A Appendix

A.1 Performance evaluation for Ψ-learning

In this section, we evaluate the performance of the proposed ψ-learning method on edge identification based on simulations. The performance comparison between ψ-learning and other existing methods can be found in Liang et. al.. Here we mainly evaluate the diagnostic ability and false discovery rate (FDR) for different settings of sample size n and variable size p.

We consider the small-world network structure to simulate the brain connectivity. We used R package *igraph* to construct the adjacency matrices $E_{p\times p}$'s and set the average neighborhood size to be 2, i.e., the total number of edges in a graph is $2p$, where p is the number of nodes (variables). Then the precision matrix $\Omega_{p\times p}$ should reflect the adjacency matrix $E_{p\times p}$ in the sense that $E_{ij} = 1 \Leftrightarrow \Omega_{ij} \neq 0$. We generate the values of all non-zero elements in the precision matrix independently from the uniform distribution in the interval $[0.5, 1] \cup [-1, -0.5]$. We set the diagonal values as $1 + |\Lambda_{min}(\Omega)|$ to ensure the positive definiteness. Then, we get the covariance matrix $\Sigma = \Omega^{-1}$ and generate $X \sim N(0, \Sigma)$.

Consider the following three (n, p) settings: $(50, 100)$,$(50, 200)$,$(100, 200)$ and set $\alpha_1 = 0.1$ and $\alpha_2 = 0.05$. Under each setting, we simulated 50 datasets independently. Figure A.1 shows the ROC curve of the ψ-learning under various settings. It can reflect the diagnostic ability of a binary classifier. Let us define an experiment from P positive instances and N negative instances for some condition. In our case, the positive instance represents there is an edge between two nodes. The four outcomes are summarized in Table A.I. In a ROC curve, the x, y axes stand for the specificity and sensitivity, respectively, and their definitions are given as follow:

$$\text{specificity} = \frac{\text{TN}}{\text{FN} + \text{TN}}, \quad \text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Table A.II calculates the areas under the ROC curve (AUC) and the FDR for each setting. Combined the results from Figure A.1 and Table A.II, we can see as the variable size p increases , it may affect the performance of ψ-learning, but the performance is steady overall. As the sample size n increases, it largely improves the diagnostic ability. Under all conditions, the FDR is steadily controlled around 0.05.

Table A.I: Outcomes of a binary decision

	Actual positive (P)	Actual negative (N)
Predicted positive	True positive (TP)	False positive (FP)
Predicted negative	False negative (FN)	True negative (FN)

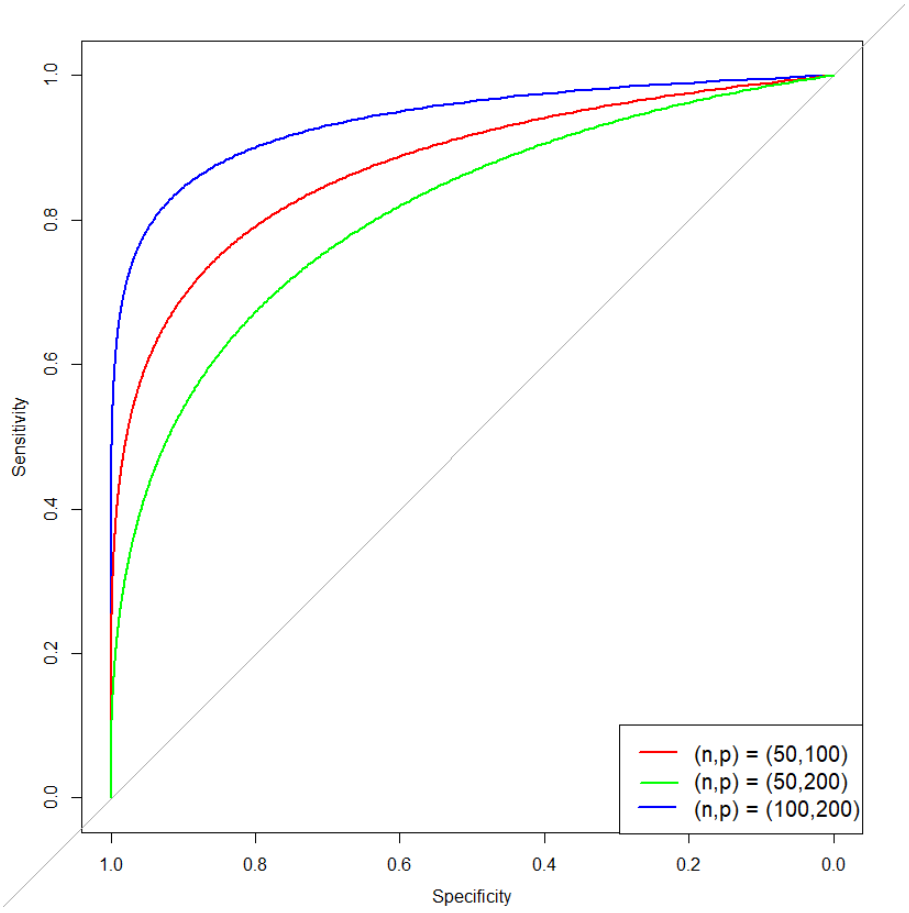


Figure A.1: The ROC curves for different (n, p) settings

Table A.II: Average areas under the ROC curves resulting from different settings

(n, p)	AUC (SD)	FDR (SD)
$(50, 100)$	0.8797 (0.0130)	0.06418 (0.04572)
$(50, 200)$	0.8433 (0.01697)	0.05604 (0.05753)
$(100, 200)$	0.9427 (0.0058)	0.04514 (0.02162)

A.2 Validation of the Ψ-learning

In order to test the reliability and stability of our model, the data for each group (case, control) were separated into two sets randomly with equal probability, denote the newly generated sets as case 1, case 2, control 1, control 2, respectively. Similar to Haxby et. al. [1], we assessed the degree of similarity among these four datasets using the correlation coefficients. The results are given in Table A.III. The two sets from the same group have a high degree of similarity, while the sets from different groups are barely correlated to each other. Furthermore, we calculated the global measures for the networks generated by the four sets (see Table A.IV) and conducted permutation tests on those statistics to see the significant difference (Table A.V). As we can see from the tables, the values of the network statistics from the same groups are close to each other, while they vary significantly from each other across groups. The results of similarity comparisons using correlation coefficients and network statistics showed that the proposed model is reliable.

Table A.III: Correlations among the four datasets

	Case 1	Case 2	Control 1	Control 2
Case 1	-	0.3135	0.03201	0.04669
Case 2	-	-	0.04297	0.03057
Control 1	-	-	-	0.52188

Table A.IV: Global network measures for the four datasets

	Case 1	Case 2	Control 1	Control 2
Density	7.70×10^{-5}	6.58×10^{-5}	4.54×10^{-4}	5.12×10^{-4}
Mean clustering coefficient	0.0155	0.0082	0.3727	0.2889
global efficiency	0.0126	0.0141	0.3310	0.5218

Table A.V: The p-values of the difference mean of various network measures

	Density	Mean clustering coefficient	Global efficiency
Between case	0.91	0.94	0.99
Between control	0.67	0.36	0.33
Across groups[1]	< 0.01	< 0.01	0.01

[1] Note that for the comparison of across groups, we took the mean difference of each measure and calculate the corresponding p-value.

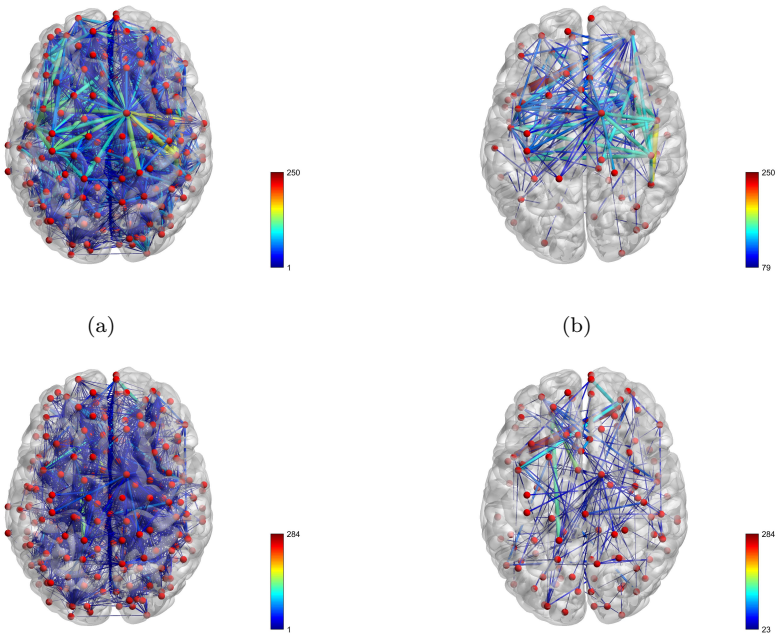
A.3 Brain network construction through the power atlas

In this session, we parcellated the fMRI data into 264 functionally defined regions by Power et. al [2]. Standard preprocessing steps were applied using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/), including motion correction, spatial normalization to standard MNI space and spatial smoothing with an 3mm FWHM Gaussian kernel. The influence of motion (6 parameters) was further addressed using a regression procedure, and the functional time series were band-pass filtered using a 0.01Hz to 0.1Hz frequency range. Finally, $p = 8878$ voxels were left after a t-test between the case and control groups and 264 ROIs were extracted based on the the power atlas [2].

We used the same parameter settings as in Session III, which are $\gamma = 25$, $\alpha_1 = 0.1$ and $\alpha_2 = 0.01$. Similar to the analysis using the AAL template, we first summarized some global network statistics as shown in Table A.IV. Figure A.2 (a) and (c) showed that, compared to healthy people, the brain connectivity of the SZ patients is significantly reduced. From Figure A.2 (b) and (d), we can see that SZ patients have lost some hub organizations on the frontal lobe like postcentral gyrus, frontal superior gyrus right and rolandic operculum left. We also applied the pairwise hypothesis test between the case and control groups for a solid statistical comparison. We followed the steps mentioned in Section 2.3, setting the significance level $\alpha_3 = 0.01$. Figure A.3 gave the axial view of the aberrant connectivity and Table A. IV listed the significantly aberrant hubs (degree $>$ mean + 3 sd).

Table A.VI: Basic measures for the case and control groups

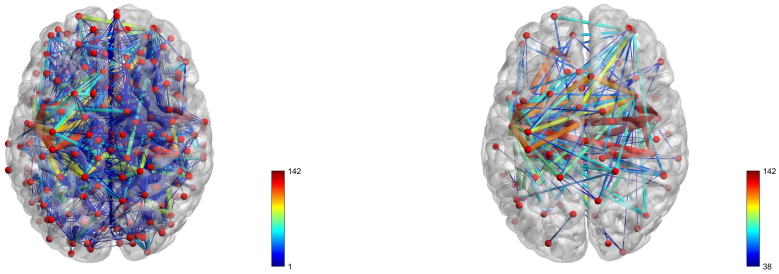
Measures	Case group	Control group	p-value
Density	0.0004	0.0018	< 0.01
Mean clustering coefficient	0.4394	0.6356	0.02
Transitivity	0.0100	0.0042	0.01
Characteristic path length	5.0000	2.1142	0.03
Global efficiency	0.1484	0.4815	0.03



(a) (b)

(c) (d)

Figure A.2: Axial view of the brain connectivity based on 264 parcellation region by Power et al.'s method. Left column shows the full connectivity for healthy people (a) and SZ patients (c). The right column shows filtered edges with sparsity = 0.005 for healthy individuals (b) and SZ patients (d). The colors of the edges show the strength of connectivity.



(a) (b)

Figure A.3: A visualization of aberrant brain connectivity and affected hub points. Figure (a) shows the axial view of the brain. Figure (b) shows the filtered edges with sparsity = 0.005.

Table A.VII: Identified aberrant hubs

	ROI	Functional system	Degree
1	Rolandic operculum left (ORG)	Auditory	4915
2	Postcentral gyrus left (PAG)	Motor and somatosensory	4913
3	Insula gyrus right (IND)	Cingulo-opercular	4913

Comparing the results of using the AAL template and the power atlas, we find that the global measures of the brain networks basically remain the same numerical level and the differences between the case and control group are both significant. The power atlas is more sensitive to capture the aberrance on the frontal lobe, while the AAL template tends to identify the abnormality on the temporal lobe. We have detected that the postcentral gyrus (PA) are significantly aberrant using both templates, which reveals the importance of PA to the SZ disease.

References

[1] J. Haxby, M. Gobbini, M. Furey, A. Ishai, J. Schouten, and P. Pietrini, "Distributed and overlapping representations of faces and objects in ventral temporal cortex.," *Science*, vol. 293, pp. 2425–2430, 2001.

[2] J. D. Power, D. A. Fair, B. L. Schlaggar, and S. E. Petersen, "The development of human functional brain networks.," *Neuron*, vol. 67(5), pp. 735–748, 2010.