Simultaneous Multi-Scale Diffusion Estimation and Tractography Guided by Entropy Spectrum Pathways

Vitaly L. Galinsky* and Lawrence R. Frank

***Abstract*—We have developed a method for the simultaneous estimation of local diffusion and the global fiber tracts based upon the information entropy flow that computes the maximum entropy trajectories between locations and depends upon the global structure of the multi-dimensional and multi-modal diffusion field. Computation of the entropy spectrum pathways requires only solving a simple eigenvector problem for the probability distribution for which efficient numerical routines exist, and a straight forward integration of the probability conservation through ray tracing of the convective modes guided by a global structure of the entropy spectrum coupled with a small scale local diffusion. The intervoxel diffusion is sampled by multi b-shell multi q-angle diffusion weighted imaging data expanded in spherical waves. This novel approach to fiber tracking incorporates global information about multiple fiber crossings in every individual voxel and ranks it in the most scientifically rigorous way. This method has potential significance for a wide range of applications, including studies of brain connectivity.**

***Index Terms*—Brain connectivity, diffusion weighted imaging (DWI), fiber tractography, magnetic resonance imaging (MRI).**

I. Introduction

A PROBLEM of significant interest in basic neuroscience research and in a wide range of clinical applications is the reconstruction of tissue fiber pathways from volumetric diffusion weighted magnetic resonance imaging (DW-MRI) data. This is an inherently ill-posed problem because the local (voxel) diffusion measurements are noisy and made on a scale significantly greater than the underlying fibers and thus there are a multitude of *possible* neural pathways between any two given points in the imaging volume that might be consistent with the experimental data. The question then is to find the paths that are most *probable*. Current fiber tractography methods generally fall into two categories: 1) deterministic methods, typically based on some form of streamline construction (e.g., [1]–[3]) or 2) probabilistic methods, also generally based on streamline construction, but with the most likely principal diffusion direction determined from a posterior distribution of principal diffusion directions (e.g., [4]–[7]). These algorithms are "local" in the sense that the computations are done at each voxel and some small neighborhood around it and thus are not informed by the final path that is created, and thus are not capable of assessing the probability of the final path amongst all possible paths. In most cases, these algorithms are inherently based upon some underlying relation to a random walk which guides the evolution of the trajectories.

Recently, interest has grown in more "global" methods that aim to take into account the probabilities of the final paths by incorporating the path probabilities into the estimation process. These methods typically are based upon parameterizations of the diffusion field, or the anatomical connections they imply, that extend spatially beyond the voxel dimensions and subsequently take the form of either improving the local computations by the incorporation of more spatially extended path lengths (e.g., [8], [9]) or on the extremization of a cost function over a multitude of possible paths [10]–[13]). These global methods usually (with some exceptions [11]) do not take the random walk viewpoint but rather view the entire system as possessing some underlying structure, characterized by local interactions or potentials, that can be elucidated by optimizing some cost function (e.g., energy) over multiple configurations of that system.

The original diffusion tensor imaging (DTI) model assumes that the measurements in each voxel provide an estimate of a single real, 3×3 symmetric diffusion tensor D from whose eigenstructure can be derived both a meaningful measure of the anisotropy (here characterized by the fractional anisotropy FA [14]) and a principal eigenvector that can be used as a proxy for the fiber orientation [14]. Then DTI is the simplest underlying model for diffusion data, is predicated on a single fiber model for the voxel content, and is equivalent to a Gaussian model for diffusion (e.g., [14]). (To be more accurate, DTI can be viewed as the next simplest mathematical framework, while a scalar framework is the simplest that can be used for modeling diffusion data. Also, there may be significant deviations from Gaussian diffusion both on microscopic and on meso-scales. Thus, effectively even DTI may have a deviations from Gaussian due to i.e., cellular boundaries with less than 100% permeability). However, the DTI model is not sufficient to capture more realistic possibilities of complex fiber crossings needed for clinical applications [15]. To estimate local diffusion

Manuscript received October 03, 2014; revised December 04, 2014; accepted December 07, 2014. Date of publication December 18, 2014; date of current version December 18, 2014. This work was supported in part by the National Science Foundation (NSF) under Grant DBI-1143389, Grant DBI-1147260, EF-0850369, and Grant PHY-1201238, and in part by the National Institutes of Health (NIH) under Grant R01 MH096100. Data used for preparation of Fig. 11 were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: D. Van Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research ; and by the McDonnell Center for Systems Neuroscience at Washington University . *Asterisk indicates corresponding author.*

*V. L. Galinsky is with the Center for Scientific Computation in Imaging, and Electrical and Computer Engineering Department, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: vit@ucsd.edu).

L. R. Frank is with the Center for Scientific Computation in Imaging, Center for Functional MRI, University of California at San Diego, and VA San Diego Healthcare System, La Jolla, CA 92093 USA (e-mail: lfrank@ucsd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2014.2380812

directions in each voxel (streamline directions) several high angular resolution diffusion imaging (HARDI) [16] methods are typically used. These methods represent an extension of the original DTI acquisition framework [17] to higher angular resolutions appropriate not only for detection of main fiber orientation, but also for attempting to resolve more complex intravoxel fiber architecture such as multiple crossing fibers [18]–[22].

In recent years, there has been significant interest in developing DW-MRI methods capable not only of estimating angular fiber distributions from multidirectional diffusion imaging (multiple q-angles) [16], [19]–[23]), but also find spatial scales with multiple diffusion weightings (multiple b-shells) [24]–[28]. While it has long been recognized that the most general nonparametric (model-free) approach is to measure the displacement probability density function or diffusion propagator directly [29], [30], the natural extension of this to imaging wherein 3-D Cartesian sampling of q-space is used to obtain the 3-D displacement probability density function (dPDF) at each voxel [24], is prohibitively expensive from the standpoint of data acquisition. This recognition has recently spawned more practical methods for obtaining an estimate of the dPDF, often called the ensemble average propagator (EAP), from more practical multi-shell, multi-directional acquisitions [25]–[28].

Nevertheless, despite these advances, a critical simplification that is made in all current methods used to estimate either the intravoxel diffusion characteristics (via the EAP, for example) or to estimate the underlying global structure (tractography) is the assumption that these two estimation procedures are *independent*. Thus one first estimates the intravoxel diffusion, then applies a tractography algorithms. For example multiple b-shell effects, used in obtaining the EAP, are used only to infer directional multiple fiber information for input into streamline tractography algorithms (see e.g., [31]). However, this distinction between local and global estimation is artificial and limiting, since both the local (voxel EAP) information and the global structure (tracts) are from the same tissue, just seen at different scales. In practical applications of human DW-MRI data this artificial division of information on local and global gives rise to problems at the data interpretation stage: e.g., spurious white matter loss due to local underestimation of fiber anisotropy, incorrect tracking of fiber orientation due to apparent overlapping of angular distributions, etc.

In this paper we revisit the problem of local diffusion estimation and fiber tractography with the specific goal to include multiple spatial and temporal scales that can be deduced from multiple b-shell DW-MRI measurements in addition to just angular (multi-)fiber orientation [32], [33]. In many practical applications, either one or two spatial locations (or regions) are known *a priori*. In neuroscience applications, for example, two regions may be functionally connected (as measured, perhaps, by FMRI) and the diffusion weighted MRI data is being used to assess the degree (if any) of the structural connectivity between two functionally connected regions. We therefore reconsider two common formulations of fiber tractography: 1) initial value, i.e., finding fibers that start at some chosen area of the brain, 2) boundary value, i.e., finding fibers that connect two preselected brain regions. Thus we recast the fiber tractography algorithm as the determination of the most probable path either starting at a selected location or connecting two spatial locations, and seek a general probabilistic framework that can accommodate various local diffusion models and yet can incorporate the structure of extended pathways into the inference process. In this case, the problem of tractography from DWI data can be reformulated as the determination of the probability of paths on a 3-D lattice between two given points where the probability of a path passing through any particular point is *not* equiprobable, but is weighted according to the locally measured diffusion characteristics.

The essential problem at the core of the tractography problem is the estimation of macroscopic structure from microscopic measurements. In this paper we present a formulation of the tractography problem based upon a recently formulated general theory for understanding information flow in a disordered lattice. This theory, called *entropy spectrum pathways*, or *ESP*[34], is used to infer the spectra of the most probable global pathways (in this case, fiber tracts) in a nonuniform lattice (the sampled DWI data) based upon prior information about the local coupling structure of lattice (in this case estimated from the local measurements of the diffusion). The method is generalized to utilize multi-scale diffusion information that is available in multi-shell DWI datasets by extending the mechanism of streamlines generation using a Hamiltonian formalism and a diffusion-convection (Fokker-Plank) description of signal propagation though multiple scales [34]–[36].

II. Reformulation of the EAP Problem

As shown below, the ESP framework allows for the incorporation of both measured data and prior information into the estimation procedure. It is thus essential that the description of the data be as general and complete as possible. A general description of the measured DW-MRI data is provided by the EAP formalism [25]–[28]. In this section we reformulate the problem in order to provide a very general characterization amenable to numerical implementation, and to bring out some of the essential spatial scales that inform our application of ESP.

The DW-MRI signal $W(\boldsymbol{r},\boldsymbol{q})$ measured in both $\boldsymbol{r}$ and $\boldsymbol{q}$ space can be expressed in terms of both the spin density $\rho(\boldsymbol{r})$ and the average propagator $p_\Delta(\boldsymbol{r},\boldsymbol{R})$ using the narrow pulse approximation [30] as

$$W(\boldsymbol{r},\boldsymbol{q}) = \int Q(\boldsymbol{r},\boldsymbol{R}) e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} d\boldsymbol{R} \tag{1}$$

where $\boldsymbol{r}$ is the voxel coordinate, $\boldsymbol{q} = \gamma \boldsymbol{G}\delta/2\pi$, with $\boldsymbol{G}$ and δ being the strength and duration of the diffusion-encoding gradient and γ the gyromagnetic ratio of protons and the function $W(\boldsymbol{r},\boldsymbol{q})$ is the Fourier transform (in the diffusion displacement coordinate $\boldsymbol{R}$, defined as a change in particle position over time t, $\boldsymbol{R} = \boldsymbol{r}(t_0+t) - \boldsymbol{r}(t_0)$) of the weighted spin density function [23]

$$Q(\boldsymbol{r},\boldsymbol{R}) = \rho(\boldsymbol{r}) p_\Delta(\boldsymbol{r},\boldsymbol{R}) \tag{2}$$

that scales (or weights) the spin density with the average propagator $p_\Delta(\boldsymbol{r},\boldsymbol{R})$ at each observed voxel.

To find an expression for the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ we will use the plane wave expansion in spherical coordinates with $\boldsymbol{q} = q\hat{\boldsymbol{q}}$ and $\boldsymbol{R} = R\hat{\boldsymbol{R}}$, where $q = \|\boldsymbol{q}\|$ and $R = \|\boldsymbol{R}\|$

$$e^{i\boldsymbol{q}\cdot\boldsymbol{R}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(qR) Y_l^{m*}(\hat{\boldsymbol{q}}) Y_l^m(\hat{\boldsymbol{R}}) \tag{3}$$

where $j_l(qR)$ is the spherical Bessel function of order l and $Y_l^m(\hat{\boldsymbol{q}}) = Y_l^m(\Omega_{\hat{\boldsymbol{q}}}) = Y_l^m(\theta_q, \phi_q)$ is the spherical harmonic with θ_q and ϕ_q being the polar and azimuthal angles of the vector $\boldsymbol{q}$, and similarly for the vector $\boldsymbol{R}$.

The product $j_l(qx)Y_l^m(\hat{\boldsymbol{x}})$ represents the basis function for the spherical wave expansion [37]. These basis functions can be obtained as solutions of Helmholtz's wave equation [38]

$$\nabla^2 f + q^2 f = 0. \tag{4}$$

This representation suggests an interesting possibility of treating the problem of fiber tractography for diffusion weighted MRI data using the techniques of geometrical optics in inhomogeneous media [39]. We will discuss this point in more details in Section III-C.

The above basis functions are composed of radial (spherical Bessel j_l) and angular (spherical harmonic Y_l^m) parts, where the spherical harmonics $Y_l^m(\hat{\boldsymbol{x}})$ are the eigensolution of the angular part of the Laplacian with the eigenvalues $\lambda_l = -l(l+1)$

$$\nabla_\Omega^2 Y_l^m = \lambda_l Y_l^m. \tag{5}$$

The spherical harmonic Y_l^m of degree l and order m allows separation of the θ and ϕ variables when expressed using associated Legendre polynomials P_l^m of order m as

$$Y_l^m(\theta, \phi) = c_{l,m} P_l^m(\cos\theta) e^{-im\phi} \tag{6}$$

where $c_{l,m}$ is the normalization constant

$$c_{l,m} = \sqrt{\frac{2l-1}{4\pi} \frac{(l-m)!}{(l+m)!}}$$

chosen to guarantee the orthonormality condition

$$\int_0^{\pi} \int_0^{2\pi} Y_l^m Y_{l'}^{m'\star} \sin\theta d\theta d\phi = \delta_{ll'}\delta_{mm'}.$$

The radial component $j_l(qx)$ of (3) is obtained as the eigenfunction of the radial Laplacian

$$\nabla_r^2 j_l = -\left(q^2 + \frac{\lambda_l}{r^2}\right) j_l \tag{7}$$

with the orthonormality conditions

$$\int_0^{\infty} j_l(qx) j_l(q'x) x^2 dx = \frac{\pi}{2q^2}\delta(q - q').$$

This allows us to reconstruct the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ using the spherical wave decomposition as

$$Q(\boldsymbol{r}, \boldsymbol{R}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l Y_l^m(\hat{\boldsymbol{R}}) s_{lm}(\boldsymbol{r}, R) \tag{8}$$

where

$$s_{lm}(\boldsymbol{r}, R) = \int W(\boldsymbol{r}, \boldsymbol{q}) j_l(qR) Y_l^{m*}(\hat{\boldsymbol{q}}) d\boldsymbol{q}. \tag{9}$$

This representation offers a concise and intuitively clear quantitative description of the local diffusion in terms of a clearly defined expansion order on which can be based decisions of optimal fitting. We would like to mention, that although there are several different bases proposed to describe MR signal in q-space [25], [40], [41], finding the best representation of q-signal on partially acquired grid was not an intent of our paper. We reused existing fast and robust algorithms that we have developed for this computation [37]. Our implementation is flexible and does have a choice of several filters able to significantly reduce ringing artifacts.

It should be kept in mind that the typical scales for the voxel coordinate $\boldsymbol{r}$ and for the dynamic displacement $\boldsymbol{R}$ in current diffusion weighted MR experiments are vastly different. For the time scale over which the individual measurements in DWI are typically made (≈ 50 ms) the free diffusion root mean squared distance is $\langle x^2 \rangle^{1/2} \approx 20\ \mu$ and thus much smaller than typical voxel dimensions ($\approx 1\ \text{mm}^3$, at best). Hence, with high degree of accuracy it can be assumed that the average propagator in the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ only influences the nearest neighbor voxels through the dynamic displacement $\boldsymbol{R}$ dependence. The entropy spectrum pathway (*ESP*) formalism presented in [34] is well suited for taking nearest neighbors into account. We would like to emphasize that although the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ only influences the nearest neighbors through the dynamic displacement $\boldsymbol{R}$ dependence in the current sample acquisition framework, it also contains the global scale variations through the $\boldsymbol{r}$ dependence. Hence, even in the limit when the diffusions times are vanishingly small (such as in high gradient strength systems) our approach does not break down, but simply operates as a global method.

III. Tractography Guided by Entropy Spectrum Pathways

A. Summary of ESP Theory

The entropy spectrum pathways (ESP) theory [34] is an extension of the maximum entropy random walk [42], [43] and concerns the very general problem of random walks on a defective or disordered lattice. There are several key findings provided by the ESP theory. First, the pathways of the random walk are determined by prior information concerning the structure and relationships of the lattice points, and therefore the ESP represents a flow of *information*, rather than representing an actual physical process. This view facilitates the use of ESP within a wide range of practical problems related to connectivity. Second, it is possible to characterize multiple pathways, ranked according to their entropy, all of which contribute to the flow of information on the lattice. Thirdly, the interesting localization phenomenon previously noted [42], [43] can be understood in terms of the eigenstructure of the lattice. Fourthly, the local interactions that inform the generation of global structure can be based upon whatever coupling information the user has available. This coupling can take on a general form, and it was

shown that this property can be understood in terms of potential theory [34]. In the case of nearest neighbor coupling with a "binary" potential (on or off), the problem reduces to the computation of the eigenstructure of the adjacency matrix of accessible (nondefective) lattice locations. But the more general formulation of ESP facilitates the use in practical applications such as we present here. In this section we will briefly summarize and reformulate the ESP formalism within the context of the tractography problem.

The objective is to calculate the probability that a spin, or "particle," starting from an initial spatial location $x(t_0)$ at initial time $t_0 = 0$ diffuses to a second location $x(t)$ at a later time t. While the underlying structure we wish to estimate is assumed continuous (being comprised of tissue fibers), the spatial locations x at which the measurements are acquired are assumed to be from DWI images and thus discretized to a 3-D Cartesian spatial grid. However, the temporal discretization to be employed is a fictitious construct used to implement a random walk model, due to the above mentioned difference in scales of the voxel coordinates and the dynamic displacement. Our space-time points are defined on the true voxel spatial grid but on a diffusion pseudo-time grid whose increments are much larger than the experimental time scale. Simulating diffusion in this way lends itself to two different interpretations. One view of this process is to see it as diffusion that is allowed to take place for much longer than the measurement time, or, equivalently, that the process is in equilibrium and thus long time behavior is well-represented by the snapshots in time provided by the experimental data. However, another viewpoint, that of ESP, and the one we adopt, is that the simulation is of the flow of information constrained by the physical measurements. The entire process is one of estimating a macroscopic phenomena from local measurements, using by prior information.

The ESP theory ([34]) ranks the optimal paths within a disordered lattice according to their path entropy. Associated with the kth path is a transition probability

$$p_{ij}^{(\infty)} = \sum_{k=1}^{n} p_{ijk}^{(\infty)} \tag{10}$$

$$\text{where} \qquad p_{ijk}^{(\infty)} = \frac{Q_{ij}}{\lambda_k} \frac{\psi_j^{(k)}}{\psi_i^{(k)}} \tag{11}$$

where λ_k and $\psi^{(k)}$ are eigenvalues and eigenvectors of the coupling matrix Q defined as

$$Q_{ij} = e^{-\Lambda_{ij}}. \tag{12}$$

Hence each transition probability is associated with a standard eigenvalue equation

$$\sum_j Q_{ij} \psi_j^{(k)} = \lambda_k \psi_i^{(k)}. \tag{13}$$

This matrix defines the interactions between locations on the lattice and is called the *coupling matrix*. The Λ_{ij} are Lagrange multipliers that define the interactions that can be seen as local potentials that depend on some function of the spatial locations x_{ij} on the lattice. For each transition matrix (11) there is a unique stationary distribution associated with each path k

$$\mu^{(k)} = \left[\psi^{(k)}\right]^2 \tag{14}$$

that satisfies

$$\mu_i^{(k)} = \sum_j \mu_j^{(k)} p_{ijk}^{(\infty)} \tag{15}$$

the first of which, $\mu^{(1)}$, corresponds to the maximum entropy stationary distribution. Considering only $\mu^{(1)}$, note that if the Lagrange multipliers take the form

$$\Lambda_{ij} = \begin{cases} 0, & \text{connected} \\ \infty, & \text{not connected} \end{cases} \Rightarrow Q_{ij} = e^{-\Lambda_{ij}} = \begin{cases} 1 \\ 0 \end{cases} \tag{16}$$

then Q becomes simply an adjacency matrix, the single maximum entropy distribution constructed from this adjacency matrix is maximum entropy random walk [42]. However, one major significance of the ESP theory to the present problem is that is ranks *multiple paths*, and these paths can be constructed from *arbitrary* coupling schemes through Q_{ij}. For each of the stationary distributions is associated a path related to the localization of information related to the eigenstructure of the disordered lattice. The key feature is that the *local* transition probabilities between nodes depend on the *global* structure of the graph through the eigenvectors $\psi^{(k)}$. In practical applications, the lattice can be described in terms of n pathways constructed from the first n eigenvectors of the potential matrix (in decreasing order of the eigenvalues).

We emphasize that we do not presume to be explicitly modeling the diffusion over the paths, since we know that the diffusion length over the typical time-scale of a DWI experiment is typically far smaller than a voxel dimensions. Rather, this procedure is viewed as one of estimation and thus the construction of the ranked maximum entropy paths—those that are most unbiased with respect to the measured data and the prior information (the lattice couplings) while satisfying the initial and/or final conditions. In this view the problem is one of estimating the global connectivity from the local diffusion characteristics.

B. Multiple Scale Coupling

The estimation of the local and global tissue structure over multiple scales using DW-MRI data can be investigated within the ESP framework by viewing the data as measurements on a 3-D lattice in which each voxel is ascribed a "potential" that is related to its coupling with neighboring voxels. An important feature of the ESP theory is that this potential is very general in form. This is critical to its application in the current problem. While it was shown for a binary coupling (i.e., wherein the coupling matrix reduced to the adjacency matrix with 0/1 elements) [34], in the current problem we will incorporate a strength of coupling that reflects the local interaction of voxel data. In order to do this we first symmetrize the spin density function

$$\begin{aligned} \mathcal{Q}_{ij}(\boldsymbol{r}_i, \ell) &= \mathcal{Q}_{ji}(\boldsymbol{r}_j, \ell) \\ &= \frac{1}{2}\left[Q\left(\boldsymbol{r}_i, (\boldsymbol{r}_i - \boldsymbol{r}_j)\ell\right) + Q\left(\boldsymbol{r}_j, (\boldsymbol{r}_j - \boldsymbol{r}_i)\ell\right)\right] \end{aligned} \tag{17}$$

where ℓ represents the dimensionless ratio of scales of dynamic displacement $\boldsymbol{R}$ to the spatial (voxel) scales $\boldsymbol{r}$, and, then sum all relevant scales included in the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ by the dependence on the dynamic displacement $\boldsymbol{R}$

$$\overline{\mathcal{Q}}_{ij} = \overline{\mathcal{Q}}_{ji} = \int_{\ell_{\min}}^{\ell_{\max}} \mathcal{Q}_{ij}(\boldsymbol{r}_i, \ell) d\ell. \tag{18}$$

Here, we used a symmetric input from voxels i and j by taking the line integral of the spin density function $Q(\boldsymbol{r}, R_\ell)$ along the direction $\boldsymbol{R}_\ell = \boldsymbol{r}_i - \boldsymbol{r}_j$ between those voxels and take into account only a subset of spatial scales that can contribute to this interaction (from $\ell_{\min}$ to $\ell_{\max}$). Since the typical scales for the voxel coordinate $\boldsymbol{r}$ in current diffusion weighted MR experiments are much larger then the scales for the dynamic displacement $\boldsymbol{R}$ (20 μ versus 1 mm), the coupling can be limited to nearest neighbor effects taking $\ell_{\min} = 0$ and $\ell_{\max} = \infty$ to calculate a coupling potential $\overline{\mathcal{Q}}_{ij}^{\infty}$.

We would like to emphasize, that in nearest neighbor coupling evaluation [(18)] we do not use solid angle integration. The appropriate choice of filters and order of angular resolution in the SWD allows us to replace the costly integration of geometrically complex coupling between noisy multiple peaked dODF from neighboring voxels with the fast and simple line integration across all radial scales.

This form of the coupling potential is then used in (13) to obtain the relevant eigenvalues and eigenvectors

$$\sum_j \overline{\mathcal{Q}}_{ij}^{\infty} \psi_j^{(k)} = \lambda_k \psi_i^{(k)}. \tag{19}$$

The kth eigenvalue and eigenvector can be used to generate the transition probabilities [(11)] but in addition we also generate the scale dependent transition probabilities

$$\mathcal{P}_{ijk}(\boldsymbol{r}_i, \ell) = \frac{\mathcal{Q}_{ij}(\boldsymbol{r}_i, \ell)}{\lambda_k} \frac{\psi_j^{(k)}}{\psi_i^{(k)}} \tag{20}$$

that will be equal to the total transition probability P_{ij} when integrated over all scales ℓ. As those probabilities only describe transitions between nearest neighbors they can be expressed as a scale dependent function $\mathcal{P}_\ell(\boldsymbol{r}, \boldsymbol{R})$. We will also generate the equilibrium probabilities $\mu^{(k)} = [\psi^{(k)}]^2$.

As a concluding remark to this section, we would like to reiterate that the general problem of tractography is necessarily one of multiple scales because the local diffusion occurs on the microscale and the tracts are on a macroscale. The entire point of our approach is that it enables a characterization of the problem in terms of information at these multiple scales.

C. Generation of Optimal Paths

Our goal is either to construct a pathway between an initial spatial location a and a final spatial location b or to trace a pathway incrementally starting from an initial location a. In both cases we are interested in "the most probable" pathways, i.e., we would like to constrain our local search by the global entropy structure. Therefore, our interest is not in the final equilibrium distribution μ^* but in the pathway to it. We are therefore interested in the dynamics of the probability and want to compute the path that maximizes the entropy at each step, and thus results in the final (equilibrium) distribution μ^* at time τ.

The scale dependent transition and equilibrium probabilities obtained in the previous section can naturally define the global entropy field that shapes the flow of information and allow finding optimal paths. In the limit of long pathway lengths (or large time τ) and under the Markovian assumption, the rate of entropy change $S_\ell(\boldsymbol{r}_i)$ can be expressed at each location $\boldsymbol{r}_i$ as [44], [45]

$$S_\ell(\boldsymbol{r}_i) = -\sum_k \mu^{(k)} \sum_j \mathcal{P}_{ijk}(\boldsymbol{r}_i, \ell) \ln \mathcal{P}_{ijk}(\boldsymbol{r}_i, \ell). \tag{21}$$

The most straight forward way to include this multi scale structure of the global entropy field is by taking into account that the conservation of probability in general includes not only the diffusive component (as for example used by [30] for obtaining the expression of EAP in single mode homogeneous self-diffusion), but also has the convective part [34]–[36]

$$\partial_t P + \nabla \cdot (LP \nabla S) = \nabla \cdot D \nabla P \tag{22}$$

here P is the probability, S is the entropy, and L and D are coefficients (in general either tensors or functions of the coordinates) that characterize local convective and diffusive scales ($L = \kappa D$). This Fokker-Planck equation, with the potential equal to the entropy, connects the global structure of the probability with the local structure of the lattice through the local structure of the entropy.

The current state-of-the-art approaches used for fiber tractography in DTI/DWI data require splitting this problem in two parts: first, obtain the EAP from the diffusion only subsystem,

$$\partial_t P = \nabla_{\boldsymbol{R}} \cdot D \nabla_{\boldsymbol{R}} P \tag{23}$$

and second, solve the convective part (averaged over all the dynamic displacement scales $\boldsymbol{R}$)

$$\int \left[\partial_t P + \nabla_{\boldsymbol{r}} \cdot (LP \nabla_{\boldsymbol{r}} S)\right] d\boldsymbol{R} = 0 \tag{24}$$

by simple local tracing of one (DTI) or several (DWI) principal fiber directions. Unfortunately, this decoupling results in only the *local* diffusion information derived from EAP being used at the fiber tracking stage

To illustrate this point, we will first assume the entropy gradient fixed and will show how it leads to the current tractography. In this case the convective part of (22) in the eikonal approximation provides a simple expression for the Hamiltonian $\mathcal{H}(\omega, \boldsymbol{k}, \boldsymbol{r})$—the function of canonical coordinates that defines the dynamics. (For mechanical systems this function is simply the total energy, which is conserved in motion. For more complex systems it does not necessarily corresponds to energy, but still describes conservation laws of the system)

$$\mathcal{H}(\omega, \boldsymbol{k}, \boldsymbol{r}) = \frac{1}{V_{\mathbf{R}}} \int \left[-\omega^2 + (\boldsymbol{k} \cdot L \nabla S)^2\right] d\boldsymbol{R} \tag{25}$$

where an input from all dynamic displacement scales is included, as formally both L and S may depend on both $\boldsymbol{R}$ and $\boldsymbol{r}$

Finding the characteristics (or rays) of (22) will describe how the signals propagate and can be accomplished by integrating a set of ordinary differential equations of the Hamilton-Jacobi type

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{k}}, \qquad \frac{d\boldsymbol{k}}{dt} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{r}}. \tag{26}$$

The current fiber tractography methods in general do not emphasize or discuss the notion of global entropy, but implicitly assume the local behavior of the entropy gradient, directing it along some of the major axes of the local diffusion/convection tensor $\boldsymbol{L} = \kappa \boldsymbol{D}$, i.e., $\nabla S = \boldsymbol{\psi}$, where $\boldsymbol{\psi}$ is the eigenvector of $\boldsymbol{L} \cdot \boldsymbol{\psi} = \lambda \boldsymbol{\psi}$. Under the assumption of scale independent diffusion [i.e., $\boldsymbol{D}(\boldsymbol{r}, \boldsymbol{R}) \equiv \boldsymbol{D}(\boldsymbol{r})$] the Hamiltonian (25) then becomes

$$\mathcal{H}(\omega, \boldsymbol{k}, \boldsymbol{r}) = -\omega^2 + \lambda^2 (\boldsymbol{k} \cdot \boldsymbol{\psi})^2 \tag{27}$$

and the ray tracing equation simplifies to the following form:

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{k}} = 2\lambda^2 (\boldsymbol{k} \cdot \boldsymbol{\psi}) \boldsymbol{\psi} = C \boldsymbol{\psi}. \tag{28}$$

Ignoring the spatial dependence of the diffusion propagator (i.e., $C = \text{const}$) this equation is exactly in the form of Frenet equation commonly used for fiber tracking [3]. Thus, the current fiber tractography can be regarded as a fixed scale and spatially homogeneous limit of the more general Fokker-Plank formalism [(22)].

To develop a more general entropy based tractography several assumptions will be made. First, only solutions with the high enough probability will be considered, i.e., it will be assumed that in the region of interest the probabilities are sufficiently close to 1, so that it is possible to linearize both the probability and the entropy as

$$P = P_0 + P_1, \qquad S = S_0 + S_1 = S_0 - (1 + \ln P_0) P_1 \tag{29}$$

where $S_0 = -P_0 \ln P_0$ and the scale dependent transition probability $\mathcal{P}_\ell(\boldsymbol{r}, \boldsymbol{R})$ [(20)] can be substituted for P_0. Second, it is assumed that P_1 is a small correction to the equilibrium probability (i.e., $P_1 \ll P_0$). The linearized convective part of (22) can then be written

$$\partial_t P_1 = \nabla \cdot (P_1 \nabla P_0 (2 + \ln P_0)) + \nabla \cdot (P_0 (1 + \ln P_0) \nabla P_1). \tag{30}$$

Third, use of the scale dependent transition probability $\mathcal{P}_\ell(\boldsymbol{r}, \boldsymbol{R})$ for P_0 allows us to omit L in these expressions, as the diffusion anisotropy is already included in ESP calculations of $\mathcal{P}_\ell(\boldsymbol{r}, \boldsymbol{R})$ from the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ obtained with the spherical wave decomposition using (8) and (9). Fourth, no time dependence is assumed to be present in P_0 as it is time stationary, hence $\partial_t P_0 \equiv 0$. Fifth, we also have omitted the last term $\nabla \cdot D \nabla P_1$ from the (22) as it will not appear in eikonal approximation used for obtaining the ray tracing equations.

Equation (30) is a linear inhomogeneous hyperbolic equation, hence it has traveling wave solutions propagating along the characteristics. In order to formally find those characteristics we will assume a plane wave solution for P_1 [39], [46] in the form

$$\begin{aligned} P_1(\boldsymbol{r}, \boldsymbol{R}, t) &= A(\boldsymbol{r}, \boldsymbol{R}, t) e^{i\Psi(\boldsymbol{r}, t)} \\ \Psi(\boldsymbol{r}, t) &= \boldsymbol{k} \cdot \boldsymbol{r} - \omega t \end{aligned} \tag{31}$$

and then obtain a more general expression for the Hamiltonian as

$$\mathcal{H}(\omega, \boldsymbol{k}, \boldsymbol{r}) = \frac{1}{V_{\mathbf{R}}} \int [-\omega^2 + (\boldsymbol{k} \cdot \boldsymbol{X})^2 + (Y(\boldsymbol{k} \cdot \boldsymbol{k}) - Z)^2] d\boldsymbol{R} \tag{32}$$

where $\boldsymbol{X} = \nabla P_0 (2 + \ln P_0) + \nabla (P_0 (1 + \ln P_0))$, $Y = P_0 (1 + \ln P_0)$, $Z = \nabla \cdot \nabla P_0 (2 + \ln P_0)$, and, again we averaged over all dynamic displacement scales.

Hence, taking into account the global entropy gradient as well as the scale dependence of the diffusion coefficient, the fiber tracking in the geometrical optics limit can be represented in more general form, using (26) and (32), as

$$\frac{d\boldsymbol{r}}{dt} = \frac{2}{V_{\mathbf{R}}} \int \left[\boldsymbol{X} (\boldsymbol{k} \cdot \boldsymbol{X}) + 2\boldsymbol{k} Y (Y |\boldsymbol{k}|^2 - Z) \right] d\boldsymbol{R} \tag{33}$$

$$\frac{d\boldsymbol{k}}{dt} = -\frac{2}{V_{\mathbf{R}}} \cdot \int \left[(\boldsymbol{k} \cdot \boldsymbol{X}) \nabla (\boldsymbol{k} \cdot \boldsymbol{X}) + (|\boldsymbol{k}|^2 Y - Z)(|\boldsymbol{k}|^2 \nabla Y - \nabla Z) \right] d\boldsymbol{R}. \tag{34}$$

The first equation [(33)] traces the characteristics (rays) of the convective part of the original Fokker-Plank equation [(22)] under the influence of a local diffusion coupled with a global entropy gradient. This coupling is locally described by a vector $\boldsymbol{X}$. A second term (with $2\boldsymbol{k}Y$) provides some smoothing by adding "a push" in the direction of the wave vector $\boldsymbol{k}$. It also ensures that in voxels with isotropic diffusion (or with many fibers of different directions crossing) and without a global entropy gradient the ray will continue following this $\boldsymbol{k}$ direction. In the second equation [(34)] the spatial gradients are responsible for a change of the wave vector $\boldsymbol{k}$ direction and magnitude.

A schematic illustration of differences between traditional fiber tracking and geometrical optics-like tracking is shown in Fig. 1. The traditional approach defines tracts by integration of position-only function $\boldsymbol{\psi}$, that assigns the tangential direction of tracts to each location $\boldsymbol{r}$. For the geometrical optics approach, the integration takes into account both the orientation and multiple scales, through the dependence of $\boldsymbol{\psi}$ on directional angle $\boldsymbol{k}/|\boldsymbol{k}|$ and magnitude $|\boldsymbol{k}|$.

IV. Results

To evaluate practical aspects and performance of ESP guided fiber tractography we used our method to process several multiple shell multiple angle diffusion weighted MRI datasets acquired using either realistic MR phantom or real brain samples.

A practical realization of the method used in all of the experiments includes several stages that were implemented in C++ with multithreading (using pthread libraries) and runs both on Linux and Mac OS X. The ESP framework uses our spherical

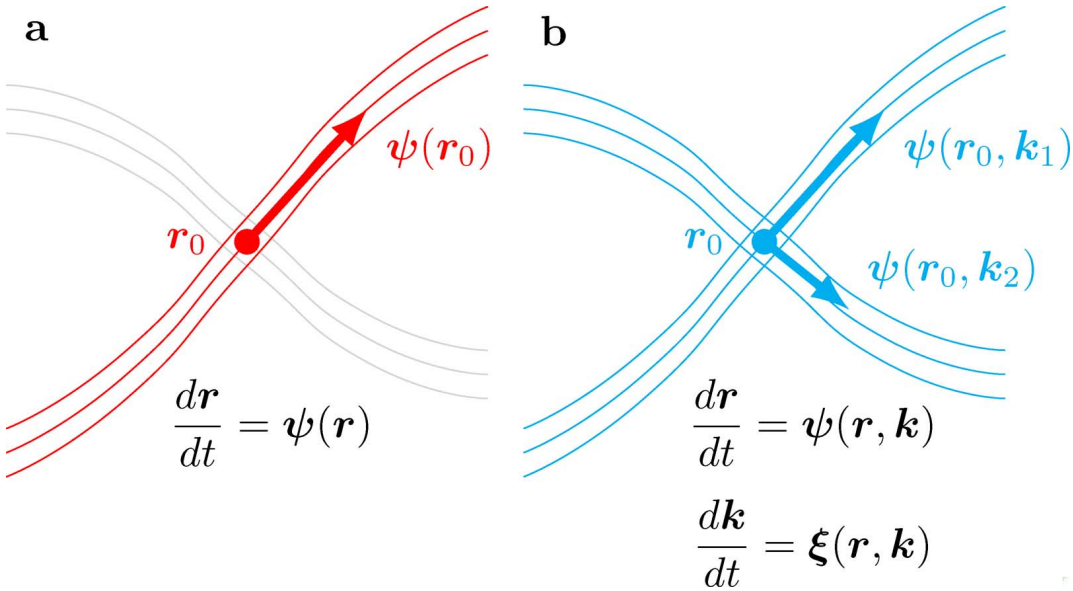

Fig. 1. Schematic illustration of (a) traditional fiber tracking based on integration of a single Frenet equation versus (b) fiber tracking that uses the geometrical optics analogy. (a) In the first case the fiber orientation vector $\boldsymbol{\psi}$ only depends on spatial location $\boldsymbol{r}_0$, hence, even at location of fiber crossing only single the most important fiber can be followed (point $\boldsymbol{r}_0$ uniquely selects single family of fibers oriented along $\boldsymbol{\psi}(\boldsymbol{r}_0)$). (b) Geometrical optics approach automatically includes dependence of $\boldsymbol{\psi}$ on both orientation $\boldsymbol{k}/|\boldsymbol{k}|$ and scale $|\boldsymbol{k}|$, hence it can effectively proceed through difficult areas of crossings of multiple fibers (the choice of fiber direction at point $\boldsymbol{r}_0$ depends on the value of the parameter $\boldsymbol{k}$, with $\boldsymbol{k}_1$ selecting the same fiber direction as in (a), and $\boldsymbol{k}_2$ corresponding to the alternative crossing family of fibers).

wave decomposition approach (SWD) [37] based on the fast Fourier and on the fast spherical Bessel transforms. The tracking algorithm uses the standard fourth order Runge-Kutta integration applied to six dimensional ($\boldsymbol{r}$, $\boldsymbol{k}$) space.

The high resolution full brain tractography processing of typical $140 \times 140 \times 96$ diffusion weighted MRI volume with four shells of 552 q-vectors using Intel Core i7-4930 K six cores (twelve threads) 3.40 GHz CPU produces around 190 K fiber tracts in 15 min. The spherical wave decomposion part of processing takes six minutes for obtaining 600 spherical wave modes at each voxel using $L_{\max}$ (or spherical harmonics order) of 10, and $N_{\max}$ (or spherical Bessels order) of 6. The tractography part uses full $140 \times 140 \times 96$ spatial grid and $7 \times 7 \times 7$ resampled k-grid and outputs $\sim$ 190 K of fiber tracts in under 9 min. The medium resolution full brain processing produces $\sim$ 90 K tracts in 5 min, with the ESP stage taking around 3 min.

The first dataset is the well-known "fiber cup" MR phantom extensively used for testing and performance evaluation of various fiber tractography approaches [47]. The phantom consists of seven fiber bundles confined in a single plane by squeezing them in between two solid disks. Diffusion-weighted image data of the phantom was acquired on the 3T Tim Trio MRI system with 3 mm isotropic resolution on $64 \times 64 \times 3$ spatial grid. Three diffusion sensitizations (at b-values $\mathrm{b} = 650/1500/2000\ \mathrm{s/mm^2}$) were collected two times for 64 different diffusion gradients uniformly distributed over a unit sphere. Several baseline ($\mathrm{b} = 0$) images were also recorded [47].

Our initial stage of processing includes restoration of the spin density function $Q(\boldsymbol{r}, \boldsymbol{R})$ using (8) and (9) from Section II. The spin density function is then used to generate symmetric scale integrated input to the coupling potential with (17) and (18). Eigenvectors and eigenvalues of the coupling potential then used for obtaining the transition probabilities using (11).

We included one of the baseline images of the fiber cup phantom in Fig. 2(a) to emphasize an interesting and important feature of the ESP approach. The baseline image (as well as other diffusion weighted images not shown here) clearly shows bright artifacts at the interface boundaries of the disks used for phantom manufacturing. The overall effects of these artifacts are significant with the brightest spots located either at the very ends of the fiber bundles where they are cut by the disks or even at the circular boundary of the disks themselves. The expanded version of one of these areas is shown in Fig. 2(d).

The currently standard analyses that are based on the maps of the apparent diffusion coefficient (ADC) and the fractional anisotropy (FA) (available in Fig. 2 of [47]) also favor those regions by assigning higher anisotropy and diffusion values. As a result many of the current streamline tracing tractography approaches do not see the actual ends of the fiber bundles and continue tracts through the circular disks interfaces.

But the transition probability map shown in Fig. 2(b) that was generated by the ESP clearly emphasizes the separation between the fiber bundles and the circular disk interface [with the same area enlarged in Fig. 2(e)]. Using simple thresholding [Fig. 2(c) and (f)] allows identification of the ends for all of the fiber bundles. What is equally important is that ESP correctly restores the contrast that has been lost at the crossing fiber areas that can clearly be seen in b0 as well as in the diffusion weighted images for different gradient directions. As previously mentioned, this helps the geometrical optics algorithm to find the correct continuation of rays in voxels with isotropic diffusion (or with many fibers of different directions crossing).

The local samples of multiple scales of the transition probabilities calculated by the ESP method are presented in Fig. 3.

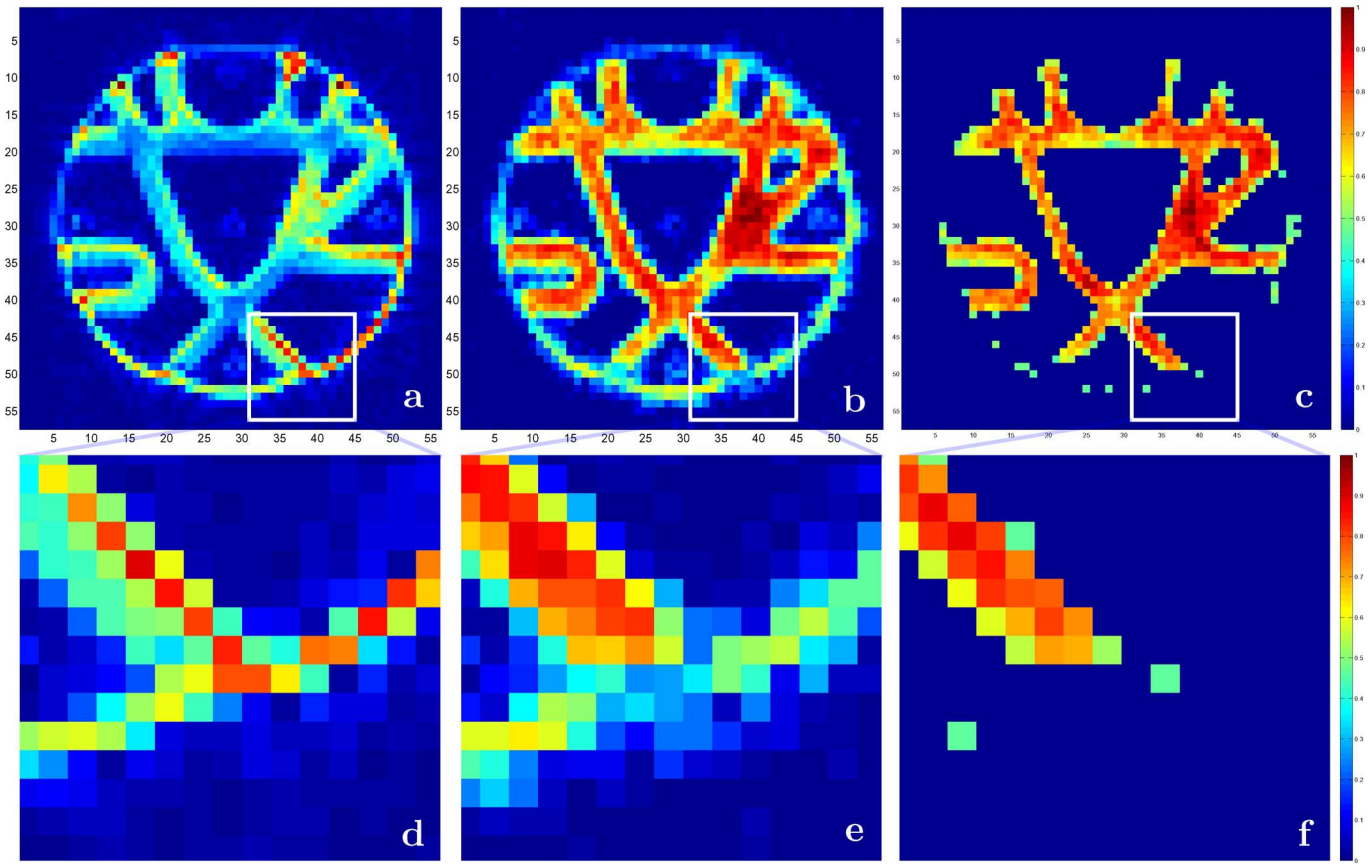

Fig. 2. Central horizontal 64×64 slice of $64 \times 64 \times 3$ MR fiber cup phantom [47]; (a) one of the baseline $b = 0$ images and (d) the expanded version of the white square area clearly showing strong signal at the fiber end and at the circular disc interface; (b) the map of the largest ESP transition probability values with (e) the enlargement of the same area showing the resolution of the fiber end; (c) and (f) the map of equilibrium probability distribution ($\mu^* = [\phi^{(1)}]^2$) thresholded at 0.45 with perfect identification of all fiber ends and overall area occupied by fiber bundles.

One of the crossing fiber areas, enlarged in the right panel, clearly shows existence of different fiber directions in different scales of the transition probabilities.

We would like to stress two important aspects of our method. First, Fig. 3 shows the multi-scale transition probabilities [as expressed by (20)] rather than the EAP (or dPDF). The transition probabilities were derived not just from the local diffusion (used in EAP/dPDF). It also takes into account the nonlocal coupling between voxels by calculating the global eigenmodes and updating the probabilities according to the structure of the corresponding global eigenvectors. In this respect, the transition probabilities are more fundamental quantities than the locally derived EAPs or dPDFs.

Second, the use of multiple scales enables the geometrical optics-like approach presented here to find the correct path even when the angular resolution is relatively low. To illustrate this fact we included in Fig. 4 two possible tracking scenarios, (a) using just single scale transition probabilities, and (b) with transition probabilities that include multiple scales. The single scale tracking finds only one bundle of fibers, and either breaks the second set of fibers or wrongly connects it to the first bundle. But with multiple scale transitional probabilities, the second set of fibers is found correctly with the geometrical optics-like tracking even at this relatively low angular resolution.

Utilization of high angular resolution locally (in an isolated voxel) and without the incorporation of multiple scales and global connectivity does not necessarily guarantee detection of crossing fibers. For example, it is not possible to detect the second direction of fibers in an EAP-like function [Fig. 4(c)] of a single voxel from the crossing area of Fig. 3(b). This EAP-like map was obtained from single voxel diffusion data using the original shell of 60 direction q-values with high angular resolution 72 mode spherical harmonics expansion. Importantly, our approach using multi scale transition probabilities and global connectivity information can identify crossings from significantly lower angular and radial resolution.

A direct comparison with the fiber cup results is shown in Fig. 5(a). Several possible fiber tracts obtained by integrating the equations of geometrical optics rays for Fokker-Plank formalism [(33) and (34)] are displayed. The phantom includes several different types of fiber crossings (at different angles), as well as kissings and splits/joints. To illustrate one peculiar feature of the approach based on geometrical optics we included the blow up of one of the fiber tracts. In the boundary area where the underlying fibers end abruptly the ray tracing algorithm may produce a reflection of the ray from the interface and proceed following the same (or neighboring) fiber backwards. In Fig. 5(b) the reflection happens at two ends of the fiber and results in a closed horseshoe-like loop that is being transversed

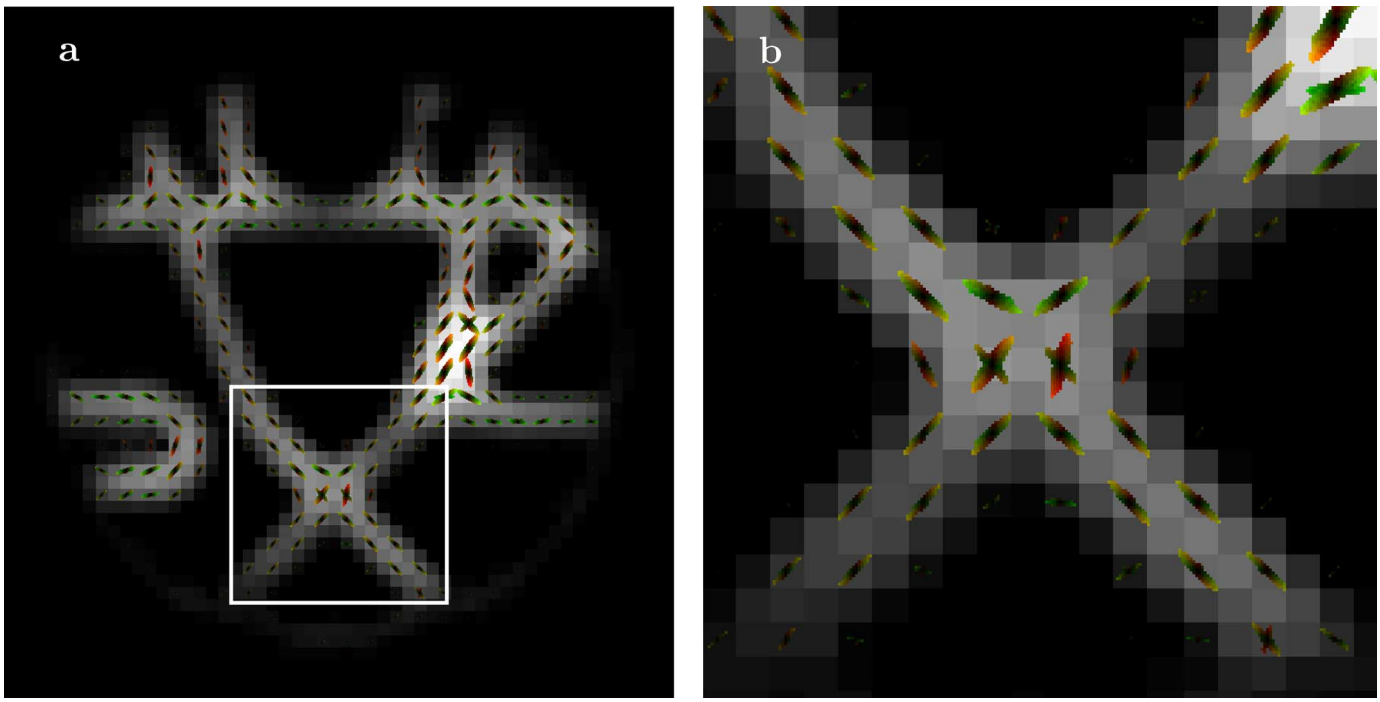

Fig. 3. (a) Local samples of different scales of transition probabilities obtained by the ESP, with (b) blow up of the center crossing fibers area.

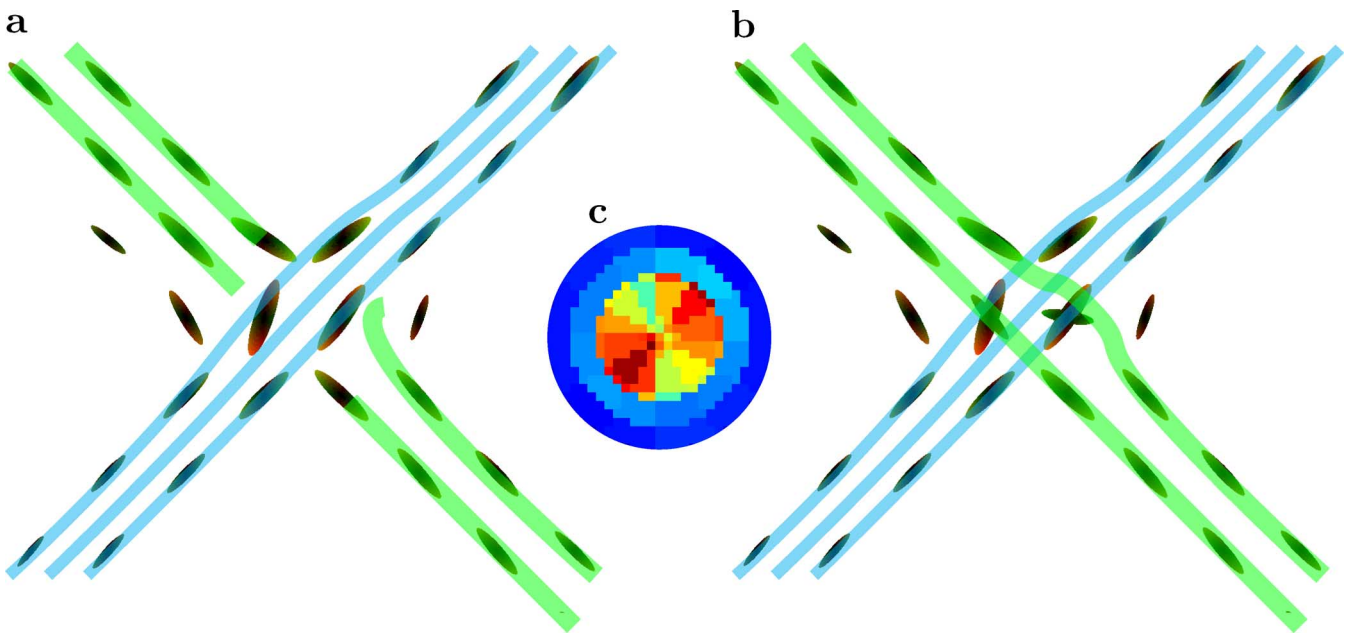

Fig. 4. (a) One possible tracking through fiber crossing area shown in Fig. 3(b) using only single scale transition probabilities. With low angular resolution only a single set of crossing fibers can be identified, and the second bundle will either be broken (or connected to the first one). (b) Tracking with multiple scales and coupling information are used in obtaining transition probabilities. Even with this relatively low angular and radial resolution the presence of multiple scales enables geometrical optics-like tracking through the difficult area. (c) A map of EAP-like function for a single voxel from the crossing area. The map was obtained using only local diffusion data from the original shell of 60 direction q-values with high angular resolution 72 mode spherical harmonics expansion. The map shows that even with high angular resolution only one fiber direction can be reliably identified. Our multi-scale approach incorporating global connectivity information can reliably identify crossing in the calculated transition probabilities.

back and forth many times. Of course, we included this close-loop tract only as an example as identifying the reflection regions and cutting the fibers at these points instead of reflecting would be a fairly easy task.

To generated fiber tracts we selected seed points [Fig. 5(d)] by thresholding the map of the ESP equilibrium probabilities. Using all 512 selected seed points the algorithm produced 372 total fiber tracts. All but two fiber tracts are topologically equivalent to one of the seven tracts shown in Fig. 5(a). Two fiber tracts (one red and one green) show the end points switched. These two incorrect tracts give 0.5% false positive error rate. In general simple post processing can remove even those two outliers using, for example, the total fiber tract length or the distinct change in the fiber curvature as a source of discrimination. Also all those fibers can be evaluated using for example Tracktometer [48]. We will surely be using it when/if we are directly involved

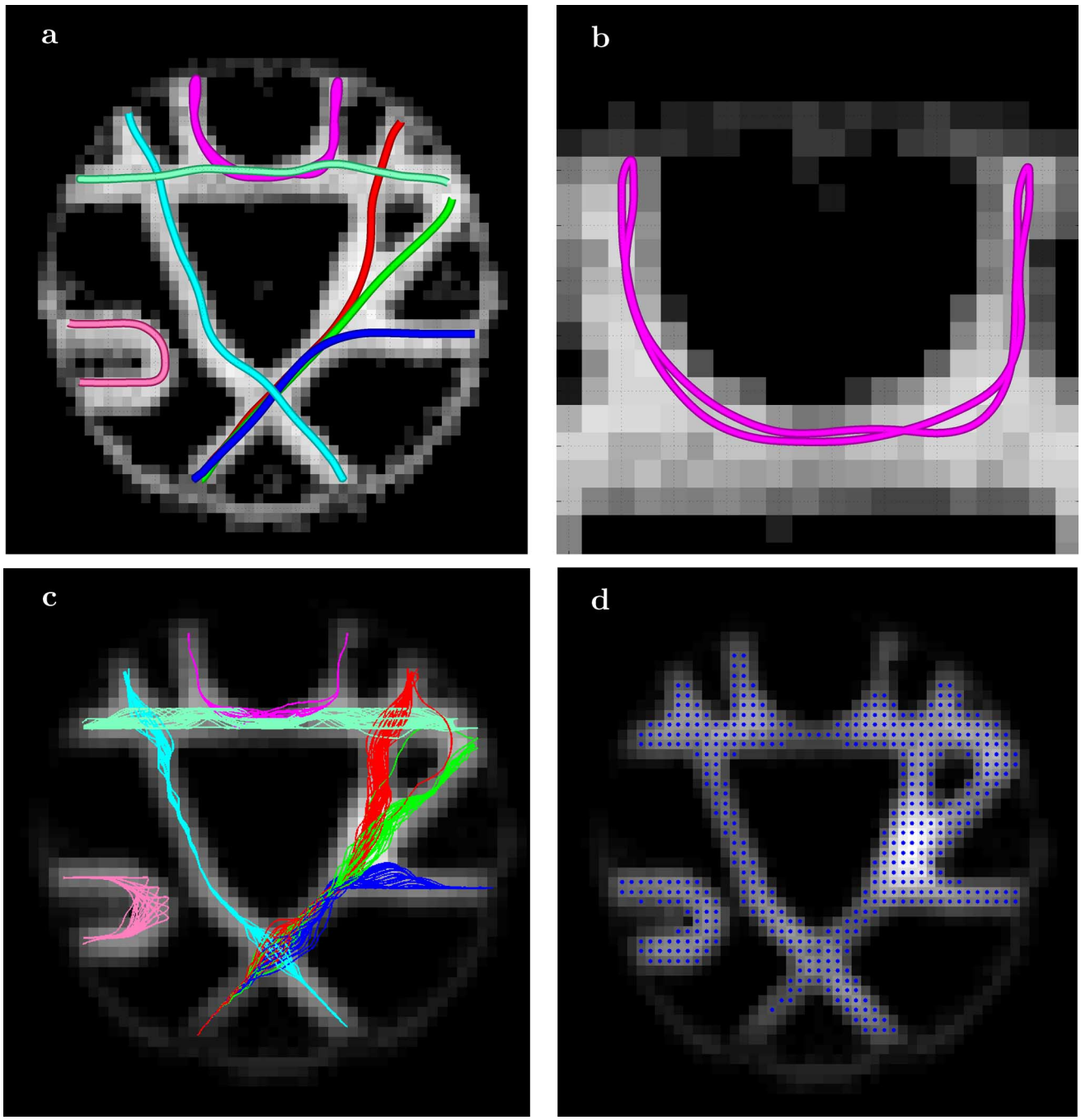

Fig. 5. (a) Several fiber tracts produced by geometrical optics ray tracing of the Fokker-Plank equation using the ESP equilibrium probability distribution shown in Fig. 2. (b) Blow up of one of the tracts showing an interesting property of the ray tracing: the tracts reaching the fiber ends can get reflected at the boundary and form a closed cycle repeating itself again and again. (c) All fiber tracts obtained using seeds (d) selected with single threshold from the ESP equilibrium probability distribution. The processing of 512 total seed produced 372 fiber tracts with only two (one red and one green) incorrectly finished at the neighboring ending point. These results correspond to slightly more than 0.5% false positive rate.

in developing, tuning, optimizing a full-fledged tractography processing pipeline. However, the main intention or this paper is to show the feasibility of a new tracking paradigm. Our motivation for using the FiberCup data was precisely *because* it facilitates comparison of our results with previously published results, which we have successfully demonstrated. Further detailed analysis on a phantom that has a rather limited connection to human brain data is ultimately of little importance.

For human brain ESP tractography we collected multi b-shell multi q-angle DWI dataset on the GE MR750 3T scanner at the UCSD Center for FMRI using a multi-band blipped-CAIPI EPI method [49] with a GRAPPA reconstruction [50]. Each data set was collected with both forward and reversed phase encoding polarity in order to perform a "topup" distortion and eddy current correction [51] using FSL [52]. The dataset contains three shells at $b = 1000, 2000,$ and 3000 s/mm^2. Each

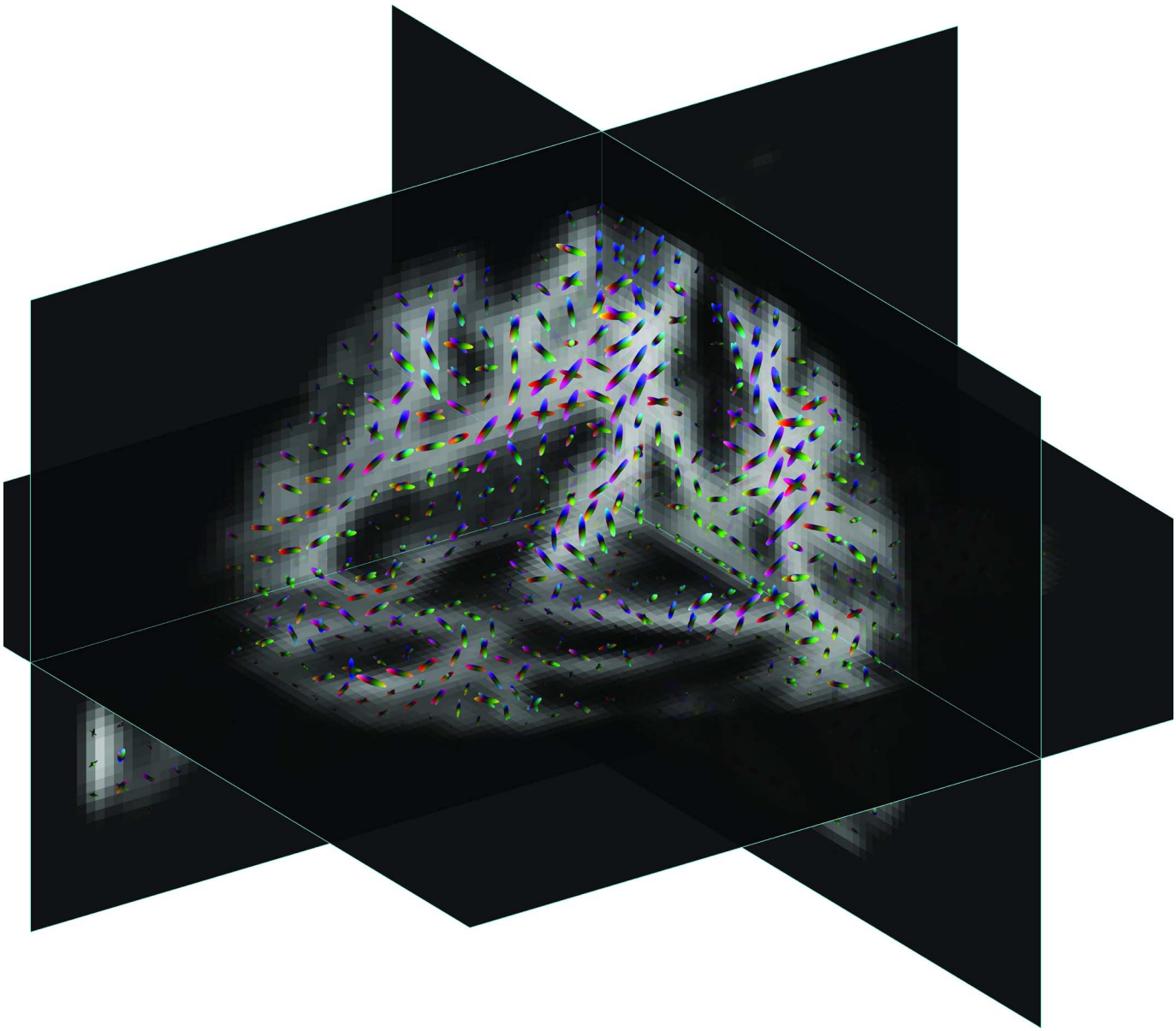

Fig. 6. Different slices of three dimensional equilibrium probabilities $\mu^{(1)}$, or square of the first ESP eigenvector (14) (shown by grayscale background), obtained using diffusion weighted images of human brain with local samples of different scales of transition probabilities shown by directionally colored ellipsoids.

b-shell uses different number of q-values, with 30 angles for $b = 1000$ s/mm^2, 45 angles for $b = 2000$ s/mm^2, and the largest at 60 angles for $b = 3000$ s/mm^2.

Several slices of three dimensional eigenvector map obtained by the ESP solution are shown in Fig. 6. The local samples of transition probabilities are shown by directionally colored ellipsoids. The multi-scale structure of the ESP approach can be used for identifying fiber crossings—it can be seen clearly in many of these samples that different scales show different main fiber direction.

To illustrate the practical ability of geometrical optics-like tracking of fibers though those "difficult" areas of multiple fibers with different orientations we generated fiber tracts for several sets of seed points (Fig. 7) located in the areas of corpus callosum and longitudinal fasciculus that are known to have multiple overlaps in both inferior and fronto-occipital regions [53]. The 135 corpus callosum seed points were grouped in blocks of three consecutive voxels coronally and two or three voxels vertically. In the fasciculus region two sets of seed were selected in left and right hemispheres with 30 seeds each again grouped in blocks by three consecutive voxels vertically and five consecutive voxels coronally. The total number of seeds were 195 voxels resulting in 195 distinct fiber tracts.

All the seeds were selected in mostly in the regions with predominantly single fiber orientation. In the multi-scale ESP guided geometrical optics-like approach, the tracking algorithm is able to follow tracts across voxels that contain a mixture of fibers of different orientations and select the correct path based on the combination of local and global parameters (Fig. 7).

For a more "difficult" starting point we first have chosen several seed voxels in the area around the splenium of the corpus callosum and up to the internal capsule. The five fiber tracts,

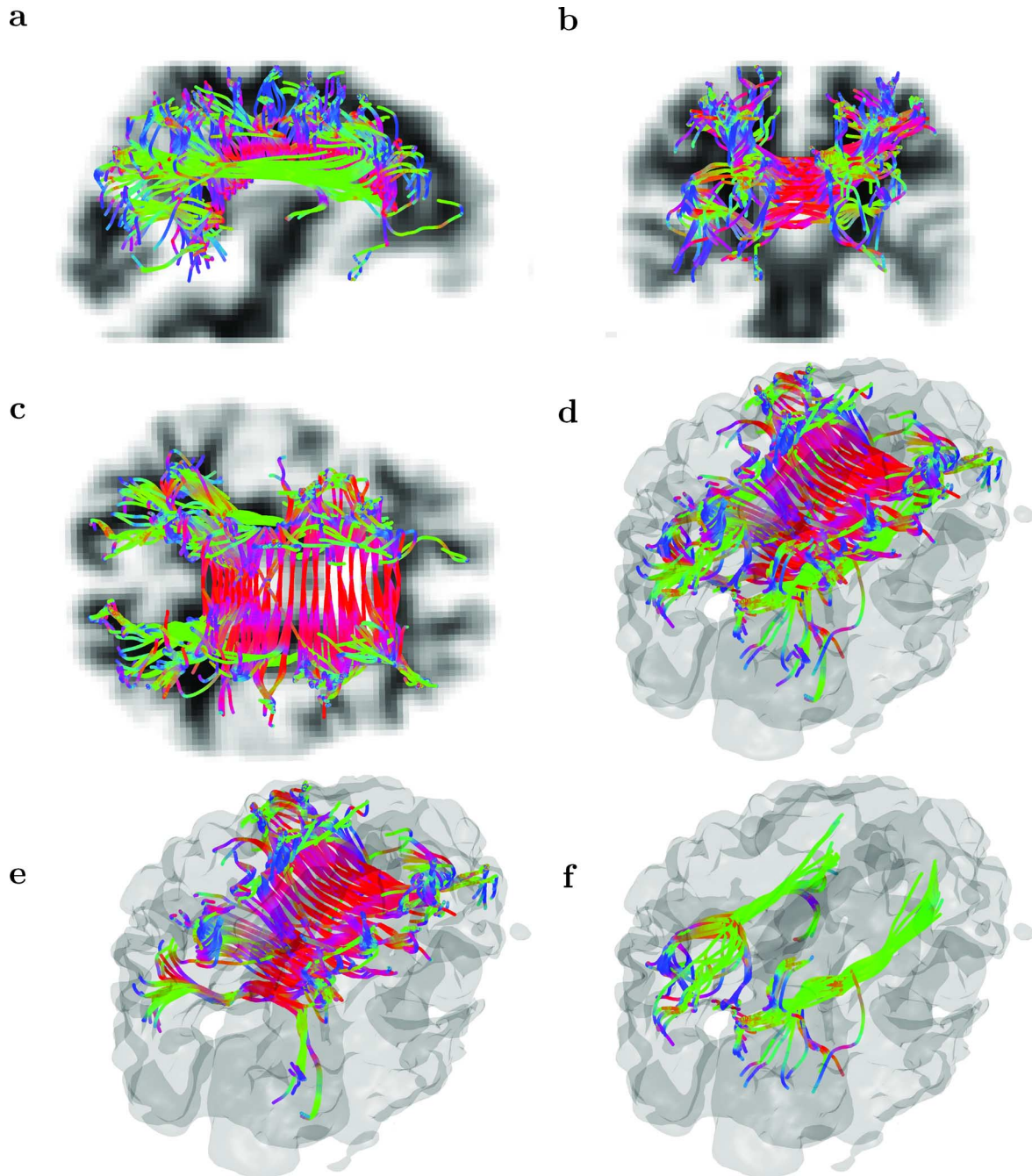

Fig. 7. Selected fiber tracts obtained by geometrical optics-like processing of ESP guided tractography, with ESP equilibrium probability distribution shown by grayish transparent background. A subset of seed points has been used to initialize fiber tracts, with 135 seeds located in corpus callosum (arranged in blocks of three consecutive voxels coronally and two or three voxels vertically), and two sets of 30 seeds localized in left and right hemispheres around the inferior fronto-occipital fasciculus areas (also blocked in by three consecutive voxels vertically and five consecutive voxels coronally), giving 195 seeds total. Panels (a)–(d) corresponds to different projections and panes (e) and (f) shows tracts with seeds in corpus callosum and fasciculus respectively. The corpus callosum originating fibers and fibers going through longitudinal fasciculus are crossing in areas occupying multiple voxels and the ESP tractography with geometrical optics-like tracking is able to correctly proceed through those voxels.

shown in Fig. 8, clearly demonstrate the ability of the algorithm to detect separate fibers crossing through rather small regions at different directions. In this figure, the fibers passing though the chosen location include those from latero–lateral (left to right and right to left), anterior–posterior, and dorsal–ventral. That example shows that the multiple scales used by ESP guided geometrical optics-like approach can help resolve crossing of multiple fibers in a small area of only several voxels extent.

To study behavior of our approach in even more difficult conditions, we selected a single seed voxel that is located into a region where the corticospinal tract crosses the corpus callosum. This is a region that is well known for crossing fiber problems, and appears often in the DTI literature. Even starting with just a single voxel seed in the “difficult” area, the multi-scale multi-modal approach is able to find and distinguish several fibers that go into different regions of brain (Fig. 9). For compar-

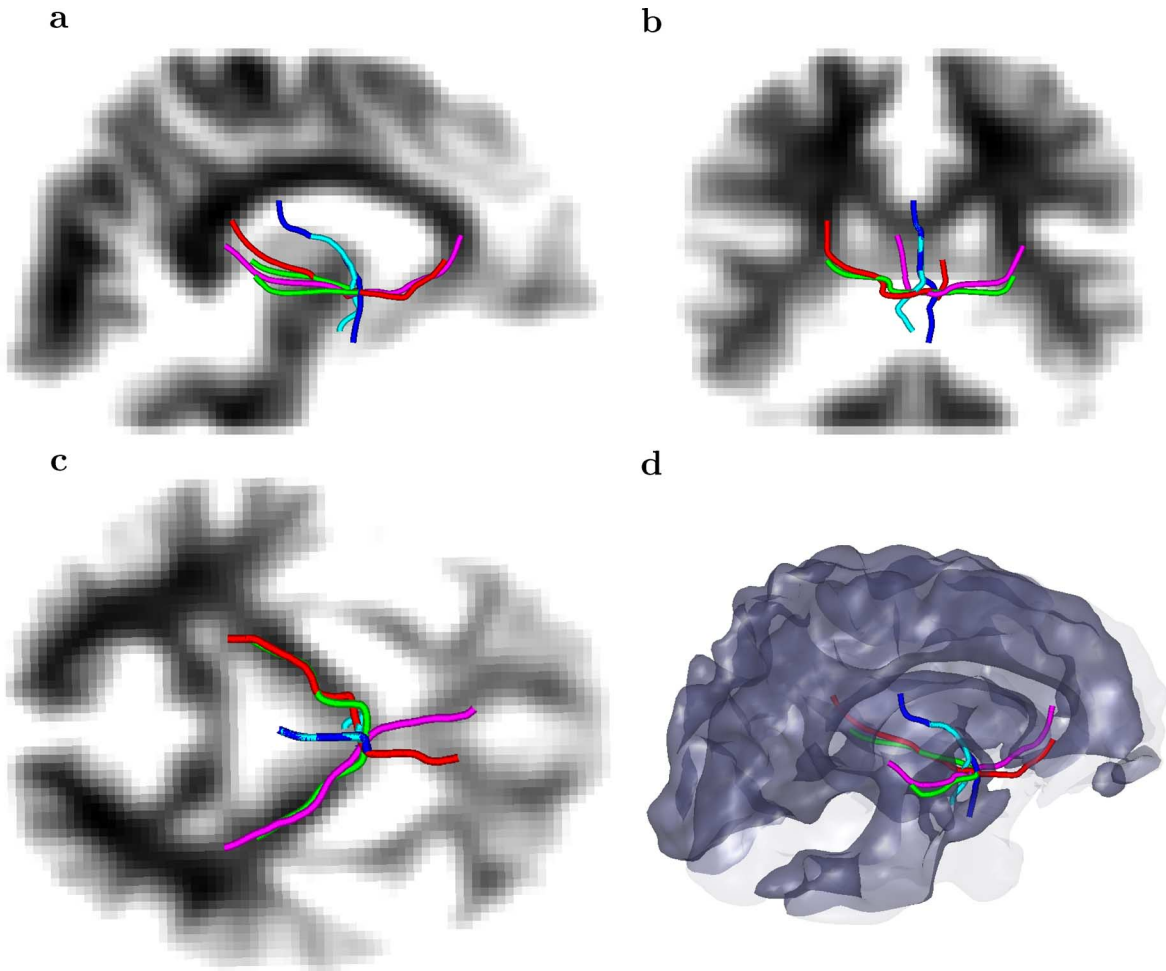

Fig. 8. Five fiber tracts going through a small region where a mixture of different fiber orientations emphasizes different directions at different scales. Five chosen tracts show coexistence of orientations ranging from a latero–lateral (left to right and right to left), an anterior–posterior, a dorsal–ventral direction, and a combinations of them.

ison we included in Fig. 9 a fiber tract that should be produced by a standard tractography approach with a single scale integration of the principal direction of the diffusion tensor (shown by white/gray tract).

Using several seeds in the small vicinity of the single seed voxel used in Fig. 9 gives more complicated behavior with a bundle of fibers going in and out of the corpus callosum, a bundle coming to/from the corticospinal tract, a bundle connecting anterior posterior regions, and some bundles going to outer regions of the brain (Fig. 10).

Finally, we applied our method to one of the diffusion imaging datasets (MGH 1010) available from the Human Connectome Project [54]. This dataset was collected on the customized Siemens 3T Connectom scanner, which is a modified 3T Skyra system (MAGNETOM Skyra Siemens Healthcare), housed at the MGH/HST Athinoula A. Martinos Center for Biomedical Imaging (see [55] for details of the scanner design and implementation). A 64-channel, tight-fitting brain array coil [56] was used for data acquisition. The dataset contains 96 slices of 140×140 matrix at four levels of diffusion sensitizations (b-values $\mathrm{b} = 1\ \mathrm{k}, 3\ \mathrm{k}, 5\ \mathrm{k},$ and $10\ \mathrm{k\,s/mm^2}$) distributed over 552 total q-vectors.

Fig. 11 shows both anisotropy maps for the primary directions of the transition probabilities and full brain tractography for the whole volume as well as for a selected set of slices. Fibers cut out through set of slices in panel (d) in general show good agreement with the primary direction of the transition probability [panel (c)], at the same time clearly indicating that many areas with seemingly single fiber direction may contain a mixture of different orientations.

V. Discussion and Conclusion

We have developed a novel diffusion estimation and fiber tractography method that is based on simultaneous estimation of global and local parameters of neural tracts from maximum entropy principles and sorting them into a series of entropy spectral pathways (ESP). The method uses local coupling between sub-scale diffusion parameters to compute the structure of the equilibrium probabilities that define the global information entropy field and uses this global entropy to update the local properties of neural fiber tracts.

We have also developed an efficient way to trace individual tracts that utilizes the multi-scale and multi-modal structure of

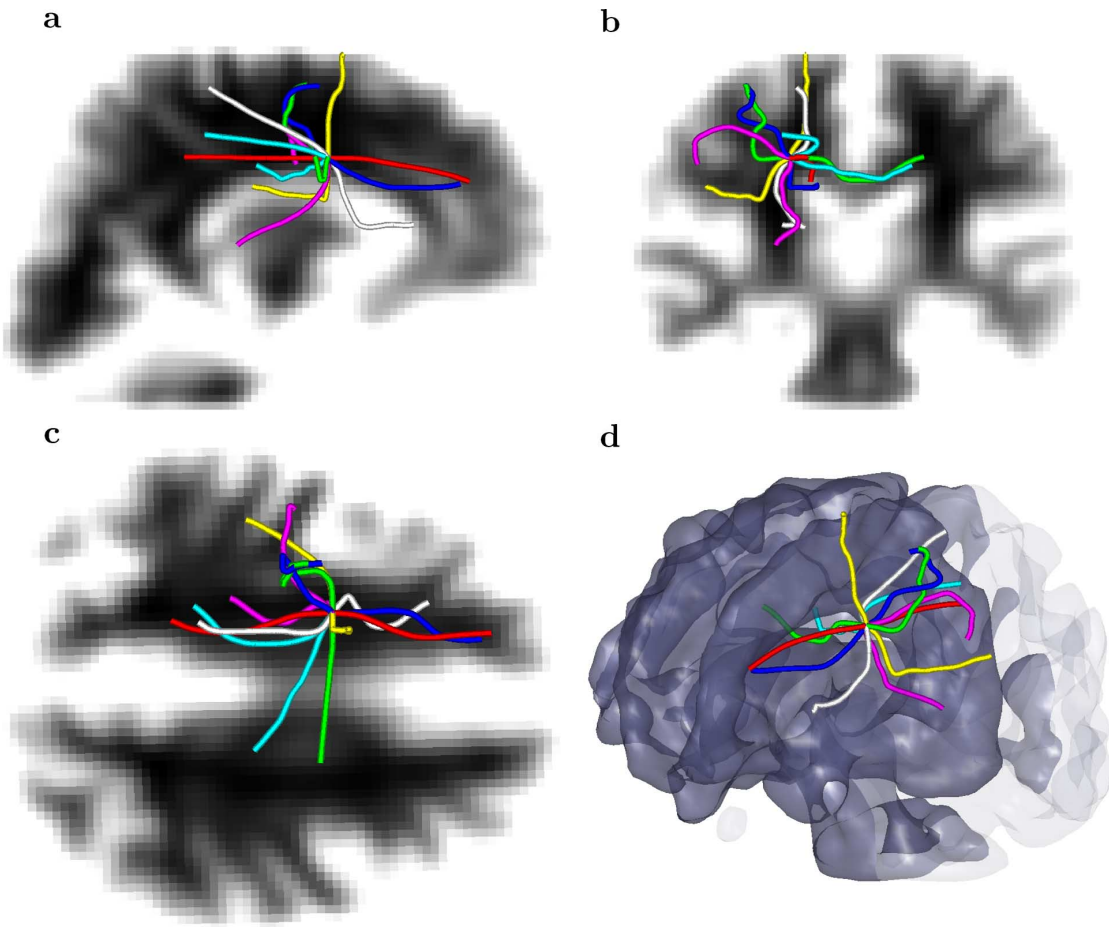

Fig. 9. Seven fiber tracts with the same single voxel seed located in the area where the corticospinal tract crosses the corpus callosum. Geometrical optics-like approach with different scales parameterized by the vector $\vec{k}$ produces a latero-lateral, an anterior–posterior, a dorsal–ventral (and mixed) tracts that originate in the same voxel. For comparison we included a fiber tact that would be traced by the standard method (shown by white/gray color) with the same voxel as a seed.

the local diffusion-convection propagation by means of an approach reminiscent of the geometrical optics ray tracing in dispersive media (either elastic or viscoelastic). This geometrical optics-like approach naturally includes multiple scales that allows fiber tracing to continue fibers through voxels with complex local diffusion properties where multiple fiber directions are unable to be adequately resolved.

One of the most important aspects of our method is that it is "global" in the sense that data from spatially extended brain structures are being used to inform both the local diffusion and generation of tracts. The typical workflow of majority of other global algorithms used in tractography, including algorithms based on the well known shortest-path algorithm on graphs by Dijkstra [57]–[60], represent the brain as a graph, where each voxel is a node, in which they have a local estimation of the diffusion process that they use as a speed function to guide a front evolution evolving from a seed point. Then, the geodesic or shortest-path between this point and any other location in the brain can be easily computed with backtracking. While there might appear to be similarities with our method, we point out that both the theoretical foundations and the numerical implementation for our approach are quite different from these schemes. It is important to realize that our method does not represent the brain as a graph. Rather, only nearest neighbor coupling has been used. However, as we have shown in [34], this is sufficient to produce long-range correlations. The local estimation of the diffusion process is not used as a speed function. Rather, the prior coupling is used to find the global eigenvalues/vectors, rank those information pathways based on a maximum entropy, and spread this global information about pathways to every voxel. This global information we use in every voxel is more than just an analog of a locally inferred speed function—it is more akin to a dispersion of fibers. For each voxel this function includes both angular and radial (scale) distributions obtained as a collective effect of all fibers that cross in a single voxel.

Our tracking process, although it might appear to look like a front evolution, is not. Our approach does not need an explicit front evolution step at all. The eigensystem calculation of the connectivity matrix provides more complete and accurate path information, than is available from the typical front evolution methods, and does more efficiently. Our tracking is performed in 6 dimensional coordinate and momentum space. Not only is the position of each fiber updated on each step, but a momentum equation is used to update the local fiber orientation as well as a rate of orientation change based on a globally constructed distribution/dispersion of fibers. All the current tracking algorithm, including the shortest-path algorithms [58]–[60], only update

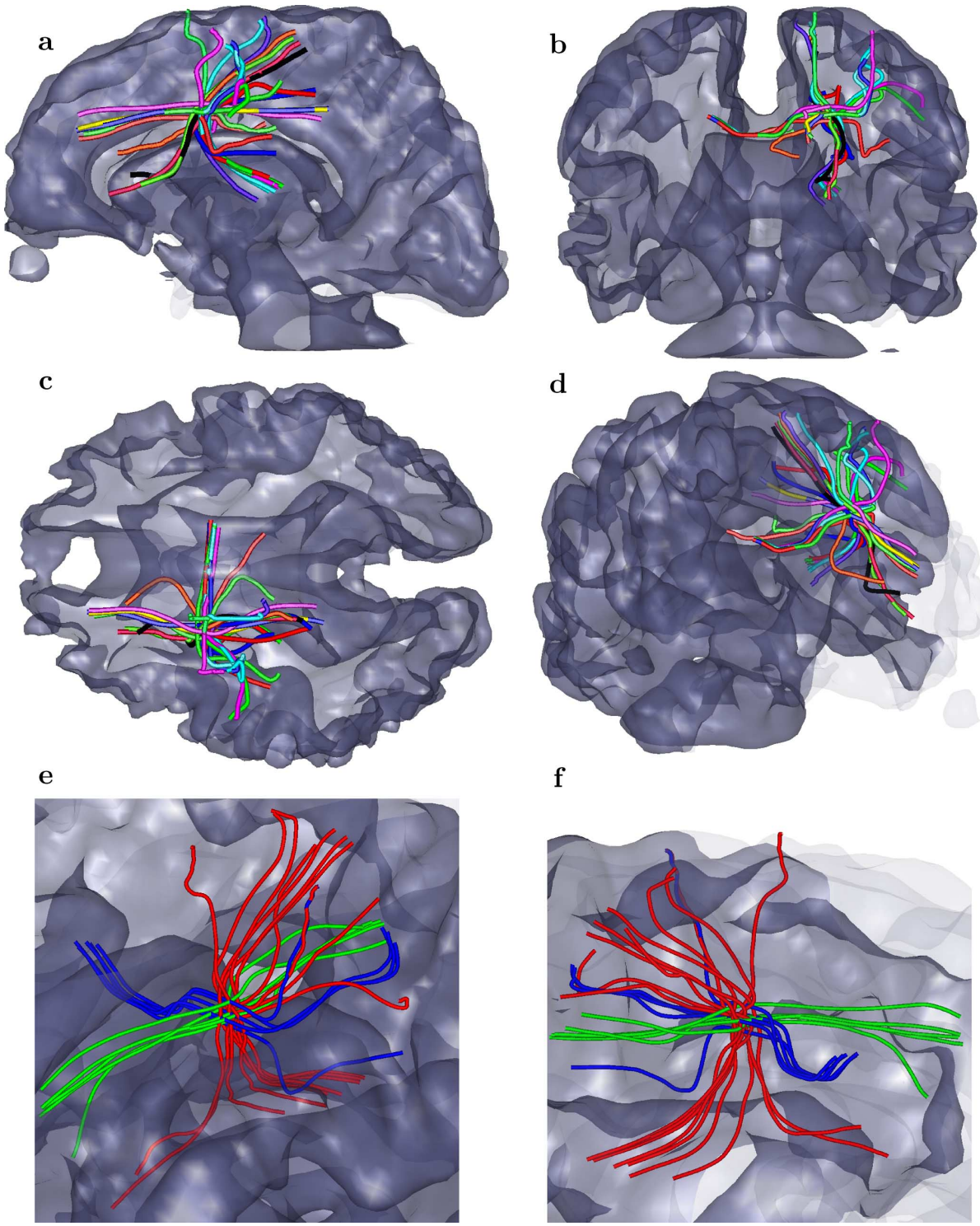

Fig. 10. Several bundles of fiber tracts obtained when seeded by several voxels in the vicinity of a single seed voxel used in the previous Fig. 9. Panels (a)–(c) show coronal, sagittal and axial planes, panel (d) shows 3-D view and panels (e) and (f) show different projections of zoomed area with fibers grouped in going from/to corpus callosum (blue), coming to/from the corticospinal tract (red), and longitudinal fibers (green).

a position of fiber assuming its orientation defined by static (fixed at each voxel position) speed function.

Although we presented here both the theoretical foundation and a number of practical examples that characterize performance and accuracy of our approach, the main limitation of the method and of the overall study is the lack of a system wide analysis of a role of different parameters that can be updated both during data acquisition and during reconstruction stages on the optimality of diffusion estimates as well as on the overall tractography results.

It is important to reiterate that we have formulated the analysis problem in this paper as one of inference where the goal is to make the most accurate estimates of both the local diffusion and the extended fiber tracts based only upon the available data and any relevant prior information. This is just the logic of probability theory [61] and the theoretical basis for our method

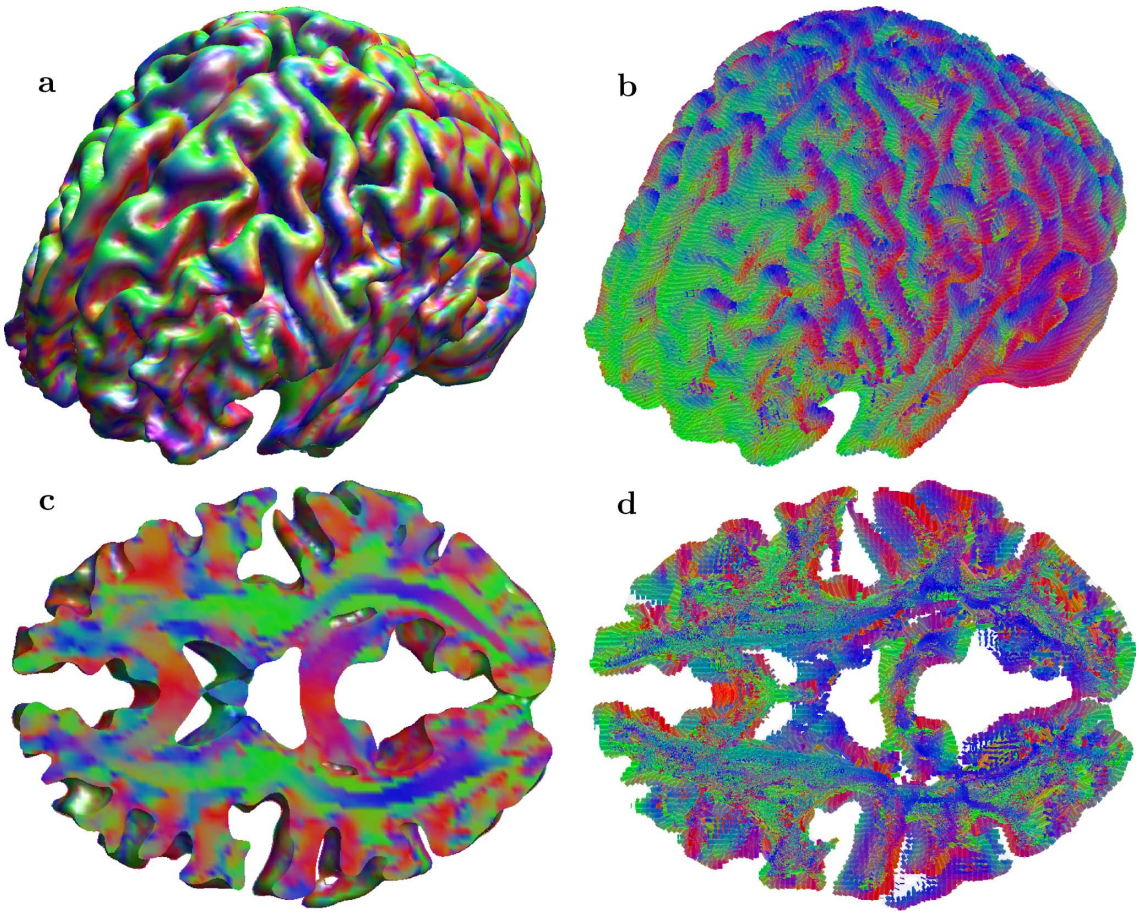

Fig. 11. Results of full brain ESP tractography obtained using the Human Connectome Project diffusion imaging dataset (MGH 1010). Panel (a) shows anisotropy maps for primary direction of the transition probabilities for the whole volume and panel (b) shows full brain tractography also for the whole volume. Panel (c) and (d) provide detailed views of for a set of slices ($46 \leq z \leq 51$).

is a probabilistic analysis of information flow in a lattice [34]. The key result of that paper is that *local coupling* information provides significant information about *global pathways*, which thus forms the important connection between local phenomena (diffusion) and global structures (fiber tracts). Moreover, the dynamics of how local effects inform global structures was shown to be characterized by a Fokker-Planck equation with a potential equal to the entropy [34], a formulation that had previously been put forth in a general theoretical framework [35], [36], but here finds a very practical manifestation, as it facilitates a geometric-optics tractography scheme where the relationship between the local diffusion measurements and the global fiber tract structure is made explicit.

Perhaps the most important result of our analysis is that the connection between the local and the global properties of the diffusion field are mediated by the transition probability, which emerges as a more fundamental quantity than the traditional diffusion PDF. In effect, our approach makes explicit a fact that is often implicitly assumed in diffusion analysis papers but rarely explicitly addressed: There is a fundamental logical flaw in estimating the local diffusion as if it were taking place in *independent*, isolated voxels, but then using it to generate connections between voxels based on their assumed *dependence*. Our formulation naturally incorporates the continuum of spatial scales, from local to global, and avoids unnecessary arguments related to the fact that the actual diffusion is occurring on a scale much smaller than the measurement process, and thus far smaller than the scale of the fiber structures, since we are only requiring that our macroscopic predictions from microscopic phenomena are consistent with our data and prior information.

This method has potential significance for a wide range of applications that employ diffusion weighted imaging, including studies of brain connectivity.

REFERENCES

[1] S. Mori, B. Crain, V. Chacko, and P. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, no. 2, pp. 265–269, Feb. 1999.

[2] T. Conturo *et al.*, "Tracking neuronal fiber pathways in the living human brain," *Proc. Natl. Acad. Sci. USA*, vol. 96, pp. 10 422–10 427, Aug. 1999.

[3] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "*In vivo* fiber tractography using (DT MRI) data," *Magn. Reson. Med.*, vol. 44, pp. 625–632, 2000.

[4] T. Behrens, H. Berg, S. Jbabdi, M. Rushworth, and M. Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?," *NeuroImage*, vol. 34, no. 1, pp. 144–155, 2007.

[5] D. Jones, "Tractography gone wild: Probabilistic fibre tracking using the wild bootstrap with diffusion tensor MRI," *IEEE Trans. Med. Imag.*, vol. 27, no. 9, pp. 1268–1274, Sep. 2008.

[6] M. Descoteaux, R. Deriche, T. Knosche, and A. Anwander, "Deterministic and probabilistic tractography based on complex fibre orientation distributions," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 269–286, Feb. 2009.

[7] N. Ratnarajah, A. Simmons, and A. Hojjatoleslami, "Novel approach for improved tractography and quantitative analysis of probabilistic fibre tracking curves," *Med. Image Anal. J.*, 2011.

[8] S. Jbabdi, M. W. Woolrich, J. L. R. Andersson, and T. E. J. Behrens, "A Bayesian framework for global tractography," *NeuroImage*, vol. 37, no. 1, pp. 116–129, Aug. 2007.

[9] I. Aganj *et al.*, "A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography," *Med. Image Anal.*, 2011.

[10] J. Mangin *et al.*, "A framework based on spin glass models for the inference of anatomical connectivity from diffusion-weighted MR data—A technical review," *NMR Biomed.*, vol. 15, no. 7-8, pp. 481–492, 2002.

[11] B. Kreher, I. Mader, and V. Kiselev, "Gibbs tracking: A novel approach for the reconstruction of neuronal pathways," *Magn. Reson. Med.*, vol. 60, no. 4, pp. 953–963, 2008.

[12] P. Fillard, C. Poupon, and J. Mangin, "A novel global tractography algorithm based on an adaptive spin glass model," in *Proc. MICCAI*, 2009, pp. 927–934.

[13] M. Reisert, I. Mader, C. Anastasopoulos, M. Weigel, S. Schnell, and V. Kiselev, "Global fiber reconstruction becomes practical," *NeuroImage*, vol. 54, no. 2, pp. 955–962, 2011.

[14] P. Basser and C. Pierpaoli, "Microstructual and physiological features of tissues elucidat ed by quantitative diffusion tensor MRI," *J. Magn. Reson. B.*, vol. 111, pp. 209–219, 1996.

[15] S. Farquharson *et al.*, "White matter fiber tractography: Why we need to move beyond DTI," *J. Neurosurg.*, vol. 118, no. 6, pp. 1367–1377, Jun. 2013.

[16] D. Tuch, T. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magn. Reson. Med.*, vol. 48, pp. 577–582, 2002.

[17] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, no. 1, pp. 259–267, Jan. 1994.

[18] D. S. Tuch, "Q-ball imaging," *Magn. Reson. Med.*, vol. 52, no. 6, pp. 1358–1372, Dec. 2004.

[19] L. Frank, "Characterization of anisotropy in high angular resolution diffusion weighted MRI," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1083–1099, Jun. 2002.

[20] E. Ozarslan and T. Mareci, "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 955–965, 2003.

[21] A. Anderson, "Measurement of fiber orientation distributions using high angular resolution diffusion imaging," *Magn. Reson. Med.*, vol. 54, no. 5, pp. 1194–1206, 2005.

[22] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, "Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications," *Magn. Reson. Med.*, vol. 56, no. 2, pp. 395–410, 2006.

[23] F. C. Yeh, V. J. Wedeen, and W. Y. Tseng, "Generalized q-sampling imaging," *IEEE Trans. Med. Imag.*, vol. 29, no. 9, pp. 1626–1635, Sep. 2010.

[24] V. J. Wedeen, P. Hagmann, W. Y. Tseng, T. G. Reese, and R. M. Weisskoff, "Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging," *Magn. Reson. Med.*, vol. 54, no. 6, pp. 1377–1386, Dec. 2005.

[25] M. Descoteaux, R. Deriche, D. L. Bihan, J. F. Mangin, and C. Poupon, "Multiple q-shell diffusion propagator imaging," *Med. Image Anal.*, vol. 15, no. 4, pp. 603–621, Aug. 2011.

[26] S. L. Merlet and R. Deriche, "Continuous diffusion signal, EAP and ODF estimation via Compressive Sensing in diffusion MRI," *Med. Image Anal.*, vol. 17, no. 5, pp. 556–572, Jul. 2013.

[27] A. P. Hosseinbor, M. K. Chung, Y. C. Wu, and A. L. Alexander, "Bessel Fourier Orientation Reconstruction (BFOR): An analytical diffusion propagator reconstruction for hybrid diffusion imaging and computation of q-space indices," *Neuroimage*, vol. 64, pp. 650–670, Jan. 2013.

[28] N. S. White, T. B. Leergaard, H. D'Arceuil, J. G. Bjaalie, and A. M. Dale, "Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation," *Hum. Brain Mapp.*, vol. 34, no. 2, pp. 327–346, Feb. 2013.

[29] D. Cory and A. Garroway, "Measurement of translational displacement probabilities by NMR: An indicator of compartmentation," *Magn. Res. Med.*, vol. 14, p. 435, 1990.

[30] P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford, U.K.: Oxford Univ. Press, 1993.

[31] G. B. Tefera, Y. Zhou, V. Juneja, and P. A. Narayana, "Evaluation of fiber tracking from subsampled q-space data in diffusion spectrum imaging," *Magn. Reson. Imag.*, vol. 31, no. 6, pp. 820–826, Jul. 2013.

[32] S. B. Vos, M. A. Viergever, and A. Leemans, "Multi-fiber tractography visualizations for diffusion MRI data," *PLoS One*, vol. 8, no. 11, p. e81453, 2013.

[33] S. Jbabdi, S. N. Sotiropoulos, A. M. Savio, M. Grana, and T. E. Behrens, "Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems," *Magn. Reson. Med.*, vol. 68, no. 6, pp. 1846–1855, Dec. 2012.

[34] L. R. Frank and V. L. Galinsky, "Information pathways in a disordered lattice," *Phys. Rev. E*, vol. 89, p. 032142, Mar. 2014.

[35] H. Grabert and M. Green, "Fluctuations and nonlinear irreversible processes," *Phys. Rev. A*, vol. 19, no. 4, pp. 1747–1756, Apr. 1979.

[36] E. Jaynes, "Macroscopic prediction," in *Complex Systems – Operational Approaches in Neurobioloy, Physics, and Computers*, H. Haken, Ed. Berlin: Springer-Verlag, 1985, pp. 254–269.

[37] V. L. Galinsky and L. R. Frank, "Automated segmentation and shape characterization of volumetric data," *NeuroImage*, vol. 92, no. 0, pp. 156–168, 2014.

[38] N. Lebedev, *Special Functions and their Applications*. Mineola, NY: Dover, 1972.

[39] Y. A. Kravtsov and Y. I. Orlov, *Geometrical Optics of Inhomogeneous Media*. Berlin, Germany: Springer-Verlag, 1990.

[40] E. Ozarslan, T. M. Shepherd, C. G. Koay, S. J. Blackband, and P. J. Basser, "Temporal scaling characteristics of diffusion as a new MRI contrast: Findings in rat hippocampus," *Neuroimage*, vol. 60, no. 2, pp. 1380–1393, Apr. 2012.

[41] E. Ozarslan, T. M. Shepherd, B. C. Vemuri, S. J. Blackband, and T. H. Mareci, "Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT)," *Neuroimage*, vol. 31, no. 3, pp. 1086–1103, Jul. 2006.

[42] Z. Burda, J. Duda, J. Luck, and B. Waclaw, "Localization of the maximal entropy random walk," *Phys. Rev. Lett.*, vol. 102, no. 16, p. 160602, 2009.

[43] Z. Burda, J. Duda, J. Luck, and B. Waclaw, "The various facets of random walk entropy," *Acta Physica Polonica B*, pp. 949–987, 2010.

[44] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[45] B. McMillan, "The basic theorems of information theory," *Ann. Math. Stat.*, vol. 24, no. 2, pp. 196–219, 1953.

[46] P. D. Lax, "Asymptotic solutions of oscillatory initial value problems," *Duke Math. J.*, vol. 24, no. 4, pp. 627–646, Dec. 1957.

[47] P. Fillard *et al.*, "Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom," *Neuroimage*, vol. 56, no. 1, pp. 220–234, May 2011.

[48] M.-A. Côté *et al.*, "Tractometer: Towards validation of tractography pipelines," *Med. Image Anal.*, vol. 17, no. 7, pp. 844–857, 2013.

[49] K. Setsompop *et al.*, "Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty," *Magn. Res. Med.*, vol. 67, no. 5, pp. 1210–1224, Aug. 2011.

[50] M. A. Griswold *et al.*, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Res. Med.*, vol. 47, no. 6, pp. 1202–1210, Jun. 2002.

[51] J. L. R. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging," *NeuroImage*, vol. 20, no. 2, pp. 870–888, Oct. 2003.

[52] S. M. Smith *et al.*, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, pp. S208–S219, Jan. 2004.

[53] M. Catani and M. Thiebaut de Schotten, "A diffusion tensor imaging tractography atlas for virtual in vivo dissections," *Cortex*, vol. 44, no. 8, pp. 1105–1132, Sep. 2008.

[54] S. N. Sotiropoulos *et al.*, "Effects of image reconstruction on fiber orientation mapping from multichannel diffusion MRI: Reducing the noise floor using sense," *Magn. Reson. Med.*, vol. 70, no. 6, pp. 1682–1689, 2013.

[55] K. Setsompop *et al.*, "Pushing the limits of in vivo diffusion {MRI} for the human connectome project," *NeuroImage*, vol. 80, no. 0, pp. 220–233, 2013.

[56] B. Keil *et al.*, "A 64-channel 3t array coil for accelerated brain MRI," *Magn. Reson. Med.*, vol. 70, no. 1, pp. 248–258, 2013.

[57] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Math.*, vol. 1, no. 1, pp. 269–271, 1959.

[58] G. J. Parker, C. A. Wheeler-Kingshott, and G. J. Barker, "Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 505–512, May 2002.

[59] S. Jbabdi *et al.*, "Accurate anisotropic fast marching for diffusion-based geodesic tractography," *Int. J. Biomed. Imag.*, vol. 2008, p. 320195, 2008.

[60] A. Zalesky, "DT-MRI fiber tracking: A shortest paths approach," *IEEE Trans. Med. Imag.*, vol. 27, no. 10, pp. 1458–1471, Oct. 2008.

[61] E. Jaynes, *Probability Theory: The Logic of Science*. New York: Cambridge Univ. Press, 2003.