ON THE NON-UNIFORM COMPLEXITY OF BRAIN CONNECTIVITY

By

Gloria Haro

Christophe Lenglet

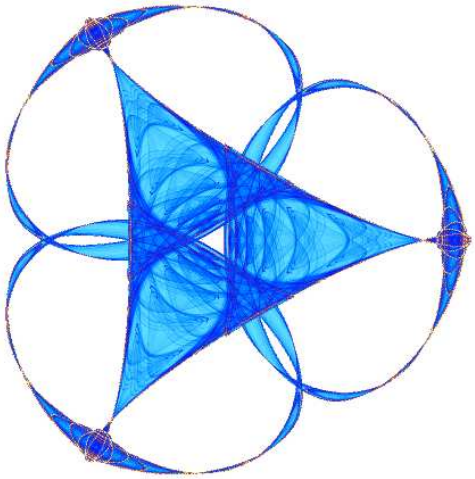
Guillermo Sapiro

and

Paul Thompson

IMA Preprint Series # 2183

(December 2007)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

ON THE NON-UNIFORM COMPLEXITY OF BRAIN CONNECTIVITY

Gloria Haro, Christophe Lenglet, Guillermo Sapiro, and Paul Thompson

UPC, Barcelona; Siemens Corporate Research; University of Minnesota; UCLA

ABSTRACT

A stratification and manifold learning approach for analyzing High Angular Resolution Diffusion Imaging (HARDI) data is introduced in this paper. HARDI data provides high-dimensional signals measuring the complex microstructure of biological tissues, such as the cerebral white matter. We show that these high-dimensional spaces may be understood as unions of manifolds of varying dimensions/complexity and densities. With such analysis, we use clustering to characterize the structural complexity of the white matter. We briefly present the underlying framework and numerical experiments illustrating this original and promising approach.
Key words: Stratification and manifold learning, DTI, HARDI, complexity, white matter connectivity.

1. INTRODUCTION

Diffusion MRI is a powerful extension of MRI that maps how local diffusion affects the MR signal, in multiple sampling directions, providing exquisite insight into local white matter fiber orientation. Water diffusion in the brain occurs preferentially along fiber bundles and is hindered in orthogonal directions, reflecting brain architecture at a microscopic scale. In the Diffusion Tensor (DT) model [2], a tensor describes local 3D diffusion as the 3×3 covariance matrix of a Gaussian distribution, modeling the averaged diffusion properties of water molecules (in a typical 1-3mm sized voxel). High Angular Resolution Imaging (HARDI) overcomes limitations of DTI for characterizing complex tissue geometries such as fiber crossing, measuring diffusion along 30-100 or more directions uniformly distributed on the sphere. From this high-dimensional signal $S_j(x), j = 1, ..., n$, where $x \in \Omega \subset \mathbb{R}^3$ is a voxel of interest and Ω the acquisition grid, spherical functions such as the ADC profile or the Orientational Distribution Function (ODF) may be approximated using a modified spherical harmonic (SH) basis [3]. The ODF provides a non-parametric model of fiber distribution, and is the radial projection of the underlying probability density function for molecular motion.

Analyzing the structure of such complex datasets will lead to a better understanding of brain tissue microstructure and connectivity. Methods such as [4, 8, 9] have been proposed to characterize the anisotropy of tissues and differentiate between isotropic, mono- and multi-fiber configurations, from the SH expansion or full profile of the ADC and from ODFs [12]. These techniques consider each voxel independently and do not try to explain the dataset's global structure.

Concepts from Riemannian geometry (e.g., [1, 6] and references therein) and manifold learning, [13], have been used to characterize the distribution of tensors, perform statistical analysis and segment DTI, e.g., using diffusion maps to cluster ODF fields [15]. These approaches consider the elements of interest (tensors, ODFs) on a single manifold (e.g., a submanifold of $\mathbb{R}^6$ in the tensor case). However, diffusion MRI data does not belong to a single manifold but to a *stratification*, i.e., the union of manifolds with different dimensions (complexities) and densities. Regions with or without fiber crossings clearly belong to manifolds with different dimensionality/complexity (requiring a different number of parameters). The single dimension (complexity) assumption is accurate only for small regions. The effort should then switch toward the study of *stratifications* [5], of which manifolds are a particular case. Studying the different manifolds in the data may also indicate the existence of different complexities in the data. Here we use stratification to *quantify* the local complexity of DTI and HARDI datasets and relate these findings to neuro-anatomical knowledge.

As we show experimentally, we can cluster diffusion MRI datasets by considering them as point clouds in $\mathbb{R}^m$ ($m \geq 6$ depends on the order of the SH approximation of ODFs), without any spatial knowledge. We show that the estimated complexity correlates with the expected fiber geometry in well-known regions of interest.

2. STRATIFICATION LEARNING

A framework for the regularized and robust estimation of non-uniform dimensionality and density in high-dimensional noisy data, i.e., stratifications, was introduced in [5]. High-dimensional sample points are modeled as a process of Translated Poisson mixtures, with regularizing restrictions, and a noise model is incorporated. Levina and Bickel [7] proposed a geometric and probabilistic method to estimate the local dimension and density of point cloud data. In [5] we showed how noise is naturally incorporated in this model, to obtain a

Work partially supported by NSF, NIH, DARPA, NGA, ONR, ARO and Juan de la Cierva Program.

dimension estimator which is robust to noise (noise brings the data outside of the manifold into the ambient space, thereby misleading the dimensionality computation). We tackle the more general case where the point cloud data is a sampling of two or more manifolds of different dimensions and densities (a stratification); we cluster the noisy data points according to these parameters. We do not assume subspaces are linear (e.g., [14]), and we simultaneously estimate the soft clustering and the intrinsic dimension and density of the clusters while being robust to noise and outliers.

If we sample an m-dimensional manifold with T points, the proportion of points falling into a ball around x_t is $\frac{k}{T} \approx \rho(x_t)V(m)R_k(x_t)^m$ [7]. The given point cloud, embedded in high dimensions D, is $X = \{x_t \in \mathbb{R}^D; t = 1, \ldots, T\}$, k is the number of points inside the ball, $\rho(x_t)$ is the local sampling density at point x_t, $V(m)$ is the volume of the unit sphere in $\mathbb{R}^m$, and $R_k(x_t)$ is the Euclidean distance from x_t to its k-th nearest neighbor (kNN). The inhomogeneous process $N(R, x_t)$, which counts the number of points falling into a small D-dimensional sphere $B(R, x_t)$ of radius R centered at x_t, is a binomial process. Under certain assumptions it can be approximated by a Poisson process and the rate λ of the counting process $N(R, x_t)$ can be expressed as $\lambda(R, x_t) = \rho(x_t)V(m)mR^{m-1}$. The local intrinsic dimension estimator at each point x_t is obtained from the Maximum Likelihood estimator based on a Poisson distribution with this rate.

Usually, noise contaminates the point samples, so the observed point process is not a simple sampling of a low-dimensional manifold but a perturbation of this sample process. We model this with a Translated Poisson Process [10], which translates an underlying (unobservable) point process into an output (observable) point process. The input and output spaces of the points are not necessarily of the same dimension. An input point at location x in the input space X is randomly translated to a location z in the output space Z, according to a conditional probability density $f(z|x)$, called the *transition density*. We consider the particular case where each point is translated independently of the others and no deletions or insertions occur in the translation process. Then, any translated Poisson process with an integrable intensity function $\{\lambda(x) : x \in X\}$ is also a Poisson process with intensity $\mu(z) = \int_X f(z|x)\lambda(x)dx$ [10].

Since the intensity of the Poisson process in our model is parametrized by the Euclidean distances of the points (and not by the points themselves), we consider a random translation in the distances. This means that we do not observe the original distances but noisy distances. Let $f(s|r)$ be the transition density which defines the random process which translates a distance r in the input space to a distance s in the observable space. If $\lambda(r, x_t)$ is the local rate of the Poisson process which defines the counting process in the input space, then $\mu(s)$, the intensity of the Poisson process in the output space, is given by $\mu(s, x_t) = \int_0^{R'} f(s|r)\rho(x_t)V(m)mr^{m-1}dr$. We consider $R' > R$ since points originally at distance greater than R from x_t can be placed within a distance less than R after the translation process.

Maximizing the likelihood of the new Translated Poisson process, we obtain the following expression for the local dimension $m(x_t)$ at point x_t when we use the k nearest neighbors (k-NN) instead of the points closer than R,

$$m(x_t) = \left[\frac{1}{k-1} \sum_{i=1}^{k-1} \frac{\int_0^{R'} f(R_i(x_t)|r) r^{m-1} \log \frac{R_k(x_t)}{r} dr}{\int_0^{R'} f(R_i(x_t)|r) r^{m-1} dr} \right]^{-1}, \quad (1)$$

where, by an abuse of notation, we identify $m = m(x_t)$ in the right hand side. This expression reduces to the Levina and Bickel estimator [7] when $f(s|r) = \delta(s - r)$, i.e., there is no translation of the original points (the noise-free case). Equation (1) is a nonlinear recursive expression in m which is difficult to solve. We approximate it by an easier to compute closed expression, with explicit bounds on the approximation,

$$m(x_t) \approx \left[\frac{1}{k-1} \sum_{i=1}^{k-1} \frac{\int_0^{R'} f(R_i|r) \log \frac{R_k}{r} dr}{\int_0^{R'} f(R_i|r) dr} \right]^{-1}, \quad (2)$$

see [5] for details.

These estimators are local as they come from a Maximum Likelihood (ML) at each point x_t based on a Translated Poisson distribution modeling the counting process in the local ball $B(R, x_t)$. We compute an ML on the whole point cloud data at the same time (not just for each point independently), based on a Translated Poisson Mixture Model, which accommodates noise and different classes (each with its own dimension and sampling density). This technique gives a soft clustering according to dimensionality and density, and estimates both quantities for each class. This Translated Poisson Mixture Model (TPMM) is solved with an EM algorithm, which leads to explicit estimations of each cluster dimensionality and density, as well as the probability of each point to belong to each cluster, see [5] for details on the theory and the very efficient computation.

We tested the framework for clustering real data in computer vision applications (scanned digits, faces under varying pose and illumination, different activities and motion in video), obtaining state-of-the-art results for the soft clustering on non-linear stratifications. Here we extend this framework to the stratification defined by HARDI, providing insight into the varying complexity of brain connections.

3. BRAIN MICROSTRUCTURE COMPLEXITY

Diffusion-weighted images $S_j, j = 1...n$, were acquired on a 4T Bruker/Siemens MRI scanner using an optimized diffusion tensor sequence.[1] 30 images were acquired, 3 with no

[1]Imaging parameters were: 21 axial slices (5 mm thick), FOV=23 cm, TR/TE=6090/91.7 ms, 0.5 mm gap, with a 128×100 acquisition matrix (1.8 mm in-plane resolution).

diffusion sensitization S_0 and n=27 diffusion-weighted images at b=1100 s/mm^2. Gradient directions were evenly distributed on the hemisphere.

Diffusion tensors were computed with the standard least-squares procedure using the linearized Stejskal-Tanner equation for free anisotropic diffusion [11]. The signal attenuation $S(\mathbf{q}_j, \tau)$ obtained with the classical Pulsed Gradient Spin Echo (PGSE) sequence, for a given diffusion gradient wavevector $\mathbf{q}_j$ and diffusion time τ, is related to the displacement probability function of water molecules $P(\mathbf{r}, \tau)$ by a Fourier transform, $S(\mathbf{q}_j, \tau) = S_0 \int_{\mathbb{R}^3} P(\mathbf{r}, \tau) e^{-2\pi i \mathbf{q}_j^T \mathbf{r}} d\mathbf{r}$. ODFs, defined as the radial projection of the PDF $P(\mathbf{r}, \tau)$, only conserve the angular information of P, which is sufficient to recover the underlying orientation distribution of fibers. ODFs were approximated by a linear transform of the signals SH coefficients $S(\mathbf{q}, \tau)$ [3]. We examined the complexity of the signal attenuation, diffusion tensors, and SH series coefficients of the ODFs. They respectively correspond to point clouds in $\mathbb{R}^{30}$, $\mathbb{R}^6$ and $\mathbb{R}^{(l+1)(l+2)/2}$, for SH series of order l. The k-NN estimation of the local dimension only weakly exploits the spatial information provided by the regular sampling of points on the acquisition grid $\Omega \subset \mathbb{R}^3$, leaving room for future improvements.

We first studied how the input model influences the consistency of the dimension/density estimation and clustering. In neighborhoods of size $k = 60$, our algorithm analyzed the raw HARDI signal (points in $\mathbb{R}^{30}$), the 4^{th} and 6^{th} order ODFs (points respectively in $\mathbb{R}^{15}$ and $\mathbb{R}^{28}$), and their sharpened versions (Fig. 1). ODF sharpening, [3], enhances the angular contrast of the spherical functions to better differentiate fiber compartments and potentially improve tractography. Clusterings from 4^{th} and 6^{th} order ODFs are almost identical, as 30 gradients may be insufficient to fit a detailed 6^{th} order model.

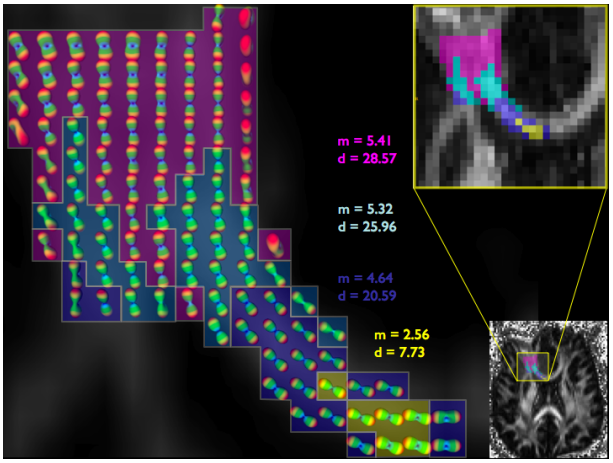


Fig. 2. *Increasing complexity in the forceps minor.*

However, clusterings obtained from the ODFs are clearly better than those from the raw HARDI data; we can readily distinguish (Fig. 1) the gray matter in green, complex white matter in purple (e.g., forceps minor/major, anterior/posterior corona radiata or superior longitudinal fasciculus), anisotropic white matter in light blue (e.g., genu/splenium of the corpus callosum or internal capsule), and highly anisotropic white matter in blue (e.g., genu of the corpus callosum, cortico-spinal tract). However, the cluster dimensions/densities do not match the expected decrease in complexity when going from complex to very anisotropic white matter (Table 1), perhaps because the raw HARDI signal is noisy. 4^{th} and 6^{th} order ODFs regularize the spherical distribution by only considering low-frequency spherical harmonics and impose some smoothness on the fitted ODFs. This translates into improved clustering results and estimated dimensions/densities nicely matching the white matter complexity. The complex white matter is perfectly labeled (purple; Fig. 1), whereas some large areas were missing (and labeled as gray matter) when working on the raw HARDI signal. Highly anisotropic areas (blue) such as the genu/splenium of the corpus callosum and cortico-spinal tract are more consistently labeled. Sharpening the 4^{th} order ODFs had little effect, but decreased clustering accuracy for the 6^{th} order ODFs, perhaps by enhancing high-frequency noise in the higher-order model.

Table 1. *Influence of the input model on complexity.*

Color	*Red*	*Green*	*Blue*	*Yellow*	*L. blue*	*Purple*
HARDI						
Dim.	1.55	4.88	5.92	4.32	5.59	5.67
Dens.	9.27	16.01	10.69	2.42	13.18	15.85
Prob.	0.65	0.18	0.005	0.002	0.026	0.088
ODF 4						
Dim.	1.33	4.53	4.64	2.56	5.32	5.41
Dens.	12.53	26.70	20.59	7.73	25.96	28.57
Prob.	0.70	0.16	0.014	0.002	0.038	0.092
ODF 6						
Dim.	1.34	4.52	4.64	2.57	5.33	5.40
Dens.	12.54	26.64	20.57	7.74	25.94	28.49
Prob.	0.70	0.16	0.014	0.002	0.037	0.092

We compared our estimates to the known complexity of white matter configurations, in the genu of corpus callosum and forceps minor. Callosal fibers are tightly packed at the interhemispheric plane, but diverge and mingle with other fiber bundles as they progress toward the frontal lobes. Our method identifies and quantifies this increase in complexity. The dimension and density of the four submanifolds increase (Fig. 2) as fibers leave the very anisotropic genu region. This clearly demonstrates our methods value for studying white matter microstructure.

Finally, we applied our algorithm to the 6-dimensional diffusion tensor dataset, clearly differentiating the gray/white matter and CSF. However, the difference in complexity between gray and white matter was low and the CSF was clearly isolated, although it was not when working with ODFs. In gray and white matter, diffusion tensors have different anisotropy but similar mean diffusivity (Fig. 3), so the model complexity stays roughly constant. Conversely,

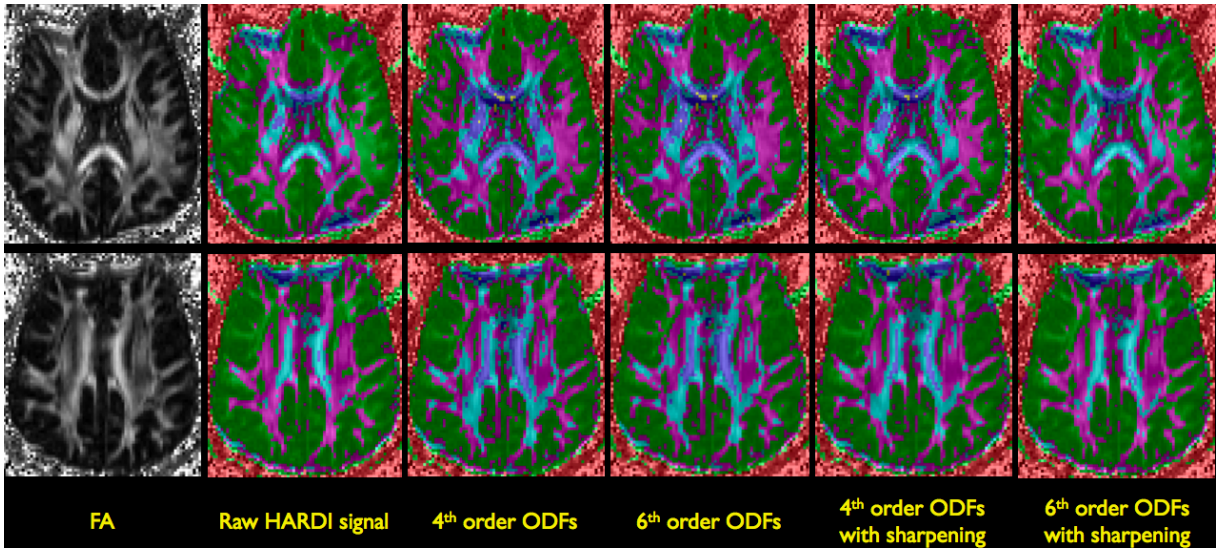


Fig. 1. *Influence of the input model on the labeling of 2 axial slices.*

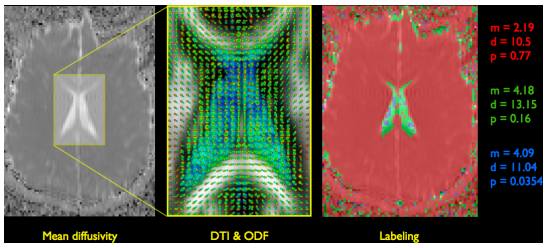


Fig. 3. *Correct labeling of the ventricles from DTI.*

HARDI differentiated gray matter from complex, anisotropic and even very anisotropic white matter, but could not clearly label the CSF where ODFs are intrinsically 2D, they have great angular resolution but lack the amplitude information of DTI (see center panel, Fig. 3).

4. CONCLUDING REMARKS

We presented a stratification learning method to study the non-uniform complexity of HARDI datasets. We labeled known neuro-anatomical areas by examining the complexity of the point clouds obtained from a set of Orientational Distribution Functions. Considering such high-dimensional data as belonging to a union of manifolds is a natural and powerful way to understand cerebral white matter connectivity.

5. REFERENCES

[1] S.P. Awate, H. Zhang, and J.C. Gee. A fuzzy, nonparametric segmentation framework for DTI and MRI analysis: With applications to DTI tract extraction. *IEEE Transactions on Medical Imaging*, 26(11):1525–1536, 2007.

[2] P.J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance*, B(103):247–254, 1994.

[3] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast and robust analytical Q-Ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.

[4] L.R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47:1083–1099, 2002.

[5] G. Haro, G. Randall, and G. Sapiro. Translated poisson mixture model for stratification learning. Technical Report 2174, IMA Report, University of Minnesota, http://www.ima.umn.edu/preprints/jul2007/jul2007.html, 2007.

[6] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision*, August 2006.

[7] E. Levina and P.J. Bickel. Maximum likelihood estimation of intrinsic dimension. In *NIPS 17, Vancouver*, 2005.

[8] E. Ozarslan and T. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution imaging. *Magnetic Resonance in Medicine*, 50:955–965, 2003.

[9] M. Rao, Y. Chen, B.C. Vemuri, and F. Wang. Cumulative residual entropy: A new measure of information. *IEEE Transactions on Medical Imaging*, 50:1220–1228, 2004.

[10] D. L. Snyder and M. I. Miller. *Random Point Processes in Time and Space*. Springer-Verlag, 1991.

[11] E.O. Stejskal and J.E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288–292, 1965.

[12] D. Tuch. Q-Ball imaging. *Magnetic Resonance in Medicine*, 52:1358–1372, 2004.

[13] R. Verma, P. Khurd, and C. Davatzikos. On analyzing diffusion tensor images by identifying manifold structure using isomaps. *IEEE Transactions on Medical Imaging*, 26(6):772–778, 2007.

[14] R. Vidal, Y. Ma, and J. Piazzi. A new GPCA algorithm for clustering subspaces by fitting, differentiating and dividing polynomials. In *Proceedings of CVPR*, pages 769–775, 2004.

[15] D. Wassermann, M. Descoteaux, and R. Deriche. Diffusion maps clustering for magnetic resonance Q-Ball imaging. *International Journal of Biomedical Imaging*, 2007.