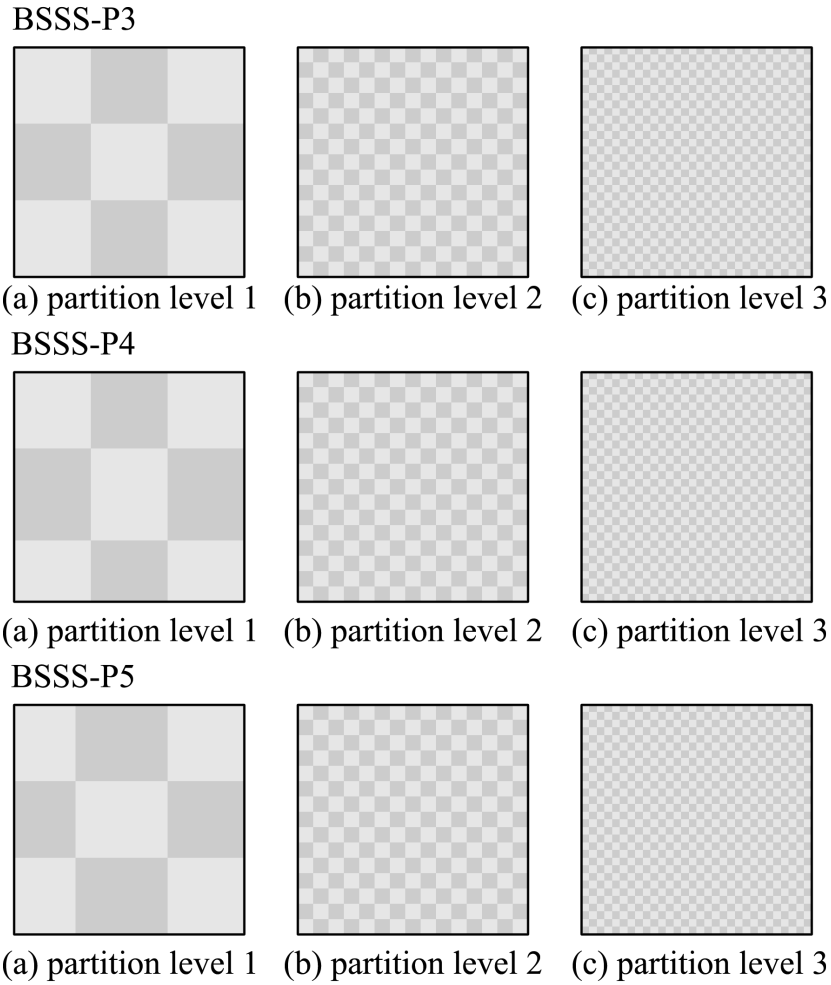

Figure 2 An illustration of the partition design at each level for BSSS-P3, BSSS-P4 and BSSS-P5. Level 3 for each case corresponds to the partition for each pixel.

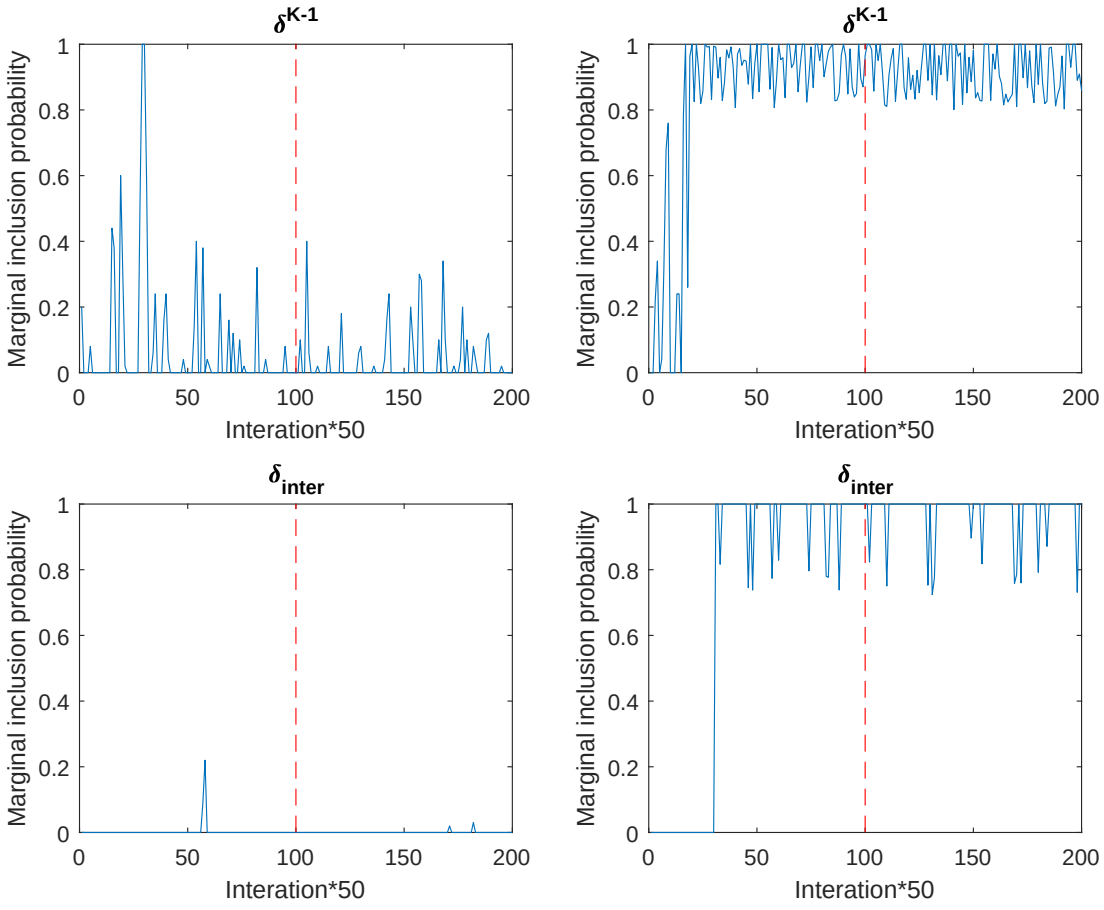

Figure 3 The trace plots over 10,000 iterations for two randomly selected main effect indicator and two randomly selected interaction indicator for their marginal inclusion probability over every 50 iterations. The burn-in period is marked by the red dashed line.

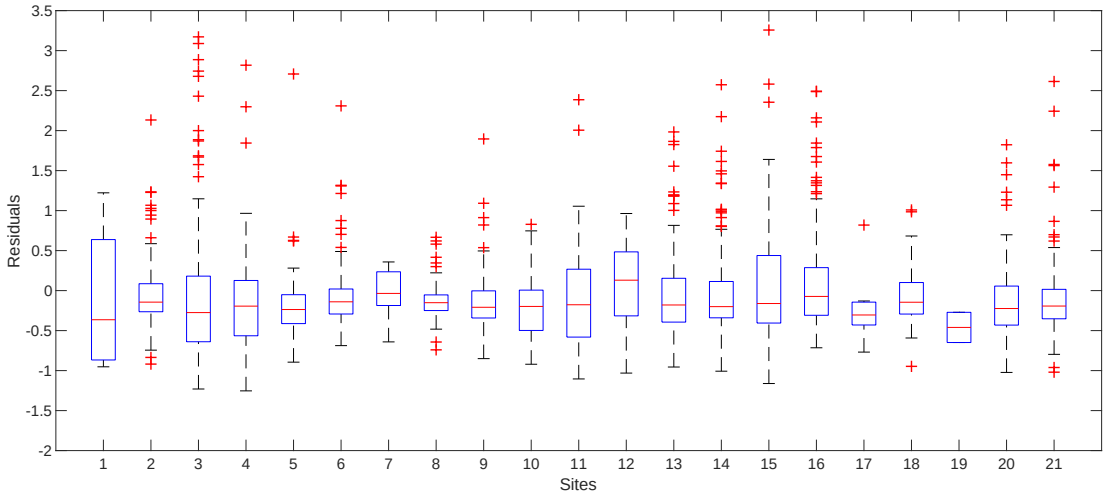

Figure 4 The box plot of residuals estimated over 21 sites.

Table 1 Additional simulation results with different partition designs for BSSS implementation are evaluated by AUCs on identifying features within main effects and interaction (AUC_{all}), main effects only (AUC_{main}) and interaction only (AUC_{inter}); and out-of-sample prediction (R^2). The Monte Carlo standard deviation for each metric is included in the parentheses.

2D Image									
		Low noise				High noise			
n	Method	AUC_{all}	AUC_{main}	AUC_{inter}	R^2	AUC_{all}	AUC_{main}	AUC_{inter}	R^2
100	BSSS-P1	0.919(0.018)	0.914(0.022)	0.966(0.021)	0.734(0.058)	0.902(0.021)	0.897(0.027)	0.957(0.025)	0.509(0.072)
	BSSS-P2	0.962(0.010)	0.969(0.008)	0.976(0.009)	0.744(0.057)	0.947(0.014)	0.962(0.012)	0.970(0.012)	0.520(0.069)
	BSSS-P3	0.946(0.078)	0.954(0.085)	0.969(0.080)	0.730(0.101)	0.876(0.153)	0.891(0.178)	0.911(0.172)	0.463(0.146)
	BSSS-P4	0.918(0.111)	0.920(0.118)	0.950(0.142)	0.717(0.124)	0.862(0.154)	0.871(0.174)	0.905(0.186)	0.467(0.123)
	BSSS-P5	0.913(0.094)	0.916(0.103)	0.955(0.109)	0.709(0.121)	0.871(0.133)	0.878(0.152)	0.928(0.149)	0.469(0.135)
300	BSSS-P1	0.939(0.017)	0.933(0.019)	0.968(0.027)	0.783(0.037)	0.931(0.012)	0.923(0.016)	0.974(0.013)	0.566(0.054)
	BSSS-P2	0.975(0.003)	0.976(0.003)	0.980(0.006)	0.788(0.037)	0.969(0.006)	0.973(0.005)	0.979(0.008)	0.572(0.054)
	BSSS-P3	0.973(0.003)	0.975(0.004)	0.982(0.005)	0.788(0.037)	0.966(0.007)	0.972(0.005)	0.982(0.007)	0.572(0.053)
	BSSS-P4	0.955(0.015)	0.953(0.020)	0.981(0.005)	0.785(0.038)	0.953(0.011)	0.953(0.015)	0.981(0.006)	0.570(0.053)
	BSSS-P5	0.955(0.017)	0.956(0.024)	0.983(0.005)	0.784(0.038)	0.950(0.013)	0.954(0.019)	0.982(0.007)	0.568(0.055)
500	BSSS-P1	0.939(0.019)	0.931(0.020)	0.967(0.024)	0.793(0.032)	0.936(0.015)	0.929(0.018)	0.971(0.020)	0.575(0.042)
	BSSS-P2	0.977(0.003)	0.977(0.003)	0.982(0.005)	0.796(0.032)	0.973(0.004)	0.975(0.004)	0.981(0.006)	0.579(0.042)
	BSSS-P3	0.975(0.003)	0.976(0.004)	0.984(0.004)	0.796(0.032)	0.970(0.005)	0.973(0.004)	0.983(0.005)	0.580(0.042)
	BSSS-P4	0.948(0.014)	0.945(0.018)	0.982(0.004)	0.794(0.032)	0.954(0.015)	0.951(0.018)	0.982(0.005)	0.578(0.042)
	BSSS-P5	0.951(0.025)	0.951(0.033)	0.984(0.004)	0.793(0.032)	0.954(0.012)	0.954(0.019)	0.983(0.005)	0.576(0.042)

Table 2 Simulation results with generation in light of the ABCD data under different noise levels. The selection accuracy is quantified by AUCs on identifying features within main effects and interaction (AUC_{all}), main effects only (AUC_{main}) and interaction only (AUC_{inter}); and prediction is evaluated by out-of-sample prediction (R^2). The Monte Carlo standard deviation for each metric is included in the parentheses.

	Method	AUC_{all}	AUC_{main}	AUC_{inter}	R^2
Low noise	BSSS	0.889 (0.003)	0.941 (0.002)	0.640 (0.008)	0.978 (0.002)
	LASSO	0.510 (0.005)	0.510 (0.005)	0.509 (0.011)	0.970 (0.002)
	horseshoe	0.563 (0.006)	0.581 (0.006)	0.559 (0.018)	0.947 (0.003)
High noise	BSSS	0.843 (0.003)	0.979 (0.001)	0.567 (0.010)	0.663 (0.004)
	LASSO	0.506 (0.006)	0.509 (0.005)	0.505 (0.010)	0.485 (0.006)
	horseshoe	0.515 (0.006)	0.526 (0.006)	0.511 (0.011)	0.359 (0.007)

Table 3 The median proportion of the voxels with the posterior inclusion probability higher than 0.95 on the AAL regions from the main and interaction effect results for the ABCD study.

	2back-0back	2back	0back	2back-0back×2back	2back-0back×0back	2back×0back
Precentral_L	0.055	0.074	0.088	0.063	0.032	0.066
Precentral_R	0.051	0.048	0.065	0.060	0.029	0.076
Frontal_Sup_L	0.016	0.020	0.022	0.056	0.046	0.066
Frontal_Sup_R	0.027	0.031	0.040	0.059	0.035	0.078
Frontal_Sup_Orb_L	0.051	0.048	0.068	0.041	0.068	0.024
Frontal_Sup_Orb_R	0.039	0.048	0.048	0.045	0.064	0.023
Frontal_Mid_L	0.008	0.010	0.008	0.046	0.036	0.061
Frontal_Mid_R	0.075	0.077	0.087	0.054	0.045	0.051
Frontal_Mid_Orb_L	0.067	0.063	0.078	0.026	0.070	0.007
Frontal_Mid_Orb_R	0.048	0.061	0.068	0.048	0.054	0.014
Frontal_Inf_Oper_L	0.107	0.104	0.153	0.086	0.037	0.071
Frontal_Inf_Oper_R	0.116	0.154	0.166	0.057	0.045	0.067
Frontal_Inf_Tri_L	0.107	0.103	0.106	0.056	0.055	0.052
Frontal_Inf_Tri_R	0.097	0.107	0.122	0.038	0.056	0.051
Frontal_Inf_Orb_L	0.147	0.159	0.181	0.052	0.060	0.018
Frontal_Inf_Orb_R	0.109	0.139	0.158	0.048	0.077	0.026
Rolandic_Oper_L	0.070	0.053	0.053	0.033	0.040	0.040
Rolandic_Oper_R	0.088	0.078	0.090	0.055	0.065	0.055
Supp_Motor_Area_L	0.053	0.061	0.053	0.070	0.021	0.096
Supp_Motor_Area_R	0.110	0.132	0.125	0.075	0.020	0.087
Olfactory_L	0.046	0.069	0.103	0.034	0.034	0.000
Olfactory_R	0.025	0.062	0.062	0.049	0.037	0.012
Frontal_Sup_Medial_L	0.125	0.134	0.138	0.051	0.046	0.061
Frontal_Sup_Medial_R	0.103	0.136	0.133	0.041	0.041	0.053
Frontal_Med_Orb_L	0.151	0.147	0.191	0.031	0.076	0.022
Frontal_Med_Orb_R	0.115	0.141	0.195	0.027	0.065	0.019
Rectus_L	0.096	0.084	0.115	0.061	0.057	0.027
Rectus_R	0.078	0.073	0.073	0.037	0.064	0.037
Insula_L	0.097	0.117	0.118	0.037	0.067	0.030
Insula_R	0.130	0.167	0.189	0.041	0.069	0.028
Cingulum_Ant_L	0.148	0.150	0.169	0.038	0.045	0.054
Cingulum_Ant_R	0.141	0.151	0.174	0.040	0.050	0.045
Cingulum_Mid_L	0.103	0.110	0.116	0.053	0.039	0.052
Cingulum_Mid_R	0.117	0.109	0.142	0.048	0.053	0.066
Cingulum_Post_L	0.080	0.036	0.080	0.080	0.036	0.066
Cingulum_Post_R	0.023	0.034	0.023	0.046	0.034	0.046
Hippocampus_L	0.147	0.150	0.194	0.051	0.070	0.011
Hippocampus_R	0.160	0.194	0.212	0.052	0.076	0.021
ParaHippocampal_L	0.080	0.091	0.091	0.056	0.077	0.024
ParaHippocampal_R	0.187	0.187	0.269	0.047	0.066	0.032
Amygdala_L	0.065	0.048	0.081	0.048	0.065	0.048
Amygdala_R	0.100	0.086	0.114	0.014	0.043	0.014
Calcarine_L	0.093	0.108	0.136	0.046	0.060	0.025
Calcarine_R	0.166	0.164	0.199	0.046	0.072	0.020
Cuneus_L	0.067	0.056	0.071	0.049	0.047	0.071
Cuneus_R	0.136	0.150	0.182	0.032	0.035	0.069
Lingual_L	0.132	0.179	0.203	0.027	0.061	0.023
Lingual_R	0.112	0.140	0.161	0.041	0.069	0.016
Occipital_Sup_L	0.068	0.068	0.073	0.053	0.053	0.076
Occipital_Sup_R	0.061	0.075	0.063	0.042	0.044	0.065
Occipital_Mid_L	0.066	0.088	0.101	0.043	0.047	0.044
Occipital_Mid_R	0.077	0.064	0.079	0.047	0.050	0.050
Occipital_Inf_L	0.075	0.131	0.104	0.030	0.056	0.019
Occipital_Inf_R	0.076	0.083	0.086	0.035	0.051	0.019
Fusiform_L	0.106	0.116	0.146	0.058	0.051	0.028
Fusiform_R	0.105	0.104	0.142	0.055	0.047	0.029
Postcentral_L	0.013	0.020	0.016	0.061	0.036	0.066
Postcentral_R	0.046	0.065	0.081	0.059	0.040	0.064
Parietal_Sup_L	0.068	0.082	0.084	0.062	0.022	0.076
Parietal_Sup_R	0.113	0.138	0.127	0.057	0.022	0.091
Parietal_Inf_L	0.057	0.063	0.063	0.046	0.029	0.063
Parietal_Inf_R	0.124	0.131	0.143	0.069	0.033	0.057
SupraMarginal_L	0.085	0.065	0.071	0.045	0.042	0.065
SupraMarginal_R	0.131	0.156	0.142	0.047	0.040	0.071
Angular_L	0.067	0.088	0.076	0.067	0.041	0.079
Angular_R	0.073	0.071	0.069	0.046	0.034	0.071
Precuneus_L	0.041	0.053	0.048	0.052	0.037	0.083
Precuneus_R	0.047	0.044	0.043	0.051	0.039	0.068
Paracentral_Lobule_L	0.090	0.116	0.130	0.076	0.028	0.071
Paracentral_Lobule_R	0.048	0.084	0.084	0.066	0.022	0.084
Caudate_L	0.090	0.101	0.122	0.036	0.079	0.043
Caudate_R	0.092	0.067	0.092	0.049	0.056	0.042
Putamen_L	0.082	0.088	0.082	0.042	0.056	0.016
Putamen_R	0.096	0.093	0.102	0.043	0.053	0.022
Pallidum_L	0.025	0.037	0.062	0.012	0.037	0.025
Pallidum_R	0.039	0.039	0.039	0.026	0.105	0.013
Thalamus_L	0.070	0.086	0.118	0.029	0.093	0.029
Thalamus_R	0.075	0.062	0.078	0.039	0.081	0.029
Heschl_L	0.069	0.069	0.042	0.014	0.056	0.028
Heschl_R	0.082	0.082	0.068	0.055	0.055	0.014
Temporal_Sup_L	0.068	0.082	0.068	0.037	0.070	0.034
Temporal_Sup_R	0.091	0.107	0.120	0.036	0.065	0.039
Temporal_Pole_Sup_L	0.076	0.079	0.120	0.042	0.071	0.039
Temporal_Pole_Sup_R	0.128	0.128	0.198	0.033	0.070	0.025
Temporal_Mid_L	0.040	0.035	0.036	0.044	0.071	0.031
Temporal_Mid_R	0.071	0.088	0.114	0.047	0.071	0.027
Temporal_Pole_Mid_L	0.113	0.090	0.158	0.045	0.036	0.027
Temporal_Pole_Mid_R	0.095	0.106	0.138	0.092	0.054	0.060
Temporal_Inf_L	0.136	0.126	0.177	0.065	0.053	0.030
Temporal_Inf_R	0.098	0.107	0.145	0.052	0.058	0.032

Table 4 The proportion of the voxels with the posterior inclusion probability higher than 0.95 on the top 5 AAL regions from 5 partition designs under the BSSS implementation for the ABCD study.

			2back-0back		
	partition 1	partition 2	partition 3	partition 4	partition 5
ParaHippocampal_R	0.187	0.203	0.184	0.193	0.142
Calcarine_R	0.166	0.185	0.173	0.166	0.076
Hippocampus_R	0.115	0.160	0.198	0.181	0.115
Frontal_Med_Orb_L	0.169	0.151	0.133	0.204	0.111
Cingulum_Ant_L	0.110	0.178	0.153	0.148	0.063
			2back		
	partition 1	partition 2	partition 3	partition 4	partition 5
Hippocampus_R	0.219	0.184	0.194	0.205	0.125
ParaHippocampal_R	0.231	0.187	0.237	0.180	0.089
Lingual_L	0.212	0.179	0.194	0.176	0.097
Insula_R	0.184	0.202	0.167	0.135	0.080
Calcarine_R	0.194	0.164	0.175	0.164	0.103
			0back		
	partition 1	partition 2	partition 3	partition 4	partition 5
ParaHippocampal_R	0.446	0.269	0.282	0.269	0.146
Hippocampus_R	0.306	0.212	0.247	0.201	0.104
Lingual_L	0.280	0.209	0.177	0.203	0.091
Calcarine_R	0.232	0.199	0.220	0.197	0.092
Temporal_Pole_Sup_R	0.275	0.188	0.198	0.228	0.080
			2back-0back×2back		
	partition 1	partition 2	partition 3	partition 4	partition 5
Temporal_Pole_Mid_R	0.100	0.100	0.063	0.074	0.092
Frontal_Inf_Oper_L	0.095	0.086	0.086	0.037	0.071
Cingulum_Post_L	0.080	0.058	0.080	0.131	0.109
Paracentral_Lobule_L	0.081	0.052	0.081	0.076	0.066
Supp_Motor_Area_R	0.068	0.071	0.075	0.095	0.077
			2back-0back×0back		
	partition 1	partition 2	partition 3	partition 4	partition 5
Pallidum_R	0.092	0.105	0.066	0.118	0.158
Thalamus_L	0.032	0.112	0.067	0.093	0.102
Thalamus_R	0.101	0.078	0.081	0.081	0.098
Caudate_L	0.050	0.083	0.086	0.079	0.079
Frontal_Inf_Orb_R	0.079	0.077	0.063	0.081	0.061
			2back×0back		
	partition 1	partition 2	partition 3	partition 4	partition 5
Supp_Motor_Area_L	0.108	0.096	0.070	0.069	0.099
Parietal_Sup_R	0.121	0.077	0.066	0.122	0.091
Supp_Motor_Area_R	0.134	0.078	0.087	0.081	0.087
Paracentral_Lobule_R	0.053	0.088	0.110	0.075	0.084
Precuneus_L	0.084	0.083	0.101	0.081	0.071

Table 5 The proportion of the voxels with the posterior inclusion probability higher than 0.95 on the top 3 functional sub-networks from 5 partition designs under the BSSS implementation for the ABCD study.

	2back-0back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Auditory	0.053	0.130	0.160	0.136	0.036
Cingulo-opercular Task Control	0.040	0.115	0.150	0.150	0.053
Visual	0.086	0.105	0.131	0.150	0.068
	2back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Auditory	0.047	0.130	0.148	0.160	0.036
Visual	0.117	0.103	0.126	0.143	0.072
Memory retrieval	0.091	0.159	0.114	0.227	0.045
	0back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Visual	0.173	0.133	0.147	0.178	0.061
Memory retrieval	0.136	0.182	0.136	0.159	0.023
Auditory	0.148	0.148	0.124	0.136	0.053
	2back-0back×2back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Cingulo-opercular Task Control	0.026	0.062	0.048	0.062	0.084
Sensory/somatomotor Hand	0.051	0.067	0.070	0.062	0.056
Auditory	0.036	0.053	0.059	0.089	0.089
	2back-0back×0back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Subcortical	0.107	0.062	0.062	0.080	0.080
Memory retrieval	0.068	0.068	0.091	0.091	0.045
Salience	0.062	0.062	0.052	0.055	0.069
	2back×0back				
	partition 1	partition 2	partition 3	partition 4	partition 5
Ventral attention	0.059	0.129	0.106	0.059	0.106
Fronto-parietal Task Control	0.072	0.049	0.062	0.075	0.068
Auditory	0.124	0.059	0.053	0.018	0.071