Partitioning the Firing Patterns of Spike Trains by Community Modularity

Hu Lu(myluhu@126.com)[a,b]
[a] School of computer science and communication engineering, 301 Xuefu Road
Jiangsu University, Zhenjiang 212003 China
[b] School of Computer Science, 825 Zhangheng Road
Fudan University, Shanghai 201203 China

Xing Hao Huang(m15251706875_1@163.com)
School of computer science and communication engineering, 301 Xuefu Road
Jiangsu University, Zhenjiang 212003 China

Yu Qing Song(yqsong@ujs.edu.cn)
School of computer science and communication engineering, 301 Xuefu Road
Jiangsu University, Zhenjiang 212003 China

Hui Wei(weihui@fudan.edu.cn)
School of Computer Science, 825 Zhangheng Road
Fudan University, Shanghai 201203 China

Abstract

The traditional clustering method utilized to partition neuronal firing patterns, including K-means and FCM algorithm, require specification of clusters numbers as priori knowledge. A new approach to analyze groups of firing patterns of neuronal spike trains based on community structure partitioning analysis and modularity function Q is examined in this study. This approach is able to automatically identify the optimal number of groups in neuronal firing patterns, realizing the true unsupervised analysis, and identify groups of neurons with similar firing patterns. The method was tested on a surrogate data set and a testing data set with firing patterns known in advance. The method was also applied to multi-electrode recording spike trains with previously unknown patterns. Results indicate this method can effectively self-determine number of pattern groups and locate firing patterns of neuronal populations based on community modularity Q.

Keywords: Community structure; Modularity; Clustering; Neuronal firing pattern

Introduction

A pressing neuroscience question exists as brain neuron encoding related to external information for learning, memory and other cognitive tasks remains unknown (Brown et al., 2004). Neurons transmit information through the form of action potentials. The distributions of action potentials firing at different times and spaces is referred to as spatial-temporal firing patterns. The study found when the same stimulation currents were injected into the soma of a cortical neuron in vitro the neuron may produce similar firing patterns (Fellous et al., 2004.). These patterns are very likely the neural basis of information representation and processing (Pillow et al., 2008). Each cortical area is composed of a large number of neurons, thus the study of brain functions has transferred from a single neuron to a large neuronal populations. The development of extracellular recording techniques has allowed a number of neurons to be recorded simultaneously by utilizing the multi-electrode recording (Buzsáki, 2004). Analysis of the intrinsic firing relationship between these neurons may reveal an existing structural or functional connectivity between the neurons (Jarrell et al., 2012). Studies have revealed brain modularity when performing cognitive tasks and that neural firing activities within the modules are correlative with several neurons exhibiting some type of similar firing patterns in a module (Schneidman et al., 2006). Discovering and analyzing neural systems holds key significance in revealing brain patterns (Lindsey et al., 1997).

Many computational and methodological challenges exist to study these firing patterns (Brown et al., 2004). Clustering methods such as the K-means clustering approach, FCM clustering algorithm and spectral clustering method have also been proposed to analyze the firing patterns (Toups & Tiesinga, 2006; Paiva et al., 2007). Clustering methods must specify the number of pattern groups in advance while the number of groups in real spike trains is unknown, limiting the practical usefulness of these methods. A new spike trains communities finding method based on the principle of community structure detection has been proposed; however, this method also utilized K-means++ clustering algorithm in resulting similarity matrix eigenvectors (Humphries, 2011). Newman (2011) applied the clustering algorithm described by Humphries (2011) to identify neuron communities during a reach and grasp task (Newman et al., 2011). Community structure detection algorithm is a graph partitioning method widely utilized in social network analysis. The recognition of community structure is a NP-hard problem with many new methods proposed to solve the problem. A modularity function Q has been proposed to optimize the community structure partitioning problem without requiring knowledge of community numbers in advance and to find the optimal partitioning of community structure by maximizing the modularity function Q (Le & Hankin, 2011).

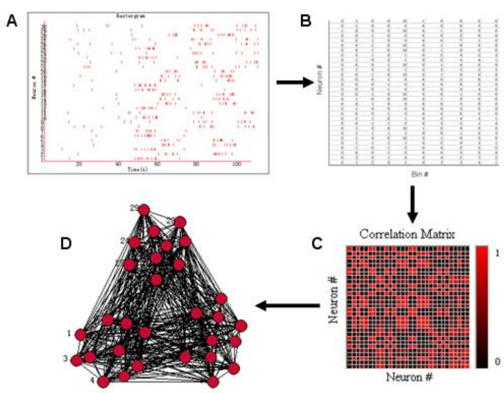

Figure 1: Schematic overview of neuronal functional network construction. (A) The original spike trains (line represents neuronal spikes). (B) Spike vector groups by binning the spike trains. (C) Correlation matrix by calculating correlations between two neuronal spike trains. (D) Retain all edges and construct a neuronal functional network. Physical locations of neurons in this Figure are random as study is limited to the functional connections between neurons.

A new neuronal firing pattern finding method based on optimization of modular function Q was proposed in this study. Neuronal functional networks constructed from neuronal spike trains can be divided into different modules through the Ncut clustering algorithm. Optimal division of spike trains was obtained by searching for the maximum value of Q using this method. Number of pattern groups was determined automatically by dividing the corresponding neuronal spike trains into similar groups, realizing the true unsupervised clustering of firing patterns. The K-means clustering algorithm was not utilized as with other existing firing pattern finding methods, thus this method not only found the similar spatial-temporal firing groups, but also divided neuronal functional networks into different community structures (Gansel & Singer, 2012; Toups & Tiesinga, 2006). Results revealed the technique was effective when tested on a surrogate data set and a testing data set with pattern structures known in advance. Finally, application to in vivo multi-electrode recording neuronal spike trains data set was performed with the firing patterns hidden and undiscovered.

Materials and Methods

Surrogate data set

Proposed methods in this study were tested utilizing spike trains data sets with known neuronal firing patterns. A spiking neuronal model was utilized to generate a data set containing the three similar firing pattern groups (Izhikevich, 2003). The number of neurons was 30. The known neuronal community structure model included 3 communities and was constructed so neuronal firing patterns in each community were similar. Each community structure contained 10 neurons as presented in the neuronal raster plot (Fig.1A). Although neuronal firing patterns in this data set are simple, pattern clustering analysis methods are required to accurately locate firing patterns.

Testing data set

A testing data set, created by Fellous (2004), composed of 90 spike trains and containing 3 pattern groups, was also utilized. A spike train simulates common patterns across trials of a single neuron. The corresponding neuronal raster plot is depicted in Fig.3A. Noise and jitter were added to the spike trains causing the spike trains to exhibit irregular firing. The data set can be obtained from sharing website (available from http://cnl.salk.edu/fellous/data/JN2004data/data.html) and can be used to test various firing pattern finding algorithm.

In vivo recording data set

Data sets above were utilized to introduce a new modularity-based method for finding firing patterns in multi spike trains. This method was applied to real in vivo spike trains data as neuronal spike trains were analyzed by utilizing the multi-electrode arrays recorded from behavioral rats performing the Y-maze working memory task. Male Spraque-Dawley rats were used. Surgical procedures were performed under sodium pentobarbital anesthesia. Microelectrode array was made by 16 microelectrodes. Microelectrode array was implanted in the left mPFC. Recording signal acquisition system used the Plexon multichannel acquisition processor system. Spike trains of neurons were obtained by the Plexon Offline Sorter. All experimental procedures were approved and monitored by the Ethical Committee of Animal Experiments at the Institute of Neurobiology, Fudan University (Shanghai, China). Trial processes were selected randomly at a time period. The time length was 42s, between 3082s and 3124 s of the whole recording time. The data set included 82 task trials with 20 neuronal spike trains analyzed.

Methods

Different types of neuronal spike trains data sets were collected including multiple neuronal spike trains recorded from a trial task and neuronal spike trains consisting of several trials of single neurons.

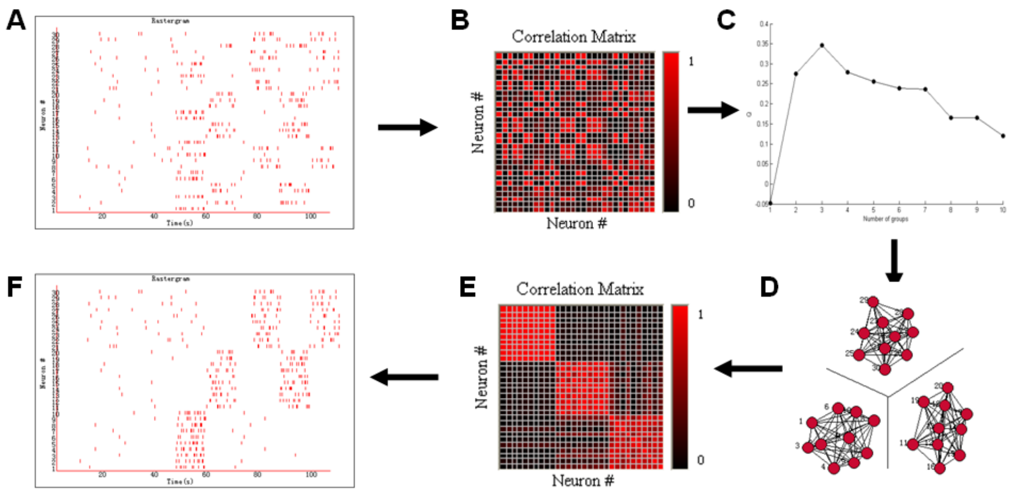

Figure 2: Utilizing the proposed method to locate firing patterns in the data set presented in Fig. 1A. (A) Represents raster plot of 30 neuronal spike trains. (B) Represents correlation matrix between neurons. (C) Distribution of Q values utilizing the Newman modularity is represented, deriving maximum value of Q when the number of communities equals 3. (D) According to the Q, we divided 30 neurons into 3 communities only give out the connections between the inter-communities. (E) Correlation matrix in Fig. 2B was sorted according to the communities, resulting in a new matrix, indicating obvious modular structure compared to Fig. 2B. (F) Raster plot of spike trains with sorting according to similar firing patterns in Fig. 2A and time window size set to 10s.

Construction of the functional connections between neuronal spike trains based on the correlation of neuronal firing were first required. Figure 1 depicts a raster plot of multi-neuronal spike trains with each mark representing the firing of a neuronal action potential. A line of marks represents a neuronal action potential sequence. Calculating correlations between pairs of neurons is the first step to constructing functional connections of neurons. The procedure was realized, in this study, by binning the spike rastergram into non-overlapping, short time windows (also refers to bins). Multiple neuronal spike trains were converted to vector groups by sizing the bin utilizing parameter δ_t and counting the number of spikes in each bin. The element of vector represents the number of spikes in each bin, as indicated in Fig. 1B. Pearson correlation coefficient r was utilized to calculate the correlation between the two spike trains:

$$r = \frac{\sum_{i=1}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}\left(x_i - \bar{x}\right)^2 \cdot \sum_{i=1}(y_i - \bar{y})^2}} \quad (1)$$

Where x_i represents the value of the ith bin of the xth spike trains. $\bar{x}$ represents the average values of all bins of the xth spike trains. The value of r is between -1 and 1.

Pearson correlation has been widely utilized to calculate correlations between brain signals in the study of brain functional networks. Focus of this research includes functional connectivity strength between neurons, regardless of the direction of functional connections, so the connectivity weights between neurons were defined as absolute values of Pearson correlation coefficient,

$$R_{ij} = |r| \quad (2)$$

The undirected, weighted neuronal functional connectivity graph was built based on the weighted correlation matrix R (Liang & Zhang, 2011). Weighted network is converted into binary by thresholding to simplify the analysis in some studies. Different group partitioning methods, based on the correlation matrix R, may be utilized to divide matrix into different clusters. Standard data clustering techniques, such as K-means may be utilized to find clusters within the comparison matrix; however, K-means requires specific clustering numbers and cluster centers in advance. The Ncut spectral clustering algorithm proposed by Shi and Malik(2000) and utilized in this study, is a graph-based partitioning method (Shi & Malik, 2000). Ncut algorithm produces a comparatively superior performance and has been applied in the brain functional networks of fMRI (Van et al., 2008; Shen et al., 2010).

Ncut algorithm application obtains a neuronal functional network partition. Evaluation of the partition quality and derivation of the number of groups hidden in spike trains was achieved utilizing the modular function Q (Newman, 2004; Leicht & Newman, 2008). When value Q reached maximum, a corresponding number of communities reflected the number of neuronal firing pattern groups. A weighted network modular function Q is defined as follows:

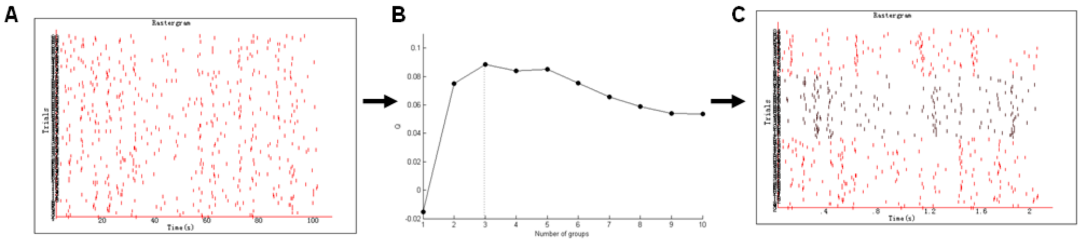

Figure 3: (A) Original raster of spike trains. (B) Distribution of different modularity Q calculated from different number of pattern groups. (C) Divide 90 trials into three firing pattern groups with time window size set to 0.2s.

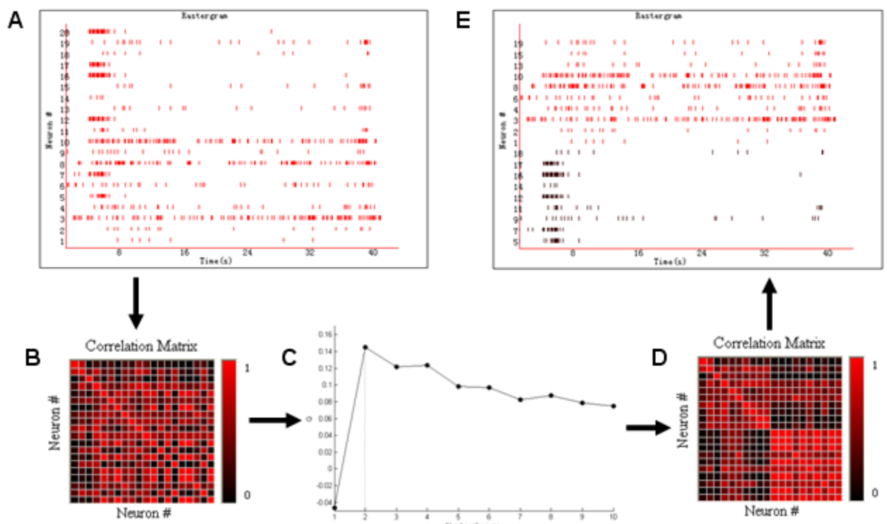

Figure 4: The time window size is set to 4s. (A) original raster plot of spike trains (B) correlation matrix of spike trains (C) distribution of different Q value (D) correlation matrix sorted according to two similar patterns (E) sorted spike trains with two similar firing patterns.

$$Q = \frac{1}{l}\sum_{i,j\in N}\left[w_{ij} - \frac{k_i k_j}{l}\right]\delta_{m_i,m_j} \quad (3)$$

Where l is the sum of values in the weighted matrix.

k_i and k_j are the degree of node i and node j, respectively. δ_{ij} is the Kronecker delta function, which equals 1 when node i and node j are in the same community and 0 when elsewhere. The community partitioning analysis process was realized by performing the following two steps alternately. A single partitioning was obtained in the first step utilizing Ncut algorithm when the number of communities was assigned to 2. The value of Q was calculated corresponding to this partitioning in the second step. One was then added to the number of communities and the above two steps repeated until the number of communities equaled the number of nodes. A distribution of Q values and corresponding partitioning was obtained and maximum modularity Q and corresponding partitioning then derived.

Results

Surrogate data set

The study method automatically identified similar firing patterns of neurons by utilizing the community partitioning algorithm, including the number of pattern groups and corresponding firing patterns, without prior knowledge of patterns contained in the data sets. Effectiveness of this method is illustrated as partitioning algorithm on a surrogate data set is tested. Figure 2 presents a description of the implementation procedure.

Prior knowledge of pattern groups does not exist in the spike trains although there were obvious firing patterns among 30 neuronal spike trains (Fig. 2). The initial structures between neurons were in disorder (Fig. 2A), thus the number of groups and the firing patterns contained in the spike trains remained unknown without applying the clustering analysis method. Using the proposed community partitioning method, the number of pattern groups was automatically identified and equaled 3 (Fig. 2C). Ncut

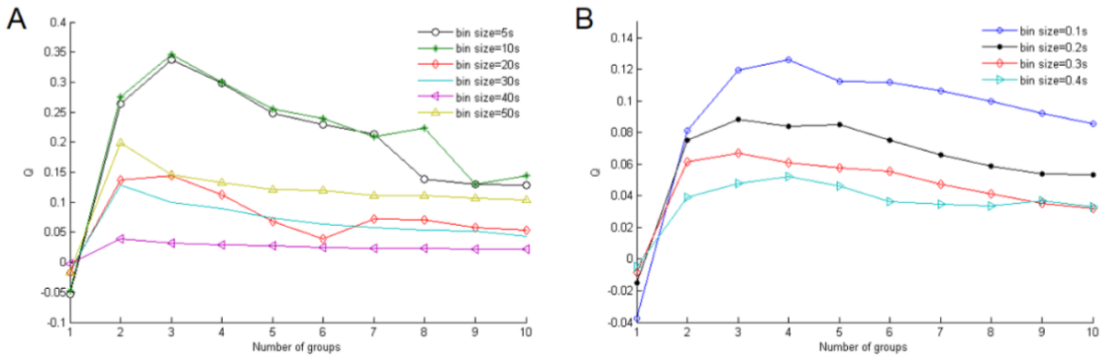

Figure 5: Distribution of different Q with different bin size t parameter selected (A) Results of surrogate data set when bin size equals 5s and 10s and optimal number of groups 3 are identified. Optimal number of pattern groups is 2, otherwise. (B) Results of testing data set.

partitioning method was then utilized to divide the 30 neurons into 3 groups according to the pattern similarity (Fig. 2D) while the original spike trains were sorted resulting in a new raster plot (Fig. 2F). Compared to Fig. 2A, Fig. 2F exhibits three obvious spatial-temporal firing patterns and Fig. 2B and Fig. 2E represent the correlation coefficient matrix of before sorting and after sorting, respectively.

Testing data set

A different data sent created by Fellous (2004) and containing a known number of groups equaling 3 was also tested. Figure 3 presents experimental results with Fig. 3A representing spike trains containing multiple trials of a neuron. Neuroscientists utilize the clustering analysis method to discover the firing pattern in the spike trains. Spike trains were divided into different communities and corresponding Q values calculated. Maximum modularity value of Q was obtained when the number of communities equaled 3. Ninety trials of spike trains were divided into 3 groups utilizing the Ncut algorithm as represented in different colors (Fig. 3C). Number of groups known in advance is not required with this method as opposed to the K-means clustering method, thus pattern discovery occurs unsupervised.

In vivo recording data set

The framework was applied to the recording spike trains in vivo and a trial process of spike trains was selected with a time period of 40 s and the number of neurons at 20.

The spike trains data set was recorded from the prefrontal cortex of a rat chronically implanted with multi electrode arrays (see methods). Structure of the data set, including the number of pattern groups, was unknown in advance, thus the proposed analysis method was applied to detect assemblies. Community modularity Q was utilized for this study as Q values were calculated when the number of communities varied. The maximum Q was found when the number of communities equaled 2 (Fig. 4C). The 20 neuronal spike trains were then divided into two groups and two firing patterns identified. (Fig. 4E)

Parameters selection

Bin size *t* of spike trains is a parameter utilized in this study as different *t* will affect the number of spikes in each bin. The impact of bin size on experimental results was analyzed as a series of different bin sizes was utilized to construct neuronal functional networks. Framework proposed in this study was then applied to detect the optimal number of groups.

Results of the two data sets were compared with the optimal number of groups known in advance. Different parameter influence on identification of the optimal number of groups was investigated revealing that different bin size affected optimal number of groups. Research for parameter selection is difficult. There is no more researches to show how to select the parameters reasonably.

Conclusion

A new method to detect multi neuronal firing patterns has been provided with the overall algorithm based on the Pearson correlation coefficient matrix, Ncut partitioning algorithm and modularity function Q. The algorithm automatically determines the number of pattern groups contained in spike trains by comparing the value of Q. Based on the maximum value of Q corresponding to potential optimal number of pattern groups, the firing of multi spike trains can be divided into different firing patterns without a priori knowledge about the number of groups or structure of spike trains. However, modularity function Q encounters the problem of resolution limit, which cannot identify some modules smaller than a certain scale. In future research, we will extend community structure partitioning methods that do not depend on modularity optimization (Lu & Wei, 2013).

Acknowledgments

We are grateful to Dr. Shengtao Yang（Fudan University）for providing us with multiple electrodes neural spike trains data. This study was supported by the National Natural Science Foundation of China (Project No. 61375122), China Postdoctoral Science Foundation (Project No. 2014M551324).Scientific Research Foundation for Advanced Talents of Jiangsu University (Project No. 14JDG040).

References

Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature neuroscience*, *7*(5), 456-461.

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nature neuroscience*, *7*(5), 446-451.

Fellous, J. M., Tiesinga, P. H., Thomas, P. J., & Sejnowski, T. J. (2004). Discovering spike patterns in neuronal responses. *The Journal of Neuroscience*, *24*(12), 2989-3001.

Gansel, K. S., & Singer, W. (2012). Detecting multineuronal temporal patterns in parallel spike trains. *Frontiers in neuroinformatics*, *6*.

Humphries, M. D. (2011). Spike-train communities: finding groups of similar spike trains. *The Journal of Neuroscience*, *31*(6), 2321-2336.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on neural networks*, *14*(6), 1569-1572.

Jarrell, T. A., Wang, Y., Bloniarz, A. E., Brittin, C. A., Xu, M., Thomson, J. N., ... & Emmons, S. W. (2012). The connectome of a decision-making neural network. *Science*, *337*(6093), 437-444.

Leicht, E. A., & Newman, M. E. (2008). Community structure in directed networks. *Physical review letters*, *100*(11), 118703.

Le Martelot, E., & Hankin, C. (2011, October). Multi-scale Community Detection using Stability as Optimisation Criterion in a Greedy Algorithm. In *KDIR* (pp. 216-225).

Liang, Z., King, J., & Zhang, N. (2011). Uncovering intrinsic connectional architecture of functional networks in awake rat brain. *The Journal of Neuroscience*, *31*(10), 3776-3783.

Lindsey, B. G., Morris, K. F., Shannon, R., & Gerstein, G. L. (1997). Repeated patterns of distributed synchrony in neuronal assemblies. *Journal of Neurophysiology*, *78*(3), 1714-1719.

Newman, M. E. (2004). Analysis of weighted networks. *Physical Review E*, *70*(5), 056131.

Newman, G. I., Aggarwal, V., Schieber, M. H., & Thakor, N. V. (2011, August). Identifying neuron communities during a reach and grasp task using an unsupervised clustering analysis. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 6401-6404). IEEE.

Paiva, A. R., Rao, S., Park, I., & Principe, J. C. (2007, August). Spectral clustering of synchronous spike trains. In *Neural Networks, 2007. IJCNN 2007. International Joint Conference on* (pp. 1831-1835). IEEE.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., & Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, *454*(7207), 995-999.

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, *440*(7087), 1007-1012.

Shen, X., Papademetris, X., & Constable, R. T. (2010). Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *Neuroimage*, *50*(3), 1027-1035.

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *22*(8), 888-905.

Toups, J. V., & Tiesinga, P. H. (2006). Methods for finding and validating neural spike patterns. *Neurocomputing*, *69*(10), 1362-1365.

Van Den Heuvel, M., Mandl, R., & Pol, H. H. (2008). Normalized cut group clustering of resting-state FMRI data. *PloS one*, *3*(4), e2001.

Hu, L., Hui, W., (2012). Detecting community structure in networks based on community coefficients. *Physica A: Statistical Mechanics and its Applications. 391*, 6156-6164.