

# **A Joint Constrained CCA Model for Network-Dependent Brain Subregion Parcellation: Supplementary Material**

Qinrui Ling<sup>1</sup>, Aiping Liu<sup>1</sup>, Yu Li<sup>1</sup>, Xueyang Fu<sup>2</sup>, Xun Chen<sup>1</sup>, Martin J. McKeown<sup>3</sup>, and Feng Wu<sup>1</sup>

<sup>1</sup> *Department of Electronic Engineering and Information Science, University of Science and  
Technology of China, Hefei, China*

<sup>2</sup> *Department of Automation, University of Science and Technology of China, Hefei, China*

<sup>3</sup> *Department of Medicine, University of British Columbia, Vancouver, Canada*

## **I. NYU DATA**

### **A. Participants and Data Acquisition**

Twenty-five right-handed participants (fifteen males and ten females, range 21–49 years old) were included with no history of psychiatric or neurological disorders, which was confirmed by a psychiatric clinical assessment. Each subject had three sessions of resting state fMRI data. The experiments were approved by the local Ethics Board and written informed consent forms were obtained from all participants prior to the study.

A SIEMENS Allegra 3.0 Tesla scanner equipped for echoplanar imaging (EPI) was used for data acquisition. During the scans, participants were instructed to stay rest with their eyes open. Each participant was collected three resting-state scans of 197 continuous EPI functional volumes with the following specifications: repetition time of 2000 ms, echo time of 25 ms, flip angle of 90°, axial slices of 39 layers, matrix size of  $64 \times 64$ , field of view of  $192 \text{ mm} \times 192 \text{ mm}$ , and acquisition voxel size of  $3.0 \text{ mm} \times 3.0 \text{ mm} \times 3.0 \text{ mm}$ . Scans 2 and 3 were conducted in a single scan session, about 30 minutes apart, and were 5-11 months after scan 1. For spatial localization and normalization, a high-resolution T1-weighted anatomical image was also obtained using a magnetization prepared gradient echo sequence with repetition time of 2500 ms, echo time of 4.35 ms, inversion time of 900 ms, and flip angle = 8°. Additionally, as complete cerebellar coverage was impossible for all participants, only those cerebellar regions acquired in all participants were included in subsequent analysis.

### **B. fMRI Data Preprocessing**

The acquired rs-fMRI data was preprocessed by DPABI (version 3.0, <http://rfmri.org/dpabi>) and SPM8 (<https://www.fl.ion.ucl.ac.uk/spm>). The first five time points were removed to make the magnetization dynamic stabilization and slice timing correction was performed. The manually reoriented functional images were then co-registered to structural images, and were normalized to the Montreal Neurological Institute (MNI) space at  $3.0 \times 3.0 \times 3.0 \text{ mm}^3$  resolution. Next, nuisance time courses were

Table S1  
THE ROIS SELECTED IN NYU DATA

No.	Name	No.	Name
1:2	Precentral	49:54	Occipital
3:16	Frontal	55:56	Fusiform
17:18	Rolandic_Oper	57:58	Postcentral
19:20	Supp_Motor_Area	59:62	Parietal
21:22	Olfactory	63:64	SupraMarginal
23:26	Frontal	65:66	Angular
27:28	Rectus	67:68	Precuneus
29:30	Insula	69:70	Paracentral_Lobule
31:36	Cingulum	71:72	Caudate
37:38	Hippocampus	73:74	Putamen
39:40	ParaHippocampal	75:76	Pallidum
41:42	Amygdala	77:78	Thalamus
43:44	Calcarine	79:80	Heschl
45:46	Cuneus	81:90	Temporal
47:48	Lingual	91:108	Cerebellum

Each region is suffixed by left(L), right(R), sup, mid or inf according to its location.

regressed from the processed data including Friston 24 head motion parameters, linear trend, white-matter signal, and cerebrospinal fluid signal. Finally, fMRI signals were spatially smoothed using  $6 \times 6 \times 6$  FWHM, and bandpass filtered between 0.01 to 0.08 Hz. In order to define the subregions in the Putamen, 106 other ROIs were selected as reference ROIs for generating the connectivity patterns of the Putamen, as shown in Table S1.

### C. Results

1) *Subregions of the Putamen:* To better evaluate the reproducibility of the proposed method, we assessed its performance on the repeated sessions of resting state fMRI data. For fair comparisons, all the subjects were registered in the common space. We parcellated the left and right Putamen separately, whose numbers of voxels were 306 and 322 respectively. 106 ROIs were chosen as reference regions to establish the connectivity patterns with the Putamen and guide its parcellations. Here, we set the number of subregions  $K = 3$ . Suppose that the defined subregions were spatially continuous whereas the reference ROIs did not require spatial continuity but required the sparsity in loadings. The Laplacian matrix was generated by the the distance matrix of each voxel in the Putamen in 3-D Euclidean space, so as to reflect the spatial continuity of voxels. The tuning parameters  $\beta_1, \beta_2, \gamma_1$  were chosen using a 5-fold cross validation (CV) method as  $\beta_1 = 0.01, \beta_2 = 2.5, \gamma_1 = 0.01$ . As iterative

Table S2  
THE SIGNIFICANTLY CORRELATED REFERENCE ROIs WITH THE LEFT PUTAMEN

	subregion 1	subregion 2	subregion 3
session 1	Insula_L	Precentral_R	Frontal_Inf_Oper_L
	Amygdala_L&R	Rolandic_Oper_L&R	Frontal_Inf_Tri_L&R
	Pallidum_L&R	Postcentral_L&R	Frontal_Inf_Orb_L&R
	Thalamus_L&R	Temporal_Sup_R	Parietal_Inf_L
session 2	Insula_L	Rolandic_Oper_L&R	Olfactory_L&R
	Amygdala_L&R	SupraMarginal_L	Frontal_Med_Orb_L
	Pallidum_L&R	Temporal_Sup_L&R	Caudate_L&R
	Thalamus_L&R		
session 3	Insula_L	Precentral_L&R	Olfactory_L&R
	Cingulum_Mid_L	Rolandic_Oper_R	Cingulum_Ant_L
	Amygdala_L&R	Supp_Motor_Area_R	Caudate_L&R
	Pallidum_L&R	Postcentral_L&R	
	Thalamus_L&R	Temporal_Sup_R	

method was adopted to solve the problem, it might fall into local rather than global minimum. We thus repeated the experiments 100 times with random initial values and chose the one with the smallest loss function. After obtaining the canonical correlation loadings by JC-CCA, we derived the soft parcellation indicated by  $u_i$  and the corresponding significantly correlated reference ROIs by  $v$ , as shwon in Table S2 and Table S3.

As demonstrated in the tables, for different sessions, the correlated reference ROIs were similar. For visualization, we selected the largest 30% weights of each  $u_i$  and used the corresponding voxels to constitute a subregion. The parcellation results of one typical subject is illustrated in Fig.S1. For the left Putamen, subregion 1 was located in the inferior anterior part, subregion 2 mainly occupied the posterior part, and subregion 3 was located in the anterior part. For the right Putamen, subregion 1 was located in the inferior part, subregion 2 was located in the inferior part, and subregion 3 was located in the superior part. As we parcellated the left and right Putamen separately, the asymmetry was observed in both parcellation and their subregional connectivity patterns.

2) *Quantitative Analysis*: In order to quantitatively measure the consistency of the distribution of different subregions, we used Kendall's coefficient of concordance ( $W$ ) for  $u_i$  to access the agreement of subregions based on rank order, which ranges from 0 (no agreement) to 1 (complete agreement). We first calculated  $W$  for each subregion of twenty-five subjects in each session, as shown in Table S4. We observe that all  $P$  values were less than 0.001, indicating that each obtained subregion had significant agreement among all subjects. Then we calculated  $W$  for each subregion of each subject in 3 sessions, as shown in Table S5. We observe that 88%, 88%, and 80% of subjects had agreement

Table S3  
THE SIGNIFICANTLY CORRELATED REFERENCE ROIS WITH THE RIGHT PUTAMEN

	subregion 1	subregion 2	subregion 3
session 1	Rolandic_Oper_R	Parietal_Sup_L&R	Precentral_L&R
	Insula_L&R	Fusiform_L&R	Amygdala_L&R
	Amygdala_R	SupraMarginal_L&R	Postcentral_L&R
	Pallidum_L&R	Paracentral_Lobule_L&R	Temporal_Sup_R
	Thalamus_R		
	Temporal_Sup_R		
session 2	Rolandic_Oper_R	Lingual_L&R	ParaHippocampal_L&R
	Insula_L&R	Fusiform_L&R	Amygdala_L&R
	Pallidum_L&R	Cerebellum_6_L&R	Temporal_Pole_Sup_L&R
	Temporal_Sup_L&R		Temporal_Mid_R
session 3	Rolandic_Oper_R	Lingual_L	Olfactory_L
	Insula_L&R	Fusiform_L&R	Hippocampus_L&R
	Cingulum_Mid_R	Paracentral_Lobule_L&R	ParaHippocampal_L&R
	Amygdala_R	Cerebellum_4_5_R	
	Pallidum_L&R	Cerebellum_6_L	
	Thalamus_R		
	Temporal_Sup_L&R		

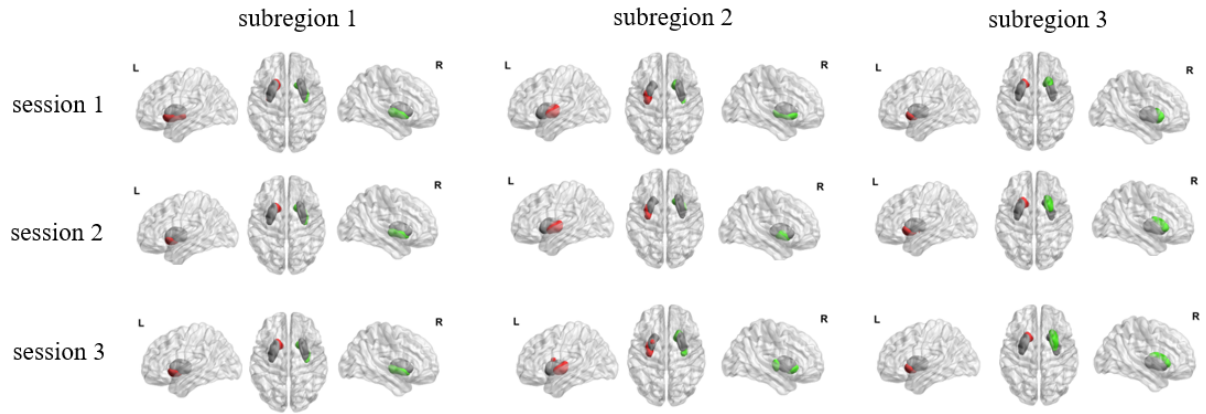


Figure S1. The hard parcellation results of one subject from three views (L, S, R). The red regions are subregions in the left Putamen and the green ones are subregions in the right Putamen.

among 3 sessions for subregion 1 to 3 in the left Putamen respectively and 96%, 84%, and 84% in the right Putamen. Given the possible variability across the subjects, the results demonstrate that for most subjects, we can get consistent parcellation among 3 sessions. The results demonstrate that the proposed method has superior reproducibility for joint functional brain subregion parcellation.

Further we calculated the correlation coefficients  $r$  which measured the functional connectivity

Table S4  
KENDALL'S COEFFICIENT OF CONCORDANCE OF 25 INDIVIDUALS

	the left Putamen			the right Putamen		
	subregion 1	subregion 2	subregion 3	subregion 1	subregion 2	subregion 3
session 1	0.390*	0.535*	0.386*	0.540*	0.527*	0.377*
session 2	0.410*	0.457*	0.677*	0.591*	0.389*	0.399*
session 3	0.370*	0.421*	0.780*	0.499*	0.365*	0.392*

Null hypothesis is no agreement ( $W = 0$ ). '\*' represents that null hypothesis is rejected, with  $P$  value  $< 0.001$ .

Table S5  
KENDALL'S COEFFICIENT OF CONCORDANCE OF 3 SESSIONS

subject	the left Putamen			the right Putamen		
	subregion 1	subregion 2	subregion 3	subregion 1	subregion 2	subregion 3
1	0.597**	0.883**	0.867**	0.940**	0.683**	0.399 (0.008)*
2	<b>0.312 (0.786)</b>	0.933**	0.834**	0.615**	<b>0.094 (1.000)</b>	0.475**
3	0.636**	0.751**	0.927**	0.781**	0.586**	0.502**
4	<b>0.311 (0.797)</b>	0.767**	0.711**	0.820**	<b>0.217 (1.000)</b>	0.488**
5	0.804**	0.472**	0.812**	0.453**	0.512**	0.579**
6	0.572**	0.442**	<b>0.223 (1.000)</b>	0.966**	0.877**	0.459**
7	0.436**	0.398 (0.011)*	<b>0.344 (0.336)</b>	0.780**	0.459**	<b>0.160 (1.000)</b>
8	0.889**	0.887**	0.623**	0.974**	0.599**	0.774**
9	0.723**	0.951**	0.529**	0.823**	0.781**	0.947**
10	0.921**	0.864**	0.709**	0.943**	0.884**	0.608**
11	0.488**	0.857**	0.654**	0.804**	0.386 (0.026)*	0.914**
12	0.467**	<b>0.219 (1.000)</b>	0.729**	0.540**	0.469**	0.776**
13	<b>0.366 (0.116)</b>	0.609**	<b>0.306 (0.841)</b>	0.960**	<b>0.216 (1.000)</b>	0.647**
14	0.864**	<b>0.217 (1.000)</b>	0.629**	0.383 (0.035)*	0.479**	<b>0.297 (0.920)</b>
15	0.511**	0.959**	0.491**	0.569**	0.471**	0.887**
16	0.654**	0.557**	0.706**	0.637**	0.705**	0.389 (0.022)*
17	0.715**	0.751**	0.934**	0.948**	0.894**	0.748**
18	0.659**	0.731**	0.575**	0.543**	0.595**	0.830**
19	0.695**	<b>0.326 (0.599)</b>	<b>0.310 (0.804)</b>	0.909**	0.456**	<b>0.183 (1.000)</b>
20	0.548**	0.417 (0.002)*	0.649**	0.877**	0.628**	0.590**
21	0.751**	0.861**	0.971**	0.504**	0.790**	0.5560**
22	0.598**	0.490**	0.871**	<b>0.228 (1.000)</b>	0.807**	0.483**
23	0.399 (0.010)*	0.707**	0.471**	0.576**	<b>0.308 (0.828)</b>	<b>0.359 (0.168)</b>
24	0.867**	0.918**	0.715**	0.992**	0.496**	0.751**
25	0.854**	0.652**	<b>0.142 (1.000)</b>	0.786**	0.447**	0.594**

Null hypothesis is no agreement ( $W = 0$ ).  $P$  value is written in the brackets if it is larger than 0.001. '\*' represents that null hypothesis is rejected, with  $P$  value  $< 0.05$ . '\*\*' represents that null hypothesis is rejected, with  $P$  value  $< 0.001$ . The bold data indicate that null hypothesis is accepted.

Table S6  
SUBREGIONAL HOMOGENEITY IN 3 SESSIONS (%)

	the left Putamen			the right Putamen		
	subregion 1	subregion 2	subregion 3	subregion 1	subregion 2	subregion 3
session 1	53.216	47.868	42.428	51.613	47.277	36.554
session 2	52.580	47.434	42.366	50.399	45.239	34.684
session 3	54.207	49.686	44.123	51.400	46.727	35.715

strength between subregions and reference ROIs. In session 1, the mean of  $r$  was 0.822, 0.433, 0.407 for subregion 1 to 3 in the left Putamen and 0.801, 0.478, 0.453 in the right Putamen. In session 2, the result was 0.813, 0.439, 0.408 for the left and 0.797, 0.460, 0.442 for the right. In session 3, the result was 0.816, 0.431, 0.414 for the left and 0.788, 0.449, 0.412 for the right. Correlation coefficients decreased sequentially with subregion index as expected and the results were quite consistent in 3 sessions.

Finally we investigated the validity of our parcellation. A pivotal marker of parcellation validity is subregional homogeneity of functional signals. Regional homogeneity is defined as the variance explained by the first principal component of the data, which is estimated using principle component analysis (PCA). Homogeneity was respectively computed for each subject and then averaged across all subjects to yield parcellation homogeneity for a subregion, referred to as subregional homogeneity. The results are shown in Table S6. There was no significant difference among 3 sessions and we observe that subregional homogeneity decreased along with subregion index. This shows the strong validity of our method when dealing with functional subregion parcellation.

## II. ROIS SELECTED FOR UBC DATA

The 48 ROIs used in the UBC data are provided in Table S7.

Table S7  
THE 48 ROIs SELECTED IN UBC DATA

No.	Name	No.	Name
1	Left-Cerebellum-Cortex	25	ctx_lh_G_temp_sup-Lateral
2	Left-Thalamus-Proper	26	ctx_lh_G_temp_sup-Plan_polar
3	Left-Caudate	27	ctx_rh_G_temp_sup-Lateral
4	Left-Putamen	28	ctx_rh_G_temp_sup-Plan_polar
5	Left-Pallidum	29	ctx_lh_G_temp_sup-Plan_tempo
6	Right-Cerebellum-Cortex	30	ctx_rh_G_temp_sup-Plan_tempo
7	Right-Thalamus-Proper	31	ctx_lh_G_and_S_cingul-Ant
8	Right-Caudate	32	ctx_rh_G_and_S_cingul-Ant
9	Right-Putamen	33	ctx_lh_G_cingul-Post-dorsal
10	Right-Pallidum	34	ctx_lh_G_cingul-Post-ventral
11	ctx_lh_precuneus	35	ctx_rh_G_cingul-Post-dorsal
12	ctx_rh_precuneus	36	ctx_rh_G_cingul-Post-ventral
13	ctx_lh_insula	37	R_M1
14	ctx_rh_insula	38	L_M1
15	ctx_lh_G_front_middle	39	R_S1
16	ctx_rh_G_front_middle	40	L_S1
17	ctx_lh_G_pariet_inf-Angular	41	R_SMA_proper
18	ctx_rh_G_pariet_inf-Angular	42	L_SMA_proper
19	ctx_lh_G_pariet_inf-Supramar	43	R_Pre_SMA
20	ctx_rh_G_pariet_inf-Supramar	44	L_Pre_SMA
21	ctx_lh_G_parietal_sup	45	R_PMd
22	ctx_rh_G_parietal_sup	46	L_PMd
23	ctx_lh_G_temp_sup-G_T_transv	47	R_PMv
24	ctx_rh_G_temp_sup-G_T_transv	48	L_PMv