
STAT768 Final Project Report: Graph Structure Smoothing via wavelet on non-Euclidean Space for Group Analysis of Networks

Won Hwa Kim
`wonhwa@cs.wisc.edu`

Abstract

Human brain is a very complicated network system connected by billions of neurons. Recent development of data acquisition techniques in medical applications has enabled to obtain information of the actual physical connectivity of the brain network, such as diffusion tensor image (DTI). However, due to the noise and limited number of subjects for the data, it is generally a difficult task to find group differences in the raw data. Smoothing by a Gaussian filter is known to enhance the Gaussianity in the signal, therefore improves the result from group analysis. However, existing smoothing methods on graphs smoothes the signal that are defined on the nodes, not the edges, and it is ambiguous how one should smooth the signals defined at the edge level. In this report, I introduce a new framework of enhancing the underlying true signal by smoothing the given data using wavelets on graphs that are defined on the edges of a network. Accommodating the concept of line graph brings a new domain for analysis of network edges, where the edge weights in the original network are defined as a signal on the nodes of a line graph. In the experiment, it is shown that after applying this framework on a human brain network dataset, significant group differences shows up between bipolar and control group, where we could not detect anything using the raw data.

1 Introduction

A brain is a such complicated system that there is much more to study on its process and functions despite abundant previous researches. There has been lots of studies regarding analysis of signal defined on the brain surface such as cortical thickness or deformation of the brain structure along the development of brain data acquisition techniques such as CT, MRI, fMRI and so on. When it comes to the studies on brain network, people often use correlation to build the network. However, it would be meaningless to trying to find correlation between different brain regions when these regions do not have anatomical connections. With advancement of brain image acquisition techniques, we can obtain the physical connectivity information of the brain as Diffusion tensor imaging (DTI), considering that a brain is a natural network system with billions of neurons.

The feature of dataset is critical when it comes to detecting group differences involving certain biomarkers. If the given data is not related to the characteristic of certain illness or disorder, it would be impossible to find the group differences among the subjects. Constructing brain network using correlation has a long history [16, 18, 14], however, it would be meaningless to justify connection between different regions of brain by high correlation when they are anatomically disconnected. Since the human brain that is structured as a network of neurons, it will be much more feasible and intuitive to find the differences using the physical connectivity information assuming that any abnormal symptom will show some changes in the brain network function.

Throughout this project, I introduce a new framework to filter data that are defined on the edges of a network using wavelet. In the experiment section, this framework is used to find group differences between a group from bipolar and a group from controls, by enhancing the underlying signal. Orig-

inally, applying simple two sample t -test on the raw dataset detects nothing, whereas the same test on the processed data using this procedure finds statistically significant differences that passes multiple comparison correction. Similar research has been done in [5, 13, 7, 4], however, these methods focus on the analysis at the node level, and do not include any multiple-error correction procedure.

Wavelet has been a very powerful tool to analyze signals not only on 1-D signal or 2-D images, but also on brain imaging as well. Many studies analyze EEG signal using wavelet [2], and also construct wavelet on a sphere to analyze the signals that are defined on brain surfaces [6, 12, 1, 17]. However, these methods require Euclidean domain to construct wavelets, and do not define wavelets directly on the graph domain. Furthermore, they are not directly applicable to the analysis of connectivity, where all the necessary information lie on the edges of the network.

Contributions. In this report, we introduce a new framework for graph structure smoothing. Here, the graph smoothing refers to smoothing the function defined on the edges instead of vertices, without losing the topology of the original network. We transform the original network to a new domain, and perform smoothing there using wavelets on graphs. Furthermore, we show the power of this method by performing group analysis on a brain network data.

- (i) I introduce how this framework significantly enhances the statistical power of group analysis by performing a group test on a brain network dataset. In this dataset, performing a group test on the raw dataset shows some low p -values, but none of them survive multiple correction comparison. After processing the data using our framework, the signals showing group differences become much stronger and we are able to detect those connections that are significant, which is the main contribution of this project.
- (ii) We introduce a domain, line graph, which is useful for analyzing the network structure. In this domain, the measurement at the connection level is defined as a signal on each voxel, therefore it is convenient to apply extensive methods for analyzing signals on each node of the line graph domain.

2 Construction of wavelet in Euclidean space

The well-known Fourier series represents a function by the superposition of sine and cosine basis and provides an optimal domain to perform signal analysis in terms of frequency. Wavelet transform is conceptually similar to the Fourier transform in that they can be used to extract information in the frequency domain, however, it uses certain shape of oscillating function as a basis function with finite duration instead of the sine and cosine basis with infinite duration. The advantage of wavelet transform over Fourier transform comes from this difference of basis function. While the Fourier transform is localized in frequency only, wavelet can be localized in both time and frequency [11].

The traditional construction of wavelet is defined by a mother wavelet function ψ and a scaling function ϕ . While ψ serves as band-pass filters of different scales in the frequency domain, ϕ operates as a low-pass filter covering the low frequency components of the signal which are not covered by the band-pass filters. When these band-pass filter functions are transformed back to the original domain by the inverse transformation and translated, it becomes a localized oscillating function with finite duration, providing local support in the original domain [15]. This contrasts to the Fourier series representation of a short pulse suffering from ringing artifacts (e.g., Gibbs phenomenon) because of the non-locality (infinite duration) of sin function. Formally, the wavelet function ψ on x is a function defined by two parameters, the scale parameter s and translation parameters a

$$\psi_{s,a}(x) = \frac{1}{a} \psi\left(\frac{x-a}{s}\right). \quad (1)$$

Change in s varies the *dilation* of the wavelet, and together with a *translation* parameter a , provide a way to approximate a signal in harmonics using wavelet expansion. The function $\psi_{s,a}(x)$ forms a basis for the signal and can be used with other basis at different scales to decompose a signal. The wavelet transform of a signal $f(x)$ is defined as the inner product of the wavelet basis $\psi_{s,a}$ and signal $f(x)$, and is represented as

$$W_f(s, a) = \langle f, \psi \rangle = \frac{1}{a} \int f(x) \psi^*\left(\frac{x-a}{s}\right) dx, \quad (2)$$

where $W_f(s, a)$ is the wavelet coefficient at scale s and at location a , and the function ψ^* represents the complex conjugate of ψ . Such a transform is invertible, meaning that the original signal $f(x)$

can be reconstructed from $W_f(s, a)$ and basis function without loss of information. The inverse transformation is

$$f(x) = \frac{1}{C_\psi} \iint W_f(s, a) \psi_{s,a}(x) da ds \quad (3)$$

where $C_\psi = \int \frac{|\Psi(j\omega)|^2}{|\omega|} d\omega$ is known as the *admissibility condition constant*, Ψ is the Fourier transform of the wavelet [15], and ω denotes the frequency domain.

The scale parameter s controls the dilation of the basis. By varying s , one can produce various short to long basis functions. This is the most powerful property of wavelet, since short basis functions corresponding to high frequency components are useful to isolate signal discontinuities and long basis functions obtain detailed analysis in frequency. Moreover, unlike the single set of basis functions (sine and cosine) in the Fourier transform, wavelet transforms can have an infinite set of possible basis functions depending on which type of filters are needed. Note that the wavelet transformation mentioned so far is not applicable to non-uniform topologies such as graphs or networks. In the later section, we define analogues of the wavelet transformation in non-Euclidean space without losing these properties.

3 Wavelets in non-Euclidean space

Recent work in harmonic analysis [8] provides wavelet basis on structured data which expresses in a wide spectrum of frequencies. The solution in [8] relies on spectral graph theory and graph Fourier transform to derive a spectral graph wavelet transform (SGWT), and Cholesky approximation to relax computational burden. It is shown that SGWT formalization preserves the localization properties at fine scales as well as other wavelets specific properties.

In deriving Wavelet expansions of a signal defined on arbitrary graphs, the first question is how to define *scales* on a domain where the space is non-Euclidean. For instance, consider a function $f(n)$ defined on a vertex n of a graph. Here, defining $f(sn)$ for a scaling parameter s is conceptually difficult due to the irregularity of the domain. This problem can be tackled by defining a transformation operator to the dual domain using the graph Fourier transformation.

Let a graph $G = \{V, E, \omega\}$ be an undirected graph with a vertex set V with N vertices, an edge set E and corresponding edge weight $\omega \geq 0$. Adjacency matrix \mathbf{A} of G is given as a $N \times N$ matrix whose elements a_{ij} are given as the edge weight ω_{ij} if i th and j th nodes are connected. Degree matrix \mathbf{D} is computed as a $N \times N$ diagonal matrix whose i th diagonal is $\sum_j \omega_{ij}$, the sum of edge weights connected to the i th node. Then the graph Laplacian is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \quad (4)$$

This graph Laplacian \mathbf{L} is a positive semi-definite matrix, which shows the smoothness of the graph structure. Then the complete orthonormal basis χ_l and eigenvalues λ_l , $l \in \{0, 1, \dots, N-1\}$ can be obtained from the graph Laplacian, which forms the basis for the graph Fourier transformation. Using these basis, the forward and inverse graph Fourier transformation are defined using the eigenvalues and eigenvectors of \mathbf{L} as,

$$\hat{f}(l) = \langle \chi_l, f \rangle = \sum_{n=1}^N \chi_l^*(n) f(n) \quad (5)$$

$$f(n) = \sum_{l=0}^{N-1} \hat{f}(l) \chi_l(n) \quad (6)$$

Using these transforms, we construct spectral graph wavelets by applying band-pass filters at multiple scales and *localizing* it with an impulse function and low-pass filter for the scaling function.

Here, λ_l , the spectrum of the Laplacian, serves as an analogue of the 1-D frequency domain, where scales can be easily defined. This directly provides the key component in obtaining a multi-resolutional view of the signal localized at n . By analyzing the entire spectra in different scales, we can find which band particularly characterizes the signal of interest. And for a specific scale s , we can now construct a kernel function g which acts as band-pass filter in the frequency domain. The kernel g enables to focus on signals in certain band at scale s and restrains the influence of all others signal from other scales. When g is transformed back to the original graph domain, we directly obtain a representation of the signal for that scale. Repeating this procedure for multiple scales, the

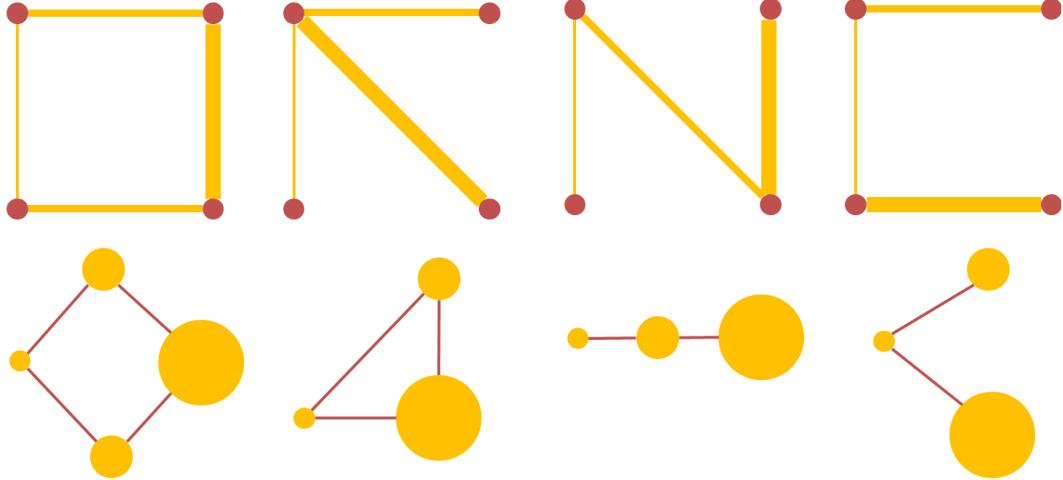


Figure 1: Examples of graphs and corresponding line graphs. Top: Graphs with red vertices and yellow edges. The thickness of edges represents the edge weights, Bottom: Corresponding line graphs with yellow vertices and red edges. The vertex size represents the signal on each vertex

set of coefficients obtained for S scales gives a multi-resolution representation for that particular vertex.

Since the transformed impulse function in the frequency domain is equivalent to a unit function, the wavelet ψ localized at vertex n can now be defined as,

$$\psi_{s,n}(m) = \sum_{l=0}^{N-1} g(s\lambda_l)\chi_l^*(n)\chi_l(m) \quad (7)$$

where m is a vertex index on the graph. With this in hand, the wavelet coefficients of a given function $f(n)$ is given by the inner product of wavelets and the given function,

$$W_f(s, n) = \langle \psi_{s,n} f \rangle = \sum_{l=0}^{N-1} g(s\lambda_l)\hat{f}(l)\chi_l(n) \quad (8)$$

Remark. Wavelet in Euclidean space has rich history in the signal processing field. However, defining wavelet in non-Euclidean space is a very recent work ([8, 3]) and has a lot of potential for many applications.

4 Line Graphs

In the graph theory, one defines the line graph $L(G)$ as a dual form of graph G . The $L(G)$ is formed by the interchange of the roles of V and E in G . Two vertices in $L(G)$ are connected when the corresponding edges in G share a common vertex. In this fashion, the line graph $L(G) = \{V_L, E_L, \omega_L\}$ has a vertex set that corresponds to the edges $\{E, \omega\}$ and a edge set that corresponds to the vertex set V in G . [9]

The transformation of $L(G)$ from a graph G is defined as the following. Let g_{ij} be the elements in the adjacency matrix \mathbf{A}_L of $L(G)$, then

$$g_{ij} = \begin{cases} 1 & \text{if } v \in V, v \curvearrowright e_i, e_j \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where v is a vertex in V and e is an edge in E . This means that when two edges share a common vertex in G , these edges are connected to each other by the common vertex. This is similar to the concept of adjacency matrix of vertices where the elements are nonzero if two vertices share a common edge, and thus also called edge-adjacency matrix. After this transformation, the isolated vertices in G become completely neglected in $L(G)$. If G is a connected graph, then $L(G)$ is unique with single

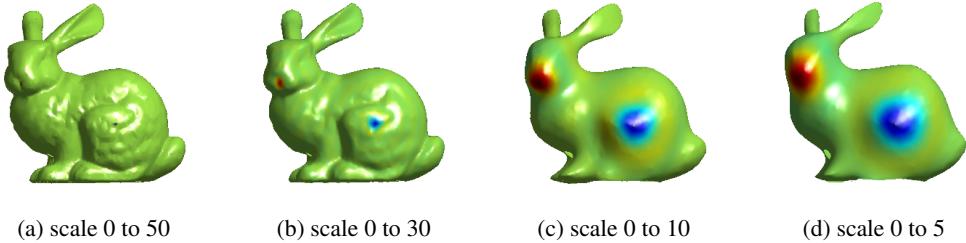


Figure 2: Toy example of surface and signal smoothing on Stanford bunny. Surface vertex coordinates and a random signal (all zeros but two points) were decomposed by 50 scales and scaling function denoted as 0 using wavelet, and was reconstructed by different scale spectra.

exception, which are triangle and 'Y' shaped graphs sharing the same line graph. Moreover, if there is no isolated vertices in G , then G and $L(G)$ have equal number of components. Because of these two properties of line graphs, we can transform the brain network to line graphs for data processing and comeback to the original network structure.

Remark. After constructing a line graph $L(G)$ of a graph G , the edges in G forms a whole new non-Euclidean domain of analysis and the edge weight ω can be defined as a function defined on each vertex in V_L , where the connection between each vertex in E_L is given from V . Four toy examples of this transformation are displayed in Fig.1. The top row shows the original graph G , and the bottom row shows the transformed line graph $L(G)$.

5 Method

Signals in real data f can be modeled as a combination of true signal and some noise, and the true signal tends to change smoothly while noise varies very rapidly in high frequency. It is obvious if the signal is represented in harmonics, i.e. Fourier series, covering a low-pass filter would remove the components of high frequency components. Using wavelet, smoothing can be efficiently performed by removing high frequency components tied to the finer scales. Previous existing methods for signal smoothing on graph structure, such as spherical harmonics, explicitly represent the the signal as a superposition of basis functions defined over regular Euclidean spaces. Although such methods have been shown to be quite robust, but loss of information is inevitable due to the 'ballooning' process, maping the original graph domain to a Euclidean space such as a sphere. Here, defining the wavelets on graphs provides a powerful tool in that it constructs basis function on the graph domain itself, and thus enables one to bypass the whole procedure of the ballooning process for analyzing certain signals defined in the non-Euclidean space.

Smoothing using wavelet comes from the inverse wavelet transformation of the resultant function that provides the smooth estimate of the signal at various scales. Rewriting (3) in terms of the graph Fourier basis,

$$f(m) = \frac{1}{C_g} \sum_l \left(\int_0^\infty \frac{g^2(s\lambda_l)}{s} ds \right) \hat{f}(l) \chi_l(m) \quad (10)$$

which sums over the entire scale s . In contrast, existing methods introduce an additional smoothness parameter (e.g., σ in case of heat kernel). Limiting the scales to the coarse scales will reconstruct the smoothed approximation of the original signal. Superposing finer scales, the complete spectrum contributes to the reconstruction and recovers the complete original signal. In this fashion, choosing optimal scales may filter noisy high frequency variations and provide the true underlying signal. This smoothing method is also well-explained in [10].

An example of smoothing using wavelet are shown in Fig.2 where we display the process of smoothing the surface of the Stanform bunny and a signal (two peaks on the mouth and the leg) defined on the surface. As we exclude finer scale components, both the surface and the signal become smooth.

In order to filter the network structure, it is inevitable to bring the network connectivity information as a signal in another domain. As decribed in section 4, transformation of a graph domain G to a line graph $L(G)$ enables us to count the edge weights as a signal defined in the domain of $L(G)$. In

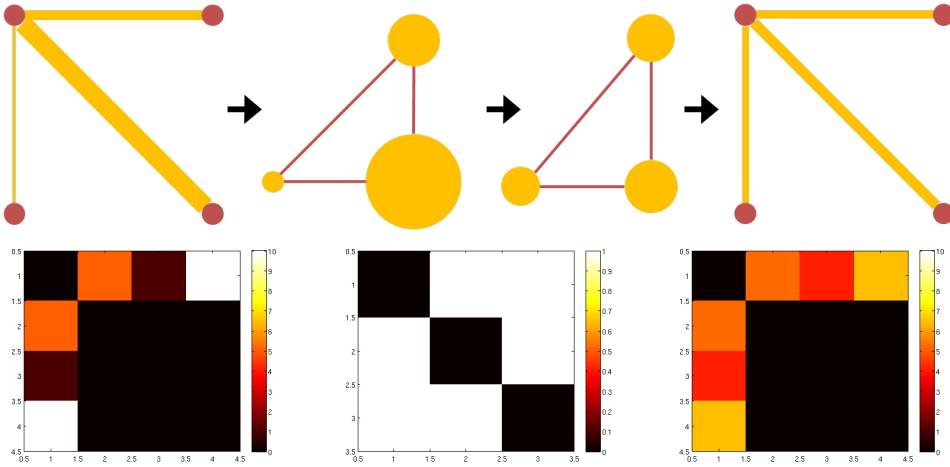


Figure 3: A toy example of graph structure smoothing. Top: Graph smoothing process. Bottom: Adjacency matrix of the original graph, Adjacency matrix of the line graph of the original graph, Adjacency matrix of the smoothed graph.

this manner, we can define the connectivity as a signal on each vertex of $L(G)$, and continue with the smoothing technique using wavelet.

A toy example of the framework for the network smoothing is given in Fig.3. The top row shows the process of the frame work. The original graph G of four vertices and three edges with corresponding edge weights (edge thickness) are transformed to a line graph $L(G)$, and the edge weights become signals (vertex size) defined on $L(G)$. The signals are smoothed along their connection, and when transformed back to the original structure. In the final stage, these signals become the smoothed edge weights without losing the original topology of G . In the bottom row, the corresponding adjacency matrices are displayed. The first matrix shows the connectivity of each vertex in G and its edge weights. The second matrix shows the relation of each edge as $L(G)$, and finally the third matrix shows the smoothed edge weights maintaining the original connection in the first matrix.

6 Application

The dataset consists of bipolar and control group, each contains 25 subjects. The dataset contains brain network information as a adjacency matrix of 87 parcellated brain regions. The edge weights in the adjacency matrix represent the number of fibers connecting one brain region to another. The followings show the result of group analysis, detecting group differences of bipolar and control group using this framework. Note that using the original raw dataset for the t -test, there exists no signal that survives multiple comparison correction.

Group Analysis. Given the dataset, one can simply apply two sample t -test on the distribution of each element (connection) of the matrix across the two groups, however, none of the connection survives multiple comparison correction and thus finds no significant difference. When the given brain networks are smoothed by our method, the variation of the data is greatly relaxed, and this can be seen by comparing the first two figures of Fig.5. It is well known that smoothing data by a linear filter imposes more Gaussianness to the data and thus enhances results from the statistical analysis.

After processing the data using the framework, it finds 7 significant connections over 12 different brain regions, while simple group analysis on the raw dataset finds no difference. In our experiment, we thresholded the connection that are less than 10 fibers, and smoothed using two different scales, including the scaling function which is defined as a heat kernel. Hotelling's T^2 -test was executed on 1911 voxels, excluding those connections on which the edge weights are 0 across all subjects. Then Bonferroni correction at 0.05 was used to reveal the significant voxels. The voxels that survive the Bonferroni threshold are shown in table 1 and those p -values in log scale are represented as the edge thickness in Fig.5. In Fig.5, the detected edges are displayed with the direction of connection strength, the red edge denoting that the edge connection is stronger in the control group, while blue edge means that the edge connection is stronger in the bipolar group.

Significant Group Difference in Brain Network			<i>p</i> -value
Stronger in Bipolar	left-amygala	ctx-lh-temporalpole	2.49e - 5
Stronger in Control	ctx-rh-paracentral	ctx-rh-postcentral	1.22e - 5
	ctx-rh-paracentral	ctx-lh-posterior cingulate	1.37e - 6
	left-caudate	ctx-rh-posterior cingulate	4.09e - 6
	left-caudate	ctx-lh-rostral anterior cingulate	1.08e - 5
	right-putamen	right-accumbens-area	8.3e - 7
	ctx-rh-caudal anterior cingulate	rh-rostral anterior cingulate	5.61e - 6

Table 1: Significant connections that show group differences between bipolar and control group, and their *p*-values.

Sample Variation Test. In order to find proper result from group test such as *t*-test or Hotelling’s T^2 -test, the assumption that the distribution of the data is Gaussian is required. We applied Shapiro-Wilk test on both the raw data and the enhanced data, and compared the result. The Shapiro-Wilk tests the null hypothesis that the data come from a normally distributed population. It turns out that among the 1911 edges that were used for the group analysis, 1544 of them showed improved Gaussiannes in the data. This explains the improvement in the group analysis where we detect significant group differences by enhancing the true signal through our framework, while it is not possible to find the differences in the raw data.

7 Conclusion

In this project, I introduced new framework to analyze network data, where the information lies on the edge of the network. Accomodating the novel idea of line graph, a new domain for the analysis is defined, where the given network edge information is defined as a function of that domain. Signal enhancing process using wavelet empowers the general group analysis using two-sample *t*-test, which is demonstrated in the experiment. The group difference that was hard to notice has become much more sensitive after processing the data through our method, and the result shows prominent differente between the bipolar and control group, whichis consistent with past studies.

References

- [1] F. Abdelnour, B. Schmidt, and T. Huppert. Topographic localization of brain activation in diffuse optical imaging using spherical wavelets. *Physics in medicine and biology*, 54(20):6383, 2009.
- [2] H. Adeli, Z. Zhou, N. Dadmehr, et al. Analysis of eeg records in an epileptic patient using wavelet transform. *Journal of neuroscience methods*, 123(1):69, 2003.
- [3] R. Coifman and M. Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21(1):53–94, 2006.
- [4] M. Cykowski, J. Lancaster, and P. Fox. A method to enhance the sensitivity of dti analyses to group differences: A validation study with comparison to voxelwise analyses. *Psychiatry Research: Neuroimaging*, 193(3):191–198, 2011.
- [5] J. GadElkarim, D. Schonfeld, O. Ajilore, L. Zhan, A. Zhang, J. Feusner, P. Thompson, T. Simon, A. Kumar, and A. Leow. A framework for quantifying node-level community structure group differences in brain connectivity networks. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 196–203, 2012.

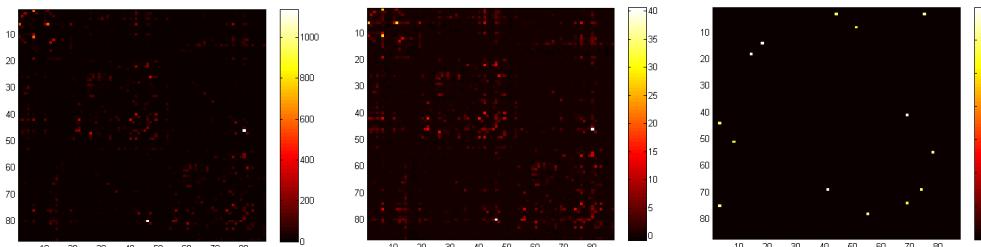


Figure 4: Group mean of raw brain network and smoothed network, and *p*-value in *log* scale from the group analysis on the smoothed network (*t*-test and Bonferroni correction at 0.05

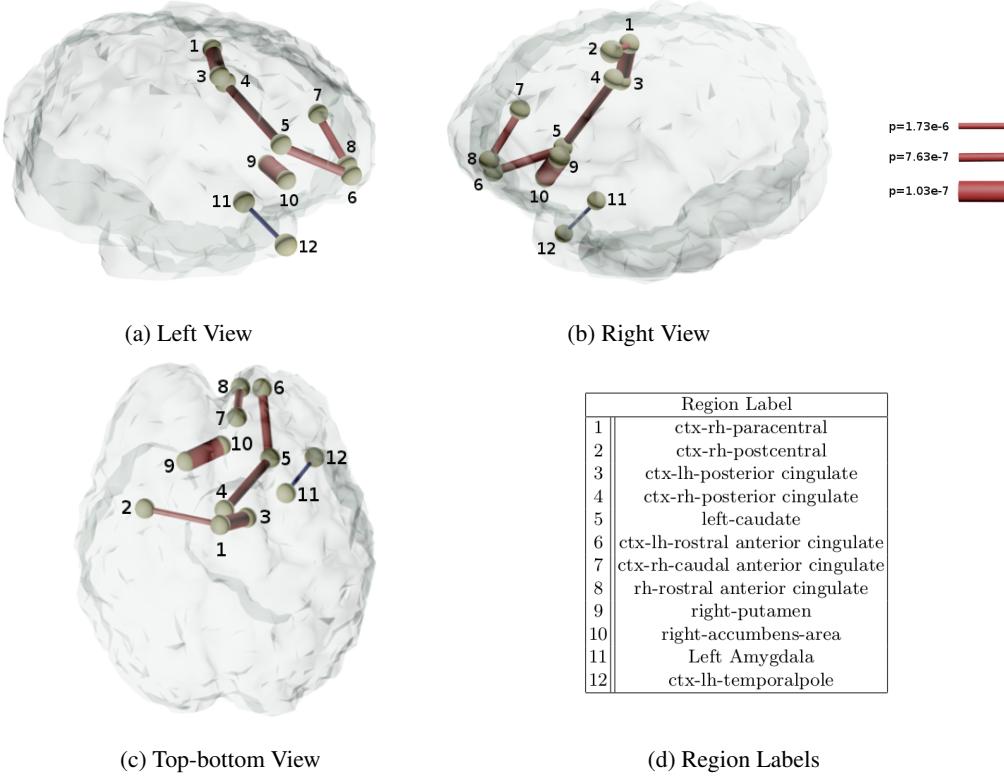


Figure 5: Main result. Anatomical connectivity showing group differences between bipolar and control group after Bonferroni threshold at 0.05. The connection thickness represents the p -value in \log scale, and the color represents sign of strength. Red: Stronger in control group, Blue: Stronger in bipolar group.

- [6] S. Gefen, O. Tretiak, L. Bertrand, G. Rosen, and J. Nissanov. Surface alignment of an elastic body using a multiresolution wavelet representation. *Biomedical Engineering, IEEE Transactions on*, 51(7):1230–1241, 2004.
- [7] C. B. Goodlett, P. T. Fletcher, J. H. Gilmore, and G. Gerig. Group analysis of dti fiber tract statistics with application to neurodevelopment. *NeuroImage*, 45(1, Supplement 1):S133 – S142, 2009. *[ce:title]Mathematics in Brain Imaging[/ce:title]*.
- [8] D. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129 – 150, 2011.
- [9] F. Harary. *Graph Theory*. 1969. Addison-Wesley, Reading, MA.
- [10] W. Kim, D. Pachauri, C. Hatt, M. Chung, S. Johnson, and V. Singh. Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination. In *Advances in Neural Information Processing Systems 25*, pages 1250–1258, 2012.
- [11] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, 11(7):674 –693, 1989.
- [12] D. Nain, M. Styner, M. Niethammer, J. Levitt, M. Shenton, G. Gerig, A. Bobick, and A. Tannenbaum. Statistical shape analysis of brain structures using spherical wavelets. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pages 209–212. IEEE, 2007.
- [13] S. Roosendaal, J. Geurts, H. Vrenken, H. Hulst, K. Cover, J. Castelijns, P. Pouwels, and F. Barkhof. Regional dti differences in multiple sclerosis patients. *NeuroImage*, 44(4):1397 – 1403, 2009.
- [14] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.
- [15] S. Haykin and B. V. Veen. *Signals and Systems, 2nd Edition*. Wiley, 2005.
- [16] C. Von Der Malsburg. The correlation theory of brain function. *Models of Neural networks II, edited by E. Domany, JL van Hemmen, and K. Schulten (Springer, Berlin, 1994) Chapter, 2:95*, 1981.
- [17] P. Yu, P. Grant, Y. Qi, X. Han, F. Ségonne, R. Pienaar, E. Busa, J. Pacheco, N. Makris, R. Buckner, et al. Cortical surface shape analysis based on spherical wavelets. *Medical Imaging, IEEE Transactions on*, 26(4):582–597, 2007.

- [18] W. Zhang, Z. Jin, G. Cui, K. Zhang, L. Zhang, Y. Zeng, F. Luo, A. Chen, and J. Han. Relations between brain network activation and analgesic effect induced by low vs. high frequency electrical acu-point stimulation in different subjects: a functional magnetic resonance imaging study. *Brain research*, 982(2):168–178, 2003.