

# Supplemental Material

## *Depression Uncouples Brain Hate Circuit*

Haojuan Tao<sup>#</sup>, Shuixia Guo<sup>#</sup>, Tian Ge<sup>#</sup>, Keith M. Kendrick<sup>#</sup>,

Zhimin Xue, Zhening Liu<sup>\*</sup>, and Jianfeng Feng<sup>\*</sup>

July 31, 2011

### **Community mining algorithm**

Network communities refer to groups of vertices within which the connecting links are dense but between which they are sparse. A community mining algorithm aims at finding all the communities from a given network. Distinct from the exiting studies in the literature, our former developed community mining algorithm described in [1] explored the notion of network communities and their properties based on the dynamics of a stochastic model naturally introduced. The relationship between the hierarchical community structure of a network and the local mixing properties of the stochastic model was established with the large-deviation theory.

Let  $(V, E)$  denote a network, where  $V = \{1, 2, \dots, n\}$  is a set of  $n$  vertices and  $E$  is a set of links. In this study, a vertex in the network corresponds to a brain region, and a

link between two vertices corresponds to the functional connectivity. Consider a stochastic process defined on the network, in which an agent freely walks from one vertex to one of its randomly selected neighbors along the links between them. Let  $X = \{X_t, t \geq 0\}$  denote the positions of the agent, and  $P(X_t = i, 1 \leq i \leq n)$  be the probability that the agent hits the vertex  $i$  after  $t$  steps. Since the next state of the agent only depends on its previous state, it follows that, for any  $i_t \in V$ ,

$$P(X_t = i_t | X_0 = i_0, X_1 = i_1, \dots, X_{t-1} = i_{t-1}) = P(X_t = i_t | X_{t-1} = i_{t-1}). \quad (1)$$

Therefore, this stochastic process is a discrete *Markov chain*. Furthermore, it is also *homogeneous* since its transition probability from vertex  $i$  to vertex  $j$  satisfies:

$$P(X_t = j | X_{t-1} = i) = p_{ij}, \quad (2)$$

where  $p_{ij}$  is a constant does not depend on  $t$ . In terms of the adjacent matrix of the network  $A = (a_{ij})_{n \times n}$ ,  $p_{ij}$  is defined by:

$$p_{ij} = \frac{a_{ij}}{\sum_j a_{ij}}, \quad (3)$$

or in matrix notation

$$\mathbf{P} = \mathbf{D}^{-1} \mathbf{A}, \quad (4)$$

where  $\mathbf{P} = (p_{ij})_{n \times n}$  is the transition probability matrix,  $\mathbf{D} = \text{diag}\{d_1, d_2, \dots, d_n\}$ , and  $d_i = \sum_j a_{ij}$ .

Now, let's consider the dynamics of the above stochastic model defined on the network. Intuitively, for a network with a community structure, its corresponding Markov chain will stick to some local mixing states or metastable states before reaches the global mixing state. The topological information related to network communities can thus be inferred from the

hitting and exiting times of the local mixing states. More specifically, for a well-formed community structure, each community is cohesive and easy to be locally mixed, corresponding to an early hitting time. On the other hand, few inter-community links lead to a late exiting time, or equivalently a long global mixing time. Let  $s_1$  denote the global mixing state and  $s_n, \dots, s_2, s_1$  be a sequence of local mixing states reaching to global mixing state. Let  $T_s^{ext}, 1 \leq s \leq n$  be the exiting time of the  $s$ -th local mixing state. Following the main result of *Large-deviation theory* developed by Varadhan, Freidlin and Wenzel [2], all local exiting times can be estimated by the spectrum of the Markov generator  $Q = I - P$ , where  $I$  is the identity matrix. Specifically,

$$T_s^{ext} \approx 1/\lambda_s, \quad s = 1, \dots, n, \quad (5)$$

where  $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  are the  $n$  non-negative real-valued eigenvalues of the generator  $Q$ . The hitting time of the  $s$ -th local mixing state can be reasonably estimated by the exiting time of the  $(s + 1)$ -th local mixing state, that is

$$T_s^{hit} = T_{s+1}^{ext} \approx 1/\lambda_{s+1}. \quad (6)$$

Therefore, any community structure of the network can be captured by the spectral properties of the Markov generator  $Q$ . For a well-formed  $K$ -community structure, the cohesion of each community indicates the quantity

$$C_K = T_K^{hit} \approx 1/\lambda_{K+1} \quad (7)$$

is small, while a good separability of the community structure leads to the measure

$$S_K = T_K^{ext} - T_K^{hit} \approx 1/\lambda_K - 1/\lambda_{K+1} \quad (8)$$

to be large. Hence, the sequence

$$CQ_k = \frac{C_k}{C_k + S_k} = \frac{\lambda_k}{\lambda_{k+1}}, \quad 1 \leq k \leq n - 1, \quad (9)$$

can be defined as the spectral signature of a network. It is clear that  $0 \leq CQ_k \leq 1$ . A small  $CQ_K$  implies a better  $K$ -community structure with better cohesion as well as separability.

The number of well-formed communities can be inferred from

$$K = \operatorname{argmin}_k CQ_k. \quad (10)$$

From the above analysis, the connection between the community structure of a network and the spectrum of its corresponding Markov generator has been uncovered. Many questions related to the characterizing the community structure of a network can be converted to observing and inferring from its spectral signature. See [1] for a more complete discussion.

Many different strategies can be used to implement the mining of the  $K$ -community structure if some  $CQ_K$  is lower than a threshold. We follow the efficient implementation proposed in [1]. This algorithm does not require calculating the eigenvalues/eigenvectors and multiplying the transition matrix, which is thus suitable for very large-scale networks. More detailed implementation of the scalable algorithm is described in [1].

Table S1: Treatment details of RMDD patients.

Treatment	Number of case
venlafaxine	6
citalopram	3
paroxetine hydrochloride	2
fluoxetine hydrochloride	1
mirtazapine	2
amitriptyline	2
amitriptyline hydrochloride + sodium valproate	1
amitriptyline + risperidone	1
paroxetine hydrochloride + quetiapine	1
venlafaxine + lithium carbonate	1
Imipramine hydrochloride + sodium valproate	1
Unavailable	3

All FEMDD patients were treatment naïve. None of the RMDD patients receiving combination treatments had schizoaffective disorder or co-morbidity with any Axis II disorders. Combination treatments were being used due to their claimed efficacy for treatment resistant depression [3–5].

## References

- [1] Yang B, Liu J, Feng J (2010) On the spectral characterization and scalable mining of network communities. *IEEE Transactions on Knowledge and Data Engineering* 99: <http://doi.ieeecomputersociety.org/10.1109/TKDE.2010.233>.
- [2] Freidlin M, Wentzell A (1998) *Random perturbations of dynamical systems*, volume 260. Springer Verlag.
- [3] Anderson I (2003) Drug treatment of depression: reflections on the evidence. *Advances in Psychiatric Treatment* 9: 11.
- [4] Mahmoud R, Pandina G, Turkoz I, Kosik-Gonzalez C, Canuso C, et al. (2007) Risperidone for treatment-refractory major depressive disorder. *Annals of internal medicine* 147: 593.
- [5] El-Khalili N, Joyce M, Atkinson S, Buynak R, Datto C, et al. (2010) Extended-release quetiapine fumarate (quetiapine XR) as adjunctive therapy in major depressive disorder (MDD) in patients with an inadequate response to ongoing antidepressant treatment: a multicentre, randomized, double-blind, placebo-controlled study. *The International Journal of Neuropsychopharmacology* 13: 917–932.

**Table S2: Grey/white matter volumes of the ROIs involved in “hate circuit” for both patients (p) and normal controls (n).**

	INS.L	PUT.L	SFGdor.L
Sum of white matter (p / n / p-value)	2465.6 / 2514.2 / 0.4707	2128.3 / 2279.3 / 0.0951	3932.0 / 4041.9 / 0.6046
Mean of white matter (p / n / p-value)	0.5717 / 0.5697 / 0.7680	0.6057 / 0.6089 / 0.6973	0.2682 / 0.2678 / 0.9854
Sum of grey matter (p / n / p-value)	7475.9 / 7394.4 / 0.4573	3206.2 / 3037.2 / 0.0886	8748.7 / 8605.1 / 0.6756
Mean of grey matter (p / n / p-value)	0.6241 / 0.6196 / 0.4964	0.6255 / 0.6136 / 0.1252	0.4452 / 0.4478 / 0.8334
	INS.R	PUT.R	SFGdor.R
Sum of white matter (p / n / p-values)	1905.4 / 1909.2 / 0.9464	2673.4 / 2801.3 / 0.2211	4996.5 / 5198.1 / 0.4594
Mean of white matter (p / n / p-value)	0.5549 / 0.5463 / 0.2810	0.6130 / 0.6101 / 0.6852	0.3083 / 0.3063 / 0.9148
Sum of grey matter (p / n / p-value)	7265.7 / 7247.1 / 0.8556	2974.6 / 2845.5 / 0.2005	9956.9 / 9803.9 / 0.6811
Mean of grey matter (p / n / p-value)	0.6129 / 0.6095 / 0.6259	0.5960 / 0.5869 / 0.2000	0.4614 / 0.4642 / 0.8074

**Table S3: Grey/white matter volumes of the ROIs involved in risk/action circuit for both patients (p) and normal controls (n).**

	IPL.L	PreCG.L	IFGperc.L
Sum of white matter (p / n / p-value)	1593.2 / 1747.6 / 0.3187	4900.5 / 5236.1 / 0.1490	1389.4 / 1462.6 / 0.3248
Mean of white matter (p / n / p-value)	0.2510 / 0.2576 / 0.6276	0.2889 / 0.2868 / 0.8579	0.2996 / 0.3005 / 0.9156
Sum of grey matter (p / n / p-value)	9155.2 / 9090.3 / 0.7603	9343.1 / 9326.7 / 0.9470	3208.0 / 3149.9 / 0.4031
Mean of grey matter (p / n / p-value)	0.5382 / 0.5366 / 0.8743	0.4399 / 0.4323 / 0.4295	0.4744 / 0.4667 / 0.3213

**Table S4: Grey/white matter volumes of the ROIs involved in emotion/reward circuit for both patients (p) and normal controls (n).**

	ORBsup.L	ORBsupmed.L	ORBsup.R	ORBsupmed.R
Sum of white matter (p / n / p-value)	1106.2 / 1184.4 / 0.2500	273.4 / 309.5 / 0.2314	1196.0 / 1248.7 / 0.4511	499.6 / 542.6 / 0.2489
Mean of white matter (p / n / p-value)	0.3035 / 0.3171 / 0.0992	0.2120 / 0.2263 / 0.4281	0.2882 / 0.3048 / 0.0615	0.2944 / 0.3133 / 0.2541
Sum of grey matter (p / n / p-value)	3391.4 / 3331.8 / 0.2850	2909.4 / 2870.3 / 0.4785	3312.9 / 3246.9 / 0.2729	3484.9 / 3421.8 / 0.2735
Mean of grey matter (p / n / p-value)	0.5128 / 0.5120 / 0.9194	0.5663 / 0.5657 / 0.9531	0.4973 / 0.4990 / 0.8201	0.5748 / 0.5762 / 0.8850