Multi-resolutional Brain Network Filtering and Analysis via Wavelets on Non-Euclidean Space [*]

Won Hwa Kim[1], Nagesh Adluru[1], Moo K. Chung[1], Sylvia Charchut[4],
Johnson J. GadElkarim[2], Lori Altshuler[3], Teena Moody[3],
Anand Kumar[2], Vikas Singh[1], and Alex D. Leow[2]

[1]University of Wisconsin, Madison, [2]University of Illinois, Chicago,
[3]University of California, Los-Angeles, [4]Southeastern Louisiana University, Hammond

Abstract. Advances in resting state fMRI and diffusion weighted imaging (DWI) have led to much interest in studies that evaluate hypotheses focused on how brain connectivity networks show variations across clinically disparate groups. However, various sources of error (e.g., tractography errors, magnetic field distortion, and motion artifacts) leak into the data, and make downstream statistical analysis problematic. In small sample size studies, such noise have an unfortunate effect that the differential signal may not be identifiable and so the null hypothesis cannot be rejected. Traditionally, smoothing is often used to filter out noise. But the construction of convolving with a Gaussian kernel is not well understood on arbitrarily connected graphs. Furthermore, there are no direct analogues of scale-space theory for graphs — ones which allow to view the signal at multiple resolutions. We provide rigorous frameworks for performing 'multi-resolutional' analysis on brain connectivity graphs. These are based on the recent theory of non-Euclidean wavelets. We provide strong evidence, on brain connectivity data from a network analysis study (structural connectivity differences in adult euthymic bipolar subjects), that the proposed algorithm allows identifying statistically significant network variations, which are clinically meaningful, where classical statistical tests, if applied directly, fail.

1 Introduction

The development of diffusion weighted imaging (DWI) and functional magnetic resonance imaging (fMRI) have laid the groundwork for ambitious initiatives towards a full characterization of the human connectome (the brain's wiring diagram) to better understand the structural and functional aspects of brain connectivity. While such large scale projects will clearly push the frontiers of

---

neuroscience forward, these efforts must necessarily go hand in hand with studies seeking to answer *more focused* questions pertaining to the variations in the connectivity structure in the context of *specific* neurodegenerative diseases, and how its manifestation is modulated by genetic and demographic factors. As such these studies operate in the *small sample size* regime, and the first order requirement on analysis methods appropriate for such applications is to maximize statistical power — in other words, the likelihood of observing a differential signal in the connectivity data given the limited size of the cohort.

Consider a brain connectivity network modeled as an undirected weighted graph denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$. The vertices, $\mathcal{V}$ may denote anatomically meaningful parcellations [6] or regions exhibiting spatially contiguous BOLD activations where as the weighted edges may correspond to temporal correlations or strength of tract connectivity[1]. Now, our interest is to perform statistical analysis on a population of such brain connectivity networks in clinically disparate groups, to understand which connections are severely affected by the disease. The overwhelming majority of current literature suggests applying standard hypothesis testing at the level of individual network edges. This approach generally works well, but when the group-wise differences are weak to begin with, one finds that after correcting for multiple comparisons, a statistically significant signal may be unidentifiable. One may smooth out the noise variance in $\mathcal{E}$ and $\mathcal{V}$, but it is still an issue in brain connectivity analysis. For example, even a small head motion in the scanner can influence the DTI connectivity information. [16]

The standard procedure to improve the signal to noise ratio is to smooth the input signal. In image processing, the measurements are defined on a uniformly sampled lattice (Euclidean space) where the standard notions of a convolution filter apply directly. The few instances in the literature which implement a smoothing process on brain network data essentially average the measurements within a the node or edge ROIs with isotropic Gaussian kernels [18,14], discarding the network structure. In computer vision, anisotropic diffusion kernels [15], heat kernels [20] and pyramids [11] have been used extensively, but mostly in the context of a regular lattice over pixels. Developments of these ideas have led to much work in scale space theory, towards deriving multi-resolution representations of the image by incremental smoothing. The question we investigate is whether such connections can be exploited to analyze brain connectivity network data with enhanced statistical sensitivity. The literature offers few strategies for filtering of signals on the edges of a brain connectivity network.

Key contributions. The most natural mathematical tool which offers multiresolution behavior, i.e., *wavelets*, was until recently, restricted to the Euclidean space. But the objects of interest here are networks with arbitrary topology (non-Euclidean). **a)** We make use of a recent harmonic analysis results to show how non-Euclidean wavelets provide tools for defining multi-scale representations of brain networks. **b)** We demonstrate an application to analyzing *structural connectivity* differences between euthymic bipolar disorder and healthy subjects. Our framework applies multi-resolutional analysis on the information defined on the edges, not on the vertices. The noise in raw connectivity data has the effect

that few edges show up as statistically significant after accounting for multiple comparisons correction. But applying the proposed method, clinically meaningful group differences can be detected at the Bonferroni threshold of $\alpha < 0.01$.

2   Non-Euclidean Wavelets

Wavelet transform is conceptually similar to the well-known Fourier transform, however, it uses a certain shape of oscillating function as a basis with finite duration instead of the sine and cosine basis with infinite duration. The traditional construction of wavelets is defined by a mother wavelet function ψ and a scaling function ϕ, which are band and low-pass filters in the frequency domain.

The wavelet function ψ on x is a function defined by two parameters, the scale parameter s and translation parameters a

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

Change in s varies the *dilation* of the wavelet, and together with a *translation* parameter a, approximates a signal in harmonics using wavelet expansion. The function $\psi_{s,a}(x)$ forms bases 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 $f(x)$,

$$W_f(s,a) = \langle f, \psi \rangle = \frac{1}{a} \int f(x)\psi^*(\frac{x-a}{s})\mathrm{d}x, \tag{2}$$

where $W_f(s,a)$ is the wavelet coefficient at scale s and at location a. The original signal $f(x)$ can be reconstructed from $W_f(s,a)$ and basis function without loss of information; the inverse transformation is

$$f(x) = \frac{1}{C_\psi} \iint W_f(s,a)\psi_{s,a}(x)\mathrm{d}a\,\mathrm{d}s \tag{3}$$

where $C_\psi = \int \frac{|\Psi(j\omega)|^2}{|\omega|}\mathrm{d}\omega$ is known as the *admissibility condition constant*, Ψ is the Fourier transform of the wavelet [9], and ω denotes the frequency domain.

Recent work in harmonic analysis [7] provides wavelet basis on structured data which expresses in a wide spectrum of frequencies. The solution in [7] relies on a graph Fourier transform to derive a spectral graph wavelet transform (SGWT). It is shown that SGWT formalization preserves the localization properties at fine scales as well as other wavelet specific properties, while addressing the bottleneck of defining *scales* on a domain where the space is non-Euclidean.

Let a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \omega\}$ be a undirected graph with a vertex set $\mathcal{V}$ with N vertices, an edge set $\mathcal{E}$ and corresponding edge weight $\omega \geq 0$. The adjacency matrix $\mathbf{A}$ of $\mathcal{G}$ is given as a $N \times N$ matrix whose elements a_{ij} are the edge weight ω_{ij} if i^{th} and j^{th} nodes are connected. The degree matrix $\mathbf{D}$ is computed as a $N \times N$ diagonal matrix whose i^{th} diagonal is $\sum_j \omega_{ij}$. The graph Laplacian from these graph matrices is defined in the usual way as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. Then, the complete orthonormal basis χ_l and eigenvalues λ_l, $l \in \{0, 1, \cdots, N-1\}$

obtained from the graph Laplacian, a self-adjoint operator, forms the basis for the graph Fourier transformation. Note that the λ_l are increasingly ordered, and are irrelevant to the order of vertex index in the graph domain. Using these basis, the forward and inverse graph Fourier transformation are defined using the eigenvalues and eigenvectors of **L** as,

$$\hat{f}(l) = \langle \chi_l, f \rangle = \sum_{n=1}^{N} \chi_l^*(n) f(n) \text{ and } f(n) = \sum_{l=0}^{N-1} \hat{f}(l) \chi_l(n) \tag{4}$$

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 analog 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. Constructing a kernel function g which acts as band-pass filter in the frequency domain, 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 set of coefficients obtained for each $s \in S$ 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,

$$\psi_{s,n}(m) = \sum_{l=0}^{N-1} g(s\lambda_l) \chi_l^*(n) \chi_l(m) \tag{5}$$

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

$$W_f(s,n) = \langle \psi_{s,n} f \rangle = \sum_{l=0}^{N-1} g(s\lambda_l) \hat{f}(l) \chi_l(n) \tag{6}$$

SGWT follows the same procedure of constructing wavelets as in the continuous wavelet transform. In the fine scale limit, SGWT maintains many of the properties of the traditional wavelet transform, including localization.

Remark. Wavelets in Euclidean space have a rich history in Signal processing. However, defining wavelets in non-Euclidean space is a recent development [7,5], and is especially interesting for network analysis in Neuroimaging.

3   Deriving a Multi-Resolution Perspective of a Network

Line Graphs. In graph theory, one defines the line graph $L(\mathcal{G})$ as a dual form of graph $\mathcal{G}$. The $L(\mathcal{G})$ is formed by interchanging the roles of $\mathcal{V}$ and $\mathcal{E}$ in $\mathcal{G}$. Two vertices in $L(\mathcal{G})$ are connected when the corresponding edges in $\mathcal{G}$ share a common vertex. The line graph $L(\mathcal{G}) = \{\mathcal{V}_L, \mathcal{E}_L, \omega_L\}$ has a vertex set for the edges $\{\mathcal{E}, \omega\}$ and a edge set that corresponds to the vertices $\mathcal{V}$ in $\mathcal{G}$ [8].
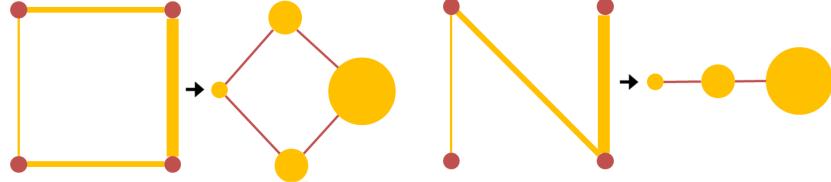
Fig. 1: Examples of graphs and the corresponding line graphs. Original graphs with vertices (red) and edges (yellow) with edge weights (thickness), and corresponding line graphs with vertices (yellow) with function (vertex size) and edges (red).

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

$$g_{ij} = \begin{cases} 1 & \text{if } v \in \mathcal{V}, \ v \frown e_i, e_j \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where v is a vertex in $\mathcal{V}$ and e is an edge in $\mathcal{E}$. This means that when two edges share a common vertex in $\mathcal{G}$, these edges are connected to each other by the common vertex. After this transformation, the isolated vertices in $\mathcal{G}$ are completely neglected in $L(\mathcal{G})$. If there are no isolated vertices in $\mathcal{G}$, then $\mathcal{G}$ and $L(\mathcal{G})$ have equal number of components. After constructing a line graph $L(\mathcal{G})$ of a graph $\mathcal{G}$, the edges in $\mathcal{G}$ form a completely new domain of analysis and the edge weight ω can be defined as a function defined on each vertex in $\mathcal{V}_L$, where the connection between each vertex in $\mathcal{E}_L$ is given from $\mathcal{V}$. Toy examples of this transformation are shown in Fig. 1.

In a measured signal, the true signal tends to change smoothly while noise varies very rapidly in high frequencies. Using wavelets, smoothing can be efficiently performed by removing high frequency components tied to the finer scales, moreover, due to the bandpass property of wavelet, we can get a multi-resolutional view of the given signal. The multi-scale view comes from the inverse wavelet transformation of the resultant function that provides the estimate of the signal at various scales. Rewriting (3) in terms of the graph Fourier basis,

$$f(m) = \frac{1}{C_g} \sum_l \left(\int_0^\infty \frac{g^2(s\lambda_l)}{s} \mathrm{d}s \right) \hat{f}(l)\chi_l(m) \tag{8}$$

which sums over the entire scale s. Limiting the scales to the coarse scales will reconstruct the smoothed approximation of the original signal, and the original signal can be reconstructed by adding finer scales.

In order to filter the network structure, it is necessary to bring the network connectivity information as a signal into another domain. As described above, the transformation of a graph domain $\mathcal{G}$ to a line graph $L(\mathcal{G})$ enables us to view the edge weights as a signal defined in the domain of $L(\mathcal{G})$. We can therefore define the connectivity as a signal on each vertex of $L(\mathcal{G})$, and continue with the smoothing technique using wavelet. An illustrative example of the framework for the network smoothing is given in Fig. 2, where the edge weights are filtered along their connection and not losing the original topology of $\mathcal{G}$. The corresponding adjacency matrices are displayed at the bottom.
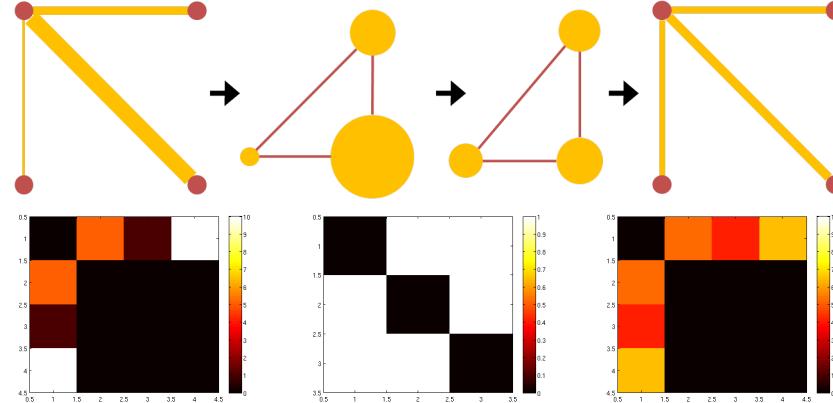
Fig. 2:   A toy example of graph structure filtering. The top panel shows the graph filtering steps: (1) Construction of the line graph, (2) filtering the signal on the line graph vertices, (3) reconstructing the filtered graph. The bottom panel shows the corresponding adjacency matrices.

In addition to the filtering, we define *Wavelet Multi-scale Descriptor* (WMD) using the wavelet wavelet coefficients over the signal defined on each vertex as

$$\text{WMD}_f(n) = \{W_f(s,n)|s \in S\} \tag{9}$$

which characterizes the signal at multi-resolutions on the vertex according to the geometry of the graph[10]. While [19] uses a sphere to obtain the descriptor, which causes data distortion by mapping process, WMD is derived based on the eigenfunction of the original graph itself, and thus avoids 'ballooning'.

4   Connectivity Differences in Bipolar Disorder

Dataset. We scanned 25 healthy subjects (13 male and 12 female; age: 42.2 ± 10.8) and 25 gender and age matched bipolar subjects (14 male, 11 female; age: 41.7 ± 12.6). All bipolar subjects received comprehensive psychiatric evaluations using the structured clinical interview for DSM disorders (SCID) and met the DSM IV criteria for bipolar I disorder (at the time of image acquisition all subjects have been in an euthymic state for at least 30 days). A Siemens 3T Trio scanner was used to acquire the brain MRI data. High resolution T1-weighted images were acquired with MPRAGE sequence (FOV $= 250 \times 250$ mm^2; TR/TE $= 1900/2.26$ ms; flip angle $= 9°$; voxel size $= 1 \times 1 \times 1$ mm^3). Diffusion weighted (DW) images were acquired using SS-SE-EPI sequences (FOV $= 190 \times 190$ mm^2; resolution $2 \times 2 \times 2$ mm^3; TR/TE $= 8400/93$ *ms*; 64 gradient directions, $b = 1000$ s/mm^2 and one minimally DW scan: b=0 image).

Structural brain networks were generated using a pipeline which integrates multiple image processing steps. First, DW images were eddy current corrected using FSL by registering all DW images to their corresponding b=0 images with 12-parameter affine transformations. This was followed by the computation of diffusion tensors and then deterministic tractography using the FACT algorithm [13] built into the DTIStudio program (maximum bending angle 60 degrees; FA cut-off 0.25). T1-weighted images were used to generate label maps using
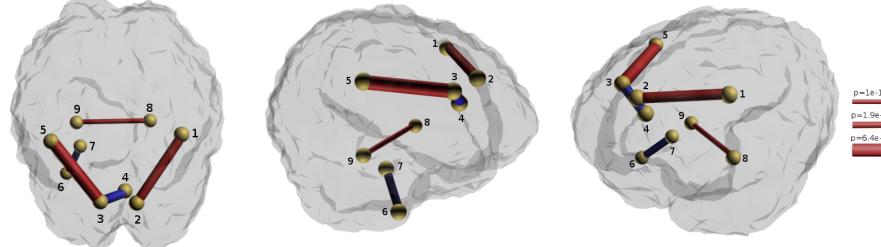
Fig. 3: Anatomical connectivity showing group differences between bipolar and controls after Bonferroni threshold at $\alpha = 10^{-7}$. Connection thickness represents the p-value in negative log scale; color gives sign of strength: red (and blue) are stronger in controls (and bipolar group). Region labels are: 1. ctx-rh-precentral, 2. ctx-rh-superiorfrontal, 3. ctx-lh-superiorfrontal, 4. ctx-rh-caudal anterior cingulate, 5. ctx-lh-precentral, 6. ctx-lh-temporalpole, 7. left amygdala, 8. right hippocampus, 9. left hippocampus.

Freesurfer. The number of tracts connecting 87 cortical/subcortical regions were used in constructing 87×87 connectivity matrix for each subject.

Group Analysis. There are total 3741 edges in the network. In order to detect connectivity differences between the two groups, we performed a Hotellings T^2- test using WMD. WMD was realized by a Mexican-hat wavelet, which was defined at 5 scales over the spectrum of λ. Since typically noise lies in high frequency, we dropped 2 scales that correspond to larger λ, and used the rest for the statistical analysis. When using the raw edge weights, we could not detect any significant difference between the two groups after accounting for Bonferroni correction at $\alpha = 0.05$ significance level. However, after applying WMD on smoothed edge weights with the proposed method, we identified 5 connections over 9 different brain regions as having significant connectivity differences at a very conservative Bonferroni correction level, 10^{-7}.

Interpretation. Results showed that relative to control subjects, bipolar patients on average exhibited weaker strength for the connections within the frontal lobe (bilateral precentral to superior frontal) as well as in fiber tracts linking the bilateral hippocampus. These findings are consistent with past studies where abnormalities in the frontal, limbic, and callosal systems have been reported (for a review, see for example [12]). Additionally, previous fMRI studies of euthymic bipolar patients have also consistently revealed frontal hypoactivation [4,3,17]. In contrast, bipolar subjects exhibit a stronger connection, relative to controls, between the left amygdala and the left temporal pole and between the left superior frontal gyrus and the right caudal anterior cingulate. Although our subjects were in euthymia at the time of the scan, these stronger connections in bipolar may be related to amygdala activation during mania as reported in [2].

5   Conclusion

In this paper, we introduced a novel signal filtering approach for brain network data that takes into account the non-Euclidean nature of the structured data. Using a line graph construction from the original network, we perform band-pass filtering of signals defined on network edges to obtain a multi-resolutional view.

The algorithm siginficantly improves the statistical sensitivity of connectivity differences using Hotelling's T^2-tests and Bonferroni correction. We believe that adapting non-Euclidean wavelets for improving the statistical properties of brain connectivity networks may improve analysis of a much wider variety of studies.

References

1. N. Adluru, M. K. Chung, N. Lange, et al. Applications of epsilon radial networks in neuroimage analyses. In *AIVT LNCS*, pages 236–247, 2012.
2. L. Altshuler, S. Bookheimer, M. A. Proenza, et al. Increased amygdala activation during mania: a functional magnetic resonance imaging study. *American Journal of Psychiatry*, 162(6):1211–1213, 2005.
3. M. A. Cerullo, C. M. Adler, M. P. Delbello, et al. The functional neuroanatomy of bipolar disorder. *International Review of Psychiatry*, 21(4):314–322, 2009.
4. C. Chen, J. Suckling, B. R. Lennox, C. Ooi, et al. A quantitative meta-analysis of fMRI studies in bipolar disorder. *Bipolar disorders*, 13(1):1–15, 2011.
5. R. R. Coifman and M. Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21(1):53–94, 2006.
6. P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. Honey, and et. al. Mapping the structural core of human cerebral cortex. *PLoS Biol*, 6(7):e159, 2008.
7. D. K. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *App. and Comp. Harmonic Analysis*, 30(2):129–150, 2011.
8. F. Harary. *Graph Theory*. Addison-Wesley, Reading, MA, 1969.
9. S. Haykin and B. Van Veen. *Signals and Systems, 2nd Edition*. Wiley, 2005.
10. W. H. Kim, D. Pachauri, C. Hatt, et al. Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination. In *NIPS*, 2012.
11. D. G. Lowe. Object recognition from local scale-invariant features. In *ICCV*, volume 2, pages 1150–1157. Ieee, 1999.
12. K. Mahon, K. E. Burdick, and P. R. Szeszko. A role for white matter abnormalities in the pathophysiology of bipolar disorder. *Neuroscience and Biobehavioral Reviews*, 34(4):533–554, 2010.
13. S. Mori, B. J. Crain, V. P. Chacko, et al. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45:256–269, 1999.
14. D. Pachauri, C. Hinrichs, M. K. Chung, et al. Topology-based kernels with application to inference problems in AD. *TMI*, 30(10):1760–1770, 2011.
15. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *TPAMI*, 12(7):629–639, 1990.
16. J. D. Power, K. A. Barnes, A. Z. Snyder, et al. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154, 2012.
17. A. Van der Schot, R. Kahn, N. Ramsey, et al. Trait and state dependent functional impairments in bipolar disorder. *Psychiatry research*, 184(3):135, 2010.
18. W. Van Hecke, A. Leemans, S. De Backer, et al. Comparing isotropic and anisotropic smoothing for voxel-based DTI analyses: A simulation study. *Human Brain Mapping*, 31:98–114, 2010.
19. P. Yu, P. E. Grant, Y. Qi, et al. Cortical surface shape analysis based on spherical wavelets. *TMI*, 26(4):582–597, 2007.
20. F. Zhang and E. Hancock. Graph spectral image smoothing using the heat kernel. *Pattern Recognition*, 41(11):3328–3342, 2008.