Voxelwise Spectral Diffusional Connectivity and Its Applications to Alzheimer's Disease and Intelligence Prediction*

Junning Li[1,**], Yan Jin[1,2,**], Yonggang Shi[1], Ivo D. Dinov[1], Danny J. Wang[3], Arthur W. Toga[1], and Paul M. Thompson[1,2]

[1] Laboratory of Neuro Imaging
[2] Imaging Genetics Center
[3] Brain Mapping Center UCLA School of Medicine, Los Angeles, CA 90095, USA
{jli,yjin,yshi,idinov,jj.wang,toga,thompson}@loni.ucla.edu

Abstract. Human brain connectivity can be studied using graph theory. Many connectivity studies parcellate the brain into regions and count fibres extracted between them. The resulting network analyses require validation of the tractography, as well as region and parameter selection. Here we investigate whole brain connectivity from a different perspective. We propose a mathematical formulation based on studying the eigenvalues of the Laplacian matrix of the diffusion tensor field at the voxel level. This voxelwise matrix has over a million parameters, but we derive the Kirchhoff complexity and eigen-spectrum through elegant mathematical theorems, without heavy computation. We use these novel measures to accurately estimate the voxelwise connectivity in multiple biomedical applications such as Alzheimer's disease and intelligence prediction.

1 Introduction

The human brain is a complex network of structurally connected regions that interact functionally. Brain connectivity can be studied from different perspectives. Functional MRI, can reveal correlated activity and even causal relationships that underlie the communication of distributed brain systems. On the other hand, diffusion weighted MRI (DWI) measures the local profile of water diffusion in tissues, yielding information on white matter (WM) integrity and connectivity that traditional structural MRI cannot provide. DWI is non-invasive, and is increasingly used to study macro-scale anatomical connections linking brain regions through fibre pathways.

Brain networks are commonly described as a mathematical graph, consisting of a collection of nodes, representing a parcellation of the brain anatomy or regions of interest (ROIs), and a set of edges between pairs of nodes, describing some property of the connection between that pair of regions. The brain exhibits several organization principles, including "small-worldness", characterized by the coexistence of dense local clustering between neighboring nodes and high global

* This work is supported by grants K01EB013633, R01MH094343, P41EB015922, RO1MH080892, R01EB008432, and R01EB007813 from NIH.
** These two authors contributed equally to this work.

K. Mori et al. (Eds.): MICCAI 2013, Part I, LNCS 8149, pp. 655–662, 2013.

efficiency (short average path length) due to few long-range connections [1]. This property results in a sparse connectivity matrix that can be explored using graph theory. A typical way to construct the connectivity matrix is to group adjacent voxels into ROIs (anatomically meaningful grey matter regions) or nodes, and count the fibres passing through each pair of nodes. Then standard measures of connectivity including small worldness, clustering, path length, and efficiency can be computed to reveal how the brain is affected by genetic factors [2] and neurological diseases such as Alzheimer's disease [3].

However, this classical approach has limitations. First, there is lack of validation of WM fibres generated by tractography. Based on different reconstruction or tracking models (tensor vs. orientation distribution function and deterministic vs. probabilistic), different tractography algorithms and variations in their parameters can lead to large differences in the resulting network measures [4]. Automatic cortical grey matter segmentation from an atlas is also susceptible to registration error. Furthermore, the spatial scale of the parcellation of the grey matter into nodes of the connectivity graph may affect connectivity measures by as much as 95% [5]. Finally, parameter thresholding in graph analysis also influences the interpretation of the results [6].

To avoid these problems, we propose a novel mathematical formulation to explore brain connectivity from a different perspective. Instead of investigating linkages among sub-regions of the brain, we use the tensor information from DWI at the voxel level. In this way, we avoid making further assumptions on tractography that diffusion images do not intrinsically provide. Then we show that the diffusion equation can be characterized by the Laplacian matrix of the tensor field. In graph theory, the Laplacian matrix is a matrix representation of a graph. It can be used to calculate the number of spanning trees for a given graph. We, therefore, circumvent the nodal parcellation problem by studying voxelwise linkage. Although others have studied voxel connectivity in its local neighborhood [7], our work focuses more on studying the brain as a whole entity. Finally, we present two important characteristics of a graph, the number of spanning trees (Kirchhoff complexity) and the eigen-spectrum, both of which can be computed without any parameter tuning. As there may be well over a million voxels in a typical image volume, the Kirchhoff complexity and eigen-spectrum can be challenging to compute. We therefore present an algorithm to calculate them efficiently. In the Experimental Section, we illustrate how to evaluate these measures in two biomedical applications (Alzheimer's disease and intelligence prediction).

2 Voxelwise Spectral Diffusional Connectivity

2.1 Diffusion Equation and Tensor

DWI yields information on WM fibres by measuring signals sensitive to the directional diffusion of water molecules. A diffusion process is usually described by the diffusion equation, which is a partial differential equation as follows

$$\frac{\partial f(x,t)}{\partial t} = \langle \nabla, T(x) \nabla f(x,t) \rangle \,, \tag{1}$$

where $f(x,t)$ is the density of the diffusing material at time t and at location x (in the continuous domain), $T(x)$ is the diffusion tensor at location x, and ∇ represents the spatial derivative operator. Here $-T(x)\nabla f(x,t)$ can be understood as the "flux", the amount of diffusing material moving through a unit surface at location x, and over a unit time interval starting at time t.

$T(x)$ fully characterizes the diffusion properties of a field. The diffusion tensor images reconstructed from DWI are voxel estimates of the diffusion field $T(x)$. To make $T(x)$ reflect the spatial density of WM fibres, we modulate diffusion tensors with its fractional anisotropy (FA).

2.2 Laplacian Matrix and Graph

To study $T(x)$ numerically, the spatially and temporally continuous process defined by Eq. (1) should be discretized. As we are interested in $T(x)$ itself, not the diffusion process $f(x,t)$, we discretize it only spatially with the finite difference method. Then the discretized version of Eq. (1) becomes $\frac{\partial f}{\partial t} = -Af$, where A is a square matrix of size n, where n is the number of voxels of interest. A should satisfy the following criteria: (1) It is self-adjoint because the diffusivity between two voxels should be independent of the direction of the flux that crosses them. (2) The sum of each row or each column needs to be zero because the total volume of the diffusion material should be preserved . (3) Its off-diagonal elements are non-positive because the molecules diffuse from high concentration to low concentration. Matrices that satisfy all the three properties are called Laplacian matrices.

Laplacian matrices and graphs have a one-to-one mapping relationship. Given the adjacency matrix of an undirected and weighted graph G whose elements $\{g_{ij}\}$ indicate the edge weight between two adjacent vertices i and j, its Laplacian matrix $A = \{a_{ij}\}$ is defined as

$$a_{ij} = \begin{cases} -g_{ij}, & \text{if } i \neq j; \\ \sum_{k=1}^{n} g_{ik}, & \text{if } i = j. \end{cases} \tag{2}$$

We can see that how the Laplacian matrix is constructed also implies that a graph can be inversely constructed from its Laplacian matrix. It is worth noting that the Laplacian matrix A of a connected graph is positive semi-definite. There is one and only one zero eigenvalue and the rest of the eigenvalues are all positive.

With this one-to-one correspondence relationship, we claim that a diffusion field can be studied via its Laplacian matrix or its corresponding graph. For example, we can study its connectivity complexity, as addressed in Section 2.3, its eigen-spectrum, as addressed in Section2.4, or its vertex centrality in future work.

2.3 Spanning Trees and Kirchhoff Complexity

A spanning tree of a connected graph G is a sub-graph connecting all the vertices in G, which does not contain any circular path. Adding one edge to a spanning tree creates a circle and deleting one edge from a spanning tree partitions the

tree into two disjoint sets. Spanning trees play important roles in graph theory. It is also related to fundamental circles and fundamental cut sets of a graph. One measurement of the complexity of a graph is the number of its spanning trees, which is called the Kirchhoff complexity [8]. The extended Kirchhoff complexity for weighted graphs is defined as

$$K(G) = \sum_{\pi \in T(G)} w(\pi), \qquad w(\pi) = \prod_{i,j \in \pi} g_{ij}, \tag{3}$$

where $T(G)$ is the set of spanning trees of an undirected weighted graph $G = \{g_{ij}\}$, π is a spanning tree, and $w(\pi)$ is the weight associated with π by multiplying all the weights of its edges.

We choose $K(G)$ to indicate the complexity of a connectivity network because it enumerates all the possible ways to connect all the vertices in a graph without circles and it also considers the effectiveness of the connection by weighting with its edge weights.

Although $K(G)$ is defined by enumeration, its calculation does not require enumeration. It can be solved with the Kirchhoff Matrix-Tree theorem [9] as follows.

Kirchhoff's Matrix-Tree Theorem: Given a connected undirected weighted graph G, its Kirchhoff complexity is

$$K(G) = \frac{1}{n}\lambda_1 \lambda_2 \cdots \lambda_{n-1} = \det(A_{-i}), \tag{4}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_{n-1}$ are the non-zero eigenvalues of the Laplacian matrix of G, A_{-i} is the matrix derived by removing the ith row and the ith column from the Laplacian matrix, and n is the number of vertices of G. Interestingly, no matter what value i takes, the result is the same.

Eq. (4) requires the calculation of the matrix determinant, but direct and exact calculation of the determinant of large matrices is not currently feasible, as it may lead to numerical overflow. Fortunately, we can calculate the logarithm of the determinant very efficiently with matrix factorization. Given a symmetric A_{-i}, we first decompose it as $A_{-i} = LDL^{\intercal}$ by the LDL decomposition where L is a square lower uni-triangular matrix and D is a diagonal matrix with rank $n-1$. Now we have two properties: $\det(A_{-i}) = \det(L)\det(D)\det(L^{\intercal}) = \det(D)$ and $\det(D) = \prod_{i=1}^{n-1} d_{ii}$ where d_{ii} is the ith diagonal element of D. Then we can calculate the logarithm of Kirchhoff complexity as

$$\ln K = \ln \det(A_{-i}) = \sum_{i=1}^{n-1} \ln d_{ii}. \tag{5}$$

2.4 Estimation of Eigenvalue Spectrum

The eigenvalues of a Laplacian matrix A not only decide the complexity of a graph (see Eq. (4)) but also convey important information on the temporal responses of the differential equation $\frac{\partial f}{\partial t} = -Af$. However, calculating the eigenvalues of a large sparse matrix demands cumbersome computation and can be

impractical. For example, at the 128 x 128 x 128 image volume size, the Laplacian matrix derived from the diffusion tensor images has approximately 2 x 10^6 rows and columns. For such a large matrix, direct and exact calculation of their eigenvalues is practically impossible. As we are only interested in the distribution of the eigenvalues instead of their exact values, we can estimate the cumulative distribution function (CDF) of the eigenvalues with Sylvester's Law of Inertia.

Sylvester's Law of Inertia: Given a symmetric and real-valued matrix A, its transformation $B = SAS^\intercal$, where S is an invertible square matrix, has the same number of positive/negative eigenvalues as A does.

Let $h(\beta)$ be the number of A's eigenvalues which are equal or smaller than β, that is, $h(\beta) = |\{\lambda_i \leqslant \beta\}|$ where λ_i's are the eigenvalues of A. To calculate $h(\beta)$, we first factorize $A_\beta = A - \beta I$ as $A_\beta = LDL^\intercal$ by the LDL decomposition, where I is the identity matrix, and L and D are denoted as in Section 2.3. Now we have three properties: (1) if λ is an eigenvalue of A, then $\lambda - \beta$ is an eigenvalue of A_β; (2) D has the same number of positive/negative eigenvalues as A_β does; (3) The eigenvalues of D are its diagonal elements. These properties implies that $h(\beta)$ equals the number of the diagonal elements of D which are less than or equal to 0, even though the diagonal elements of D are not necessarily the eigenvalues of A_β. The detailed eigen-spectrum computation algorithm is summarized in **Algorithm 1**.

Algorithm 1. Estimation of Eigen-Spectrum with m Bins

1. Calculate the largest eigenvalue $\lambda_{\max}$ of A with the power iteration method.
2. Set bin positions $\{\beta_i = \frac{i}{m}\lambda_{\max}, i = 1, \cdots m\}$ for estimating $h(\beta)$.
3. For each β_i:
   (a) Decompose $A - \beta_i I$ as $LDL^\intercal$.
   (b) $h(\beta_i) = |\{d_{kk} \leq 0\}|$ where d_{kk}'s are the diagonal elements of D.

3 Experiments

In this section, we show how to apply the theory we derived in Section 2 to two real biomedical problems.

3.1 Graph Construction and Connectivity Computation

The raw diffusion images were corrected for eddy-current induced distortions with FMRIB Software Library (FSL) and then skull-stripped using the FSL tool, BET. The FA modulated tensor field was reconstructed using our own C++ diffusion tool package developed with Segmentation & Registration Toolkit (ITK). Negative eigenvalues of the reconstructed tensors were rectified to their absolute values. Next, the FA image of each subject was linearly aligned to a single-subject International Consortium for Brain Mapping (ICBM) FA atlas from Johns Hopkins University [10] (the atlas was downsampled to the 2 x 2 x 2 mm^3resolution to facilitate the computation). The purpose of registration was to reduce the possible bias in graph construction introduced by individual volumetric differences. Tensors were linearly interpolated and re-oriented with the

preservation-of-principal-direction method [11] when the affine transformation was applied. The discretized Laplacian matrix of the transformed tensor field was constructed. Finally, the logarithm of Kirchhoff complexity was computed as described in Section 2.3 and the logarithmic eigen-spectrum was estimated according to **Algorithm 1** in Section 2.4. The Laplacian matrices are sparse and it took about 5 minutes to perform one LDL routine for a square matrix of size 2×10^6 with a 2.8 GHz Xeon CPU.

3.2 Alzheimer's Disease

Alzheimer's disease (AD) is an irreversible, progressive brain disease that destroys memory and cognition, and is the most common cause of dementia in older people. All of our subjects were recruited as part of phase 2 of the Alzheimer's Disease Neuroimaging Initiative (ADNI2) - an ongoing, longitudinal, multicenter study designed to find biomarkers for the early detection of AD. 155 subjects were categorized into four groups: normal (sex/average age: 22 male (M)/22 female (F)/72.7 years), early mild cognitive impairment (eMCI) (38M/24F/74.0), late mild cognitive impairment (lMCI) (15M/11F/73.0), and AD (15M/8F/75.8). MCI is an intermediate stage between normal aging and AD. 46 DWI volumes were acquired per subject: 5 T2-weighted b_0 image volumes and 41 diffusion-weighted volumes ($b = 1000$ s/mm^2). Each volume dimension was 256 x 256 x 59 and the voxel size was 1.37 x 1.37 x 2.7 mm^3.

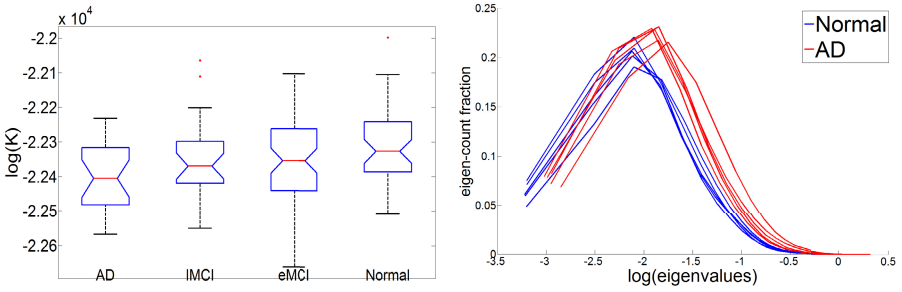

Fig. 1. The figure on the left shows the box plot of the logarithm of Kirchhoff complexity of the four groups and the figure on the right illustrates 5 representative normalized logarithmic eigen-spectra for AD patients (in *red*) and normal controls (in *blue*), respectively

Fig. 1 (left) shows a box plot of the logarithm of Kirchhoff complexity of the four groups. The minimum and the maximum of each group are displayed in *black*, the lower quartile and the upper quartile in *blue*, and the median in *red*. All outliers are marked with "+". The median logarithm of Kirchhoff complexity decreases from the normal group to both MCI groups and to the AD group. On the right, we show ten normalized logarithmic eigen-spectra of the AD group and the normal controls (5 people from each group were randomly selected).

The peaks of the spectra are shifted towards the right in the AD patients, relative to the normal controls. After excluding outliers indicated by the box plot, we also performed a t-test on the logarithm of Kirchhoff complexity between each pair of groups, and the t-statistic between AD and normal group was -3.24 (p=0.0019). The trend of decreasing global structural network connectivity from the normal group to the AD group is consistent with similar findings in functional [12] and anatomical connectivity studies [3].

3.3 Intelligence

80 pediatric subjects were included in this study, from 7 to 17 years old with an average age of 12.2 years. 30-direction DWI data was collected (b = 1000 s/mm^2). The voxel dimension was 128 x 128 x 128, with an isotropic voxel size of 2 mm.

Regression analysis was applied to study the correlation between the subjects' performance intelligence quotient (PIQ) and the logarithm of Kirchhoff complexity, in conjunction with their age variability. The response variable was PIQ, and the regressors were the logarithm of Kirchhoff complexity and age. Scatter plots relating the variables are shown in Fig. 2. Both one-factor and two-factor regressions were performed. Statistics from the regression analysis are listed in Table 1. PIQ and the logarithm of Kirchhoff complexity show statistically significant correlation ($r^2 = 0.066$ or 6.6%) at the 5% significance level. Our result is consistent with Cole *et al.*'s study [13] with functional MRI in which measures of global brain connectivity were found to explain about 5% of the normal variance in intellectual function.

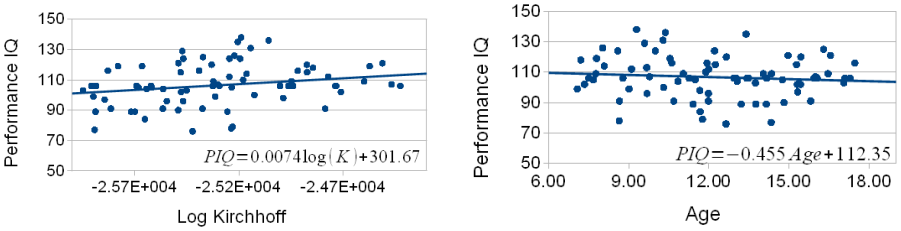

Fig. 2. Scatter plots and regression equations of PIQ against the logarithm of Kirchhoff complexity and age, respectively

Table 1. Regression statistics. The logarithm of Kirchhoff complexity is statistically significant as a predictor of PIQ in both one-factor and two-factor linear regression models.

Regression Model	Variable	Coefficient (95% CI)	r^2	t-statistic	p-value
One-factor	$\ln(K)$	$7.41 \pm 6.43 \times 10^{-3}$	0.066	2.30	**0.0246**
	Age	-0.455±1.08	0.009	-0.840	0.404
Two-factor	$\ln(K)$	$7.23 \pm 6.48 \times 10^{-3}$	0.071	2.22	**0.0292**
	Age	-0.361±1.06		-0.681	0.498

4 Conclusion and Future Work

Here we presented a new method to study overall brain connectivity at the voxel level instead of defining ROI-based nodes and using fibre guidance from tractography. Laplacian matrix of the diffusion tensor field has been proven to have a one-to-one correspondence with its connectivity graph. The voxelwise matrix is high dimensional - making a brute force solution impossible with normal laboratory computing resources. Instead, our measures, the Kirchhoff complexity and eigen-spectrum, can be computed efficiently without an unreasonable computational burden. We illustrate how to apply our measures to biological and medical questions. In our experiments, our estimates have a reasonable interpretation as indices of brain connectivity for disease characterization and intelligence prediction. Future work on voxelwise diffusion connectivity shows promise. For example, we can study the betweenness centrality of a vertex (voxel) and determine the relative importance of a voxel within the network. Or we can perform eigen-embedding to project voxels to higher-dimensional space and invent a new way to define ROIs.

References

1. Bassett, D.S., Bullmore, E.: Small-world Brain Networks. Neuroscientist 12(6), 512–523 (2006)
2. Jahanshad, N., Prasad, G., Toga, A.W., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Wright, M.J., Thompson, P.M.: Genetics of path lengths in brain connectivity networks: HARDI-based maps in 457 adults. In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L. (eds.) MBIA 2012. LNCS, vol. 7509, pp. 29–40. Springer, Heidelberg (2012)
3. Daianu, M., et al.: Analyzing the Structural k-core of Brain Connectivity Networks in Normal Aging and Alzheimer's Disease. In: 15th MICCAI NIBAD Workshop, Nice, France, pp. 52–62 (2012)
4. Bastiani, M., et al.: Human Cortical Connectome Reconstruction from Diffusion Weighted MRI: the Effect of Tractography Algorithm. NeuroImage 62(3), 1732–1749 (2012)
5. Zalesky, A., et al.: Whole-brain Anatomical Networks: Does the Choice of Nodes Matter? NeuroImage 50(3), 970–983 (2010)
6. Dennis, E.L., Jahanshad, N., Toga, A.W., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Wright, M.J., Thompson, P.M.: Test-Retest Reliability of Graph Theory Measures of Structural Brain Connectivity. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 305–312. Springer, Heidelberg (2012)
7. Zalesky, A., Fornito, A.: A DTI-Derived Measure of Cortico-Cortical Connectivity. IEEE Trans. Med. Imaging 28(7), 1023–1036 (2009)
8. Tutte, W.T.: Graph Theory. Cambridge University Press (2001)
9. Chaiken, S.: A Combinatorial Proof of the All Minors Matrix Tree Theorem. SIAM. J. 3(3), 319–329 (1982)
10. Oishi, K., et al.: Atlas-based Whole Brain White Matter Analysis using Large Deformation Diffeomorphic Metric Mapping. NeuroImage 46(2), 486–499 (2009)
11. Alexander, D.C., et al.: Spatial Transformations of Diffusion Tensor Magnetic Resonance Images. IEEE Trans. Med. Imaging 20(11), 1131–1139 (2001)
12. Supekar, K., et al.: Network Analysis of Intrinsic Functional Brain Connectivity in Alzheimer's Disease. PLoS Comput. Biol. 4(6), e1000100 (2008)
13. Cole, M.C., et al.: Global Connectivity of Prefrontal Cortex Predicts Cognitive Control and Intelligence. J. Neurosci. 32(26), 8988–8999 (2012)