

# A Random Riemannian Metric for Probabilistic Shortest-Path Tractography

Søren Hauberg<sup>1</sup>, Michael Schober<sup>2</sup>, Matthew Liptrot<sup>1,3</sup>, Philipp Hennig<sup>2</sup>, Aasa Feragen<sup>3</sup>,

<sup>1</sup> Cognitive Systems, Technical University of Denmark, Denmark

<sup>2</sup> Max Planck Institute for Intelligent Systems, Tübingen, Germany

<sup>3</sup> Department of Computer Science, University of Copenhagen, Denmark

**Abstract.** Shortest-path tractography (SPT) algorithms solve global optimization problems defined from local distance functions. As diffusion MRI data is inherently noisy, so are the voxelwise tensors from which local distances are derived. We extend Riemannian SPT by modeling the stochasticity of the diffusion tensor as a “random Riemannian metric”, where a geodesic is a distribution over tracts. We approximate this distribution with a Gaussian process and present a probabilistic numerics algorithm for computing the geodesic distribution. We demonstrate SPT improvements on data from the Human Connectome Project.

## 1 Introduction

Diffusion weighted imaging enables inference of structural brain connectivity, assuming that water diffuses *along* brain fibers, not across. At any brain location, a local brain fiber orientation distribution function (ODF) is estimated from measured diffusion along a fixed set of directions. Integration of the field of local ODFs yields a heuristic estimate of the most likely path for a brain fiber.

We study the stochasticity of the ODF estimate and its effect on the distribution of estimated tracts through Riemannian shortest-path tractography (SPT) in diffusion tensor imaging (DTI), where the ODF is a second order tensor describing a Gaussian diffusion process at any given location. In Riemannian SPT, tracts are estimated as geodesics on a Riemannian manifold, where the Riemannian metric is constructed from the diffusion tensor [9, 14, 15]. As the diffusion tensor is estimated from data, it is subject to noise. The tensor, and the induced Riemannian metric, should therefore be treated as stochastic variables, not exact representatives of the true underlying diffusion process. This is not possible in current Riemannian tractography models, which assume a deterministic metric.

**We present** a solution to *probabilistic SPT* for DTI which treats diffusion tensors as stochastic variables, and returns a distribution over the tracts connecting two given points in the brain, approximated by a Gaussian process (GP). Algorithmically, we extend recent probabilistic solutions for ordinary differential equations (ODEs) [12, 18] by allowing “noisy” evaluations of the ODE due to the noisy metric. The resulting ODE solver is extendable to other problems. The combined model yields both quantitative and qualitative improvements over standard SPT algorithms.

## 2 Shortest-Path Tractography and Random Geometries

In *shortest-path tractography* (SPT) the ODFs induce a local Riemannian metric [9, 11, 14, 15], where the most probable fiber connecting two points  $\mathbf{x}$  and  $\mathbf{y}$  is re-interpreted as the *geodesic*, or shortest path, connecting  $\mathbf{x}$  and  $\mathbf{y}$ .

**Why SPT?** Walk-based streamline methods searching for the most likely paths from a seed point to anywhere in the brain [3, 16] suffer from path-length dependency, i.e. long paths are harder to estimate than short ones. When two endpoints are known, SPT avoids this by looking for the shortest connecting path. This ensures that long and short tracts are estimated with the same sample size. As endpoints are not always known, these are complementary methods.

**Solving Riemannian SPT.** In Riemannian SPT, the diffusion tensor  $\mathbf{D}_{\mathbf{p}}$  at  $\mathbf{p}$  is commonly interpreted as the *inverse* of a Riemannian metric  $\mathbf{M}_{\mathbf{p}}$  [14, 15]. Geodesics returned by this *ad hoc* interpretation are typically “too straight” or are attracted to high diffusivity cortical spinal fluid (CSF) regions. Hao et al. [11] suggest a per-voxel scaling  $\mathbf{M}_{\mathbf{p}} = \alpha_{\mathbf{p}} \mathbf{D}_{\mathbf{p}}^{-1}$  that avoids this issue, while Fuster et al. [9] show that the *adjugate metric*  $\mathbf{M}_{\mathbf{p}} = \det(\mathbf{D}_{\mathbf{p}}) \mathbf{D}_{\mathbf{p}}^{-1}$  gives a metric under which free Brownian motion matches the diffusion implied by  $\mathbf{D}_{\mathbf{p}}$ . This reduces attraction to CSF regions [9, 19], so we use the adjugate metric.

Given a Riemannian metric  $\mathbf{M}_{\mathbf{p}}$ , the Riemannian geodesic  $\mathbf{c}: [0, 1] \rightarrow \mathbb{R}^3$  connecting two points  $\mathbf{x}$  and  $\mathbf{y}$  is found by solving the geodesic ODE

$$\ddot{c}_d(t) = f_d(t, \mathbf{c}, \dot{\mathbf{c}}) = -\Gamma_d^{\top} \cdot (\dot{\mathbf{c}}(t) \otimes \dot{\mathbf{c}}(t)), \quad d = 1, \dots, D = 3, \quad (1)$$

$$\Gamma_d = \frac{1}{2} \sum_{k=1}^D \left[ \mathbf{M}_{\mathbf{c}(t)}^{-1} \right]_{d,k} \left( \frac{\partial \text{vec } \mathbf{M}_{\mathbf{p}}}{\partial p_k} \right)_{\mathbf{p}=\mathbf{c}(t)} \in \mathbb{R}^{D^2 \times 1}, \quad (2)$$

subject to the boundary conditions  $\mathbf{c}(0) = \mathbf{x}$  and  $\mathbf{c}(1) = \mathbf{y}$ . Here  $\otimes$  is the Kronecker product,  $\mathbf{c}(t) = [c_1(t); c_2(t); c_3(t)]$  is a point on the shortest path  $\mathbf{c}$ , and  $\dot{\mathbf{c}}$  and  $\ddot{\mathbf{c}}$  denote the first and second derivative along the path, respectively.

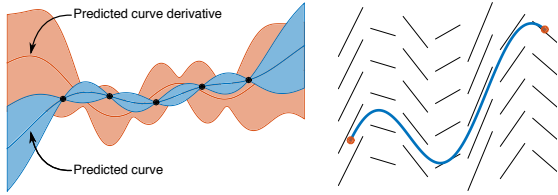
**Estimating a Random Geometry.** The Riemannian adjugate metric  $\mathbf{M}_{\mathbf{p}}$  is estimated from finite noisy data and should be considered a stochastic variable. This, however, complicates the geometric interpretation as the resulting object is now a *random Riemannian metric* [13], and not a deterministic metric. Given two endpoints  $\mathbf{x}$  and  $\mathbf{y}$ , our interest is in finding a connecting geodesic. Since the metric is stochastic, there is a distribution of geodesics connecting  $\mathbf{x}$  and  $\mathbf{y}$ .

We approximately solve the geodesic ODE (1) using a probabilistic description of the uncertainty over  $f(t, \mathbf{c}, \dot{\mathbf{c}})$ , the stochastic variable arising from Eq. 1 with a stochastic metric. We approximate  $f$  with a Gaussian distribution (using the shorthand notation  $f_{t, \mathbf{c}, \dot{\mathbf{c}}} \equiv f(t, \mathbf{c}, \dot{\mathbf{c}})$ )

$$p(f_{t, \mathbf{c}, \dot{\mathbf{c}}}) = \mathcal{N}(f_{t, \mathbf{c}, \dot{\mathbf{c}}}; m_{t, \mathbf{c}, \dot{\mathbf{c}}}, \mathbf{C}_{t, \mathbf{c}, \dot{\mathbf{c}}}). \quad (3)$$

The mean and covariance that form this approximation are computed from

$$m_d = -\mathbb{E}(\Gamma_d)^{\top} (\dot{\mathbf{c}} \otimes \dot{\mathbf{c}}) \quad \text{and} \quad C_d^2 = (\dot{\mathbf{c}} \otimes \dot{\mathbf{c}})^{\top} \text{cov}(\Gamma_d) (\dot{\mathbf{c}} \otimes \dot{\mathbf{c}}), \quad (4)$$



**Fig. 1.** *Left:* A GP regressor (blue) and its derivative (orange). *Right:* A numerical solver seeks the smooth function (blue) that best fits observed derivatives (black) while meeting the boundary constraints (red).

where  $\mathbf{m}_{t,c,\dot{c}} = [m_1; m_2; m_3]$  and  $\mathbf{C}_{t,c,\dot{c}} = \text{diag}(C_1^2, C_2^2, C_3^2)$ .

We estimate the mean and covariance of  $\mathbf{\Gamma}_d$  empirically, subsampling the gradient directions used to generate the local diffusion tensors into  $S = 20$  batches each containing 80% of the directions. For each subsample we fit a diffusion tensor field and compute  $\mathbf{\Gamma}_d$  (2). Finally, we estimate the moments  $\mathbb{E}(\mathbf{\Gamma}_d) \approx \frac{1}{S} \sum_{s=1}^S \mathbf{\Gamma}_d^{(s)}$  and  $\text{cov}(\mathbf{\Gamma}_d) \approx \frac{1}{S-1} \sum_{s=1}^S \left( \mathbf{\Gamma}_d^{(s)} - \mathbb{E}(\mathbf{\Gamma}_d) \right) \left( \mathbf{\Gamma}_d^{(s)} - \mathbb{E}(\mathbf{\Gamma}_d) \right)^\top$ .

### 3 Regressing an ODE

In the Riemannian setting, geodesics are found as the solution to the geodesic ODE (1). This is a smooth curve  $\mathbf{c}(t) : [0, 1] \rightarrow \mathbb{R}^3$ , which must be estimated numerically. We use *probabilistic numerics* that solves the ODE using Gaussian process (GP) regression [12, 18, 20]. We extend previous work [19] to handle uncertainty in the ODE due to a noisy metric.

**GP Regression.** A GP  $c(t) \sim \mathcal{GP}(\mu(t), k(t, u))$  [17] is a probability measure over real-valued functions  $c : \mathbb{R} \rightarrow \mathbb{R}$  such that any finite restriction to function values  $\{c(t_n)\}_{n=1}^N$  has a Gaussian distribution. GPs are parameterized by a mean function  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  and a covariance function  $k : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  that determines the regularity of the paths. GPs are closed under linear transformations

$$p(c) = \mathcal{GP}(c; \mu, k) \quad \Rightarrow \quad p(\mathbf{A}c) = \mathcal{GP}(\mathbf{A}c; \mathbf{A}\mu, \mathbf{A}k\mathbf{A}^\top), \quad (5)$$

where  $\mathbf{A}^\top$  denotes application of operator  $\mathbf{A}$  to the left. Given observations  $\{\mathbf{t}, \mathbf{Y}\} = \{(t_1, y_1), \dots, (t_N, y_N)\}$  of likelihood  $p(y_i | t_i) = \mathcal{N}(y_i; \mathbf{A}c(t_i), \sigma^2 \mathbf{I})$ , the posterior over  $c$  is a Gaussian process  $\mathcal{GP}(\tilde{\mu}, \tilde{k})$  with

$$\begin{aligned} \tilde{\mu}(t) &= \mu(t) + k(t, \mathbf{t})\mathbf{A}^\top(\mathbf{A}k_{\mathbf{t}\mathbf{t}}\mathbf{A}^\top + \sigma^2\mathbf{I})^{-1}(\mathbf{Y} - \mathbf{A}\mu(\mathbf{t})) \\ \tilde{k}(t, u) &= k(t, u) - k(t, \mathbf{t})\mathbf{A}^\top(\mathbf{A}k_{\mathbf{t}\mathbf{t}}\mathbf{A}^\top + \sigma^2\mathbf{I})^{-1}\mathbf{A}k(\mathbf{t}, u), \end{aligned} \quad (6)$$

where  $(k_{\mathbf{t}\mathbf{t}})_{ij} = k(t_i, t_j)$  is the  $N \times N$  covariance of input locations [17, §2.2], and similarly for  $k(t, \mathbf{t})$ . Differentiation is a linear operation so (by Eq. 5) a GP belief over  $c$  implies a GP belief over  $\partial c = \dot{c}$  as well (see left panel of Fig. 1).

Beliefs over multi-output functions  $\mathbf{c}(t) = [c_1(t); c_2(t); c_3(t)]$  can be constructed through vectorization. If the covariance structure is assumed to factorize between inputs and outputs,  $\text{cov}(c_i(t), c_j(u)) = [\mathbf{V}]_{ij} k(t, u)$ , for covariance  $\mathbf{V}$ , then the belief over  $\mathbf{c}$  can be written as  $p(\mathbf{c}) = \mathcal{GP}(\mathbf{c}; \boldsymbol{\mu}_c, \mathbf{V} \otimes k)$ .

**GP ODE solvers.** Numerical ODE solvers evaluate the ODE at finitely many points and estimate a smooth curve that *fits these observations* along with either initial or boundary values (see right panel of Fig. 1). *This is statistical regression.* This view gives rise to probabilistic ODE solvers [5, 12, 18–20] implemented using GP regression. In these solvers, uncertainty represents the approximation of not evaluating the ODE at the true solution, but at the current approximation.

At runtime, the solver repeatedly uses the current posterior mean estimate  $\tilde{\mathbf{c}}(t_i)$  for the true solution  $\mathbf{c}(t_i)$  to construct approximate noisy “observations”

$$\mathbf{y}_i = f(t_i, \tilde{\mathbf{c}}(t_i), \partial\tilde{\mathbf{c}}(t_i)) = \ddot{\mathbf{c}}(t_i) + \eta_i. \quad (7)$$

This describes that  $\tilde{\mathbf{c}}(t_i)$  is only an approximation to the true solution  $\mathbf{c}(t_i)$  as the observation  $\mathbf{y}_i$  is corrupted by an error  $\eta_i$ . Here,  $\eta_i$  is assumed to be Gaussian, and the current uncertainty over  $\mathbf{c}$  is propagated through the algorithm:

At step  $i$ , assume a posterior  $p(\mathbf{c}) = \mathcal{GP}(\mathbf{c}; \tilde{\boldsymbol{\mu}}^i, \tilde{k}^i)$ . We construct the estimates  $\tilde{\mathbf{c}}^i$  and  $\partial\tilde{\mathbf{c}}^i$  as the current “best guess”, the mean  $[\tilde{\mathbf{c}}^i; \partial\tilde{\mathbf{c}}^i] = [\tilde{\boldsymbol{\mu}}^i(t_i); \partial\tilde{\boldsymbol{\mu}}^i(t_i)]$ . An estimate for the error of this approximation is provided by the local variance

$$\text{cov} \begin{pmatrix} \tilde{\mathbf{c}}^i(t_i) \\ \partial\tilde{\mathbf{c}}^i(t_i) \end{pmatrix} = \begin{pmatrix} \tilde{k}^i(t_i, t_i) & \left. \frac{\partial\tilde{k}^i(t_i, t)}{\partial t} \right|_{t=t_i} \\ \left. \frac{\partial\tilde{k}^i(t, t_i)}{\partial t} \right|_{t=t_i} & \left. \frac{\partial^2\tilde{k}^i(t, u)}{\partial t\partial u} \right|_{t, u=t_i} \end{pmatrix} =: \boldsymbol{\Sigma}^i. \quad (8)$$

Assuming we have upper bounds  $\mathbf{U} > \partial f / \partial \mathbf{c}$  and  $\dot{\mathbf{U}} > \partial f / \partial \dot{\mathbf{c}}$  on the gradients of  $f$ ,  $\boldsymbol{\Sigma}^i$  can be used to estimate the error on  $y_i$  as  $(\mathbf{U}, \dot{\mathbf{U}})^\top \boldsymbol{\Sigma}^i (\mathbf{U}, \dot{\mathbf{U}}) =: \boldsymbol{\Lambda}^i$ , i.e.  $\eta_i \sim \mathcal{N}(0, \boldsymbol{\Lambda}^i)$ . This gives an observation likelihood function

$$p(\mathbf{y}_i \mid \ddot{\mathbf{c}}(t_i)) = \mathcal{N}(\mathbf{y}_i; \ddot{\mathbf{c}}(t_i), \boldsymbol{\Lambda}^i). \quad (9)$$

The belief is updated with Eq. 6 to obtain  $\tilde{\boldsymbol{\mu}}^{i+1}, \tilde{k}^{i+1}$ , and the process repeats.

The repeated extrapolation to construct  $\tilde{\mathbf{c}}(t_i)$  in the GP solver is similar to explicit Runge-Kutta methods as it defines a Butcher tableau. For some GP priors, the posterior mean even coincide with results from such methods [18].

**Regressing a Noisy ODE.** When the metric is *uncertain*, the geodesic ODE can only be evaluated probabilistically as  $p(f)$ . To handle this situation, we extend the probabilistic ODE solver to cope with Gaussian uncertain observations.

Due to noise, the curvature  $\mathbf{y}_i = f(t_i, \tilde{\mathbf{c}}_i, \partial\tilde{\mathbf{c}}(t_i))$  can only be estimated up to Gaussian noise  $p(\mathbf{y}_i \mid \mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}, \mathbf{C}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}) = \mathcal{N}(\mathbf{y}_i; \mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}, \mathbf{C}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}})$ . As the normal distribution is symmetric around the mean, this is a likelihood for  $y_i$ ,

$$p(\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}} \mid \mathbf{y}_i, \mathbf{C}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}) = \mathcal{N}(\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}; \mathbf{y}_i, \mathbf{C}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}). \quad (10)$$

Since both likelihoods (9, 10) are Gaussian, the latent  $\mathbf{y}_i$  can be marginalized analytically, giving the complete observation likelihood for  $\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}$

$$p(\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}} \mid \ddot{\mathbf{c}}(t_i)) = \mathcal{N}(\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}; \ddot{\mathbf{c}}_i, \mathbf{C}_{t_i, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}} + \boldsymbol{\Lambda}_i). \quad (11)$$

Solutions to noisy ODEs are then inferred by replacing Eq. 9 with Eq. 11, using  $\mathbf{m}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}} = \mathbb{E}[f_{t, \mathbf{c}, \dot{\mathbf{c}}}]$  in place of the (inaccessible) function evaluations  $\mathbf{y}_i = f_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}$ , such that the approximation error is modeled by the additive uncertainty  $\mathbf{C}_{t, \tilde{\mathbf{c}}, \partial\tilde{\mathbf{c}}}$ .

The ODE solver rely on two noise terms:  $\eta_i$  in Eq. 7 captures the numerical error, while  $\mathbf{C}$  in Eq. 10 describes the uncertainty arising from the data. While the terms are structurally similar they capture different error sources; e.g. changes in the number of grid points imply a change in  $\eta$ , while  $\mathbf{C}$  is unaffected. Both noise terms are approximated as Gaussian to ensure efficient inference.

**Adaptation to Tractography.** To compute geodesics in DTI we subsample the diffusion gradients, estimate the mean and covariance of  $\mathbf{\Gamma}_d$ , and finally solve the noisy geodesic ODE (1) numerically. For the prior covariance we use a Gaussian kernel  $k(t_i, t_j) = \exp[-(t_i - t_j)^2 / (2\lambda^2)]$  with length scale  $\lambda$ . To estimate  $\mathbf{V}$  and  $\lambda$  we maximize their likelihood with gradient descent on the positive definite cone. We initialize the prior mean with the Dijkstra path under the deterministic metric, which we pre-process with a GP smoother.

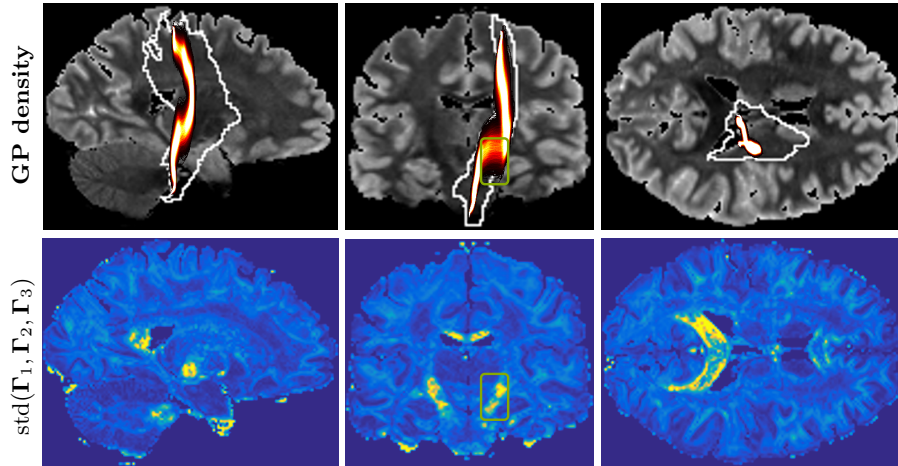
## 4 Experimental Results

Tractography was performed on pre-processed diffusion data of 20 subjects from the Q3 release of the Human Connectome Project (HCP) [7, 8, 10, 21]. The pre-processed HCP diffusion data contains 270 diffusion directions distributed equally over 3 shells with  $b$ -values = 1000, 2000 and 3000 s/mm<sup>2</sup> [21]. Diffusion directions were uniformly subsampled into 20 batches each containing 80% of the directions, where DTI tensors were computed with `dtifit` [2]. Segmentation was performed with `FAST` [23]. The cortico-spinal tract (CST) and inferior longitudinal fasciculus (ILF) used for experiments were obtained from the probabilistic expert-annotated Catani tract atlas [4]. ROI atlases were constructed in MNI152 “template space” by overlapping the tract atlas with regions from the Harvard-Oxford atlas [6]. The CST ROIs are the overlap with the brainstem, the hippocampus and the amygdala for one region, and the overlap with the superior frontal gyrus, the precentral gyrus and the postcentral gyrus for the second region. The ILF ROIs are the overlap with the temporal pole for one region, and the overlap with the superior occipital cortex, the inferior occipital cortex and the occipital pole for the second region. The constructed ROIs were warped from “template space” to “subject space” using warps provided by HCP.

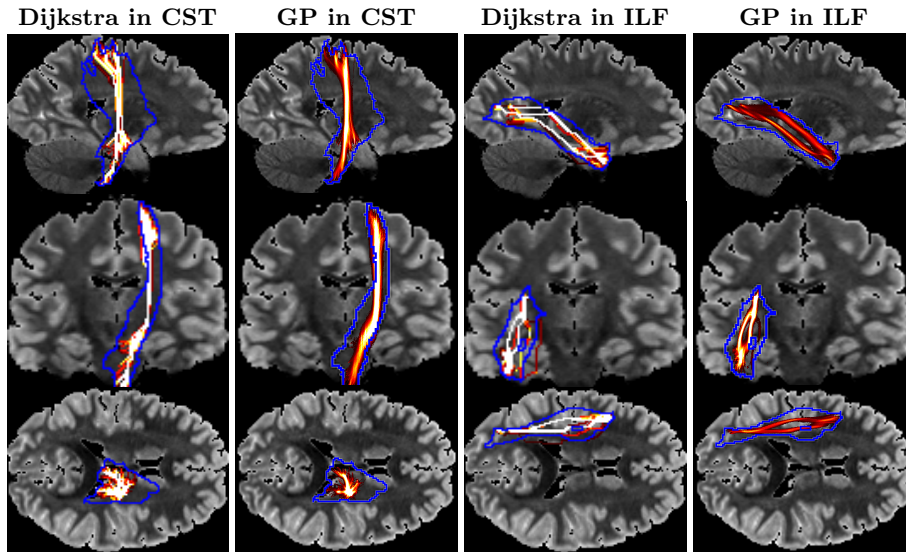
As a first illustration, Fig. 2 shows the density of a *single* geodesic within the CST projected onto a slice. The center column shows that the geodesic density roughly consists of two certain vertical line segments with an uncertain connection between them (green box). The bottom row shows the standard deviation of  $\mathbf{\Gamma}_d$ . The geodesic uncertainty appears related to data noise. This is not attained with other GP ODE solvers [19] as they model constant observation noise.

Next we sample 250 endpoint pairs and compute the corresponding geodesics. Fig. 3 shows the resulting heatmaps. The GP solution provides a more coherent picture of the tract compared to the picture generated with Dijkstra’s algorithm.

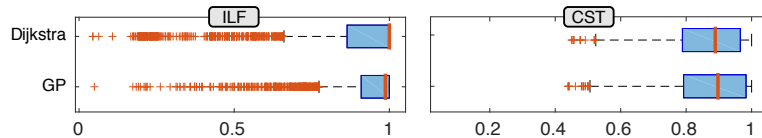
To compare solution qualities we compute the set of voxels each geodesic passes through. Taking the Catani atlas as a reference we measure the percentage of voxels which are classified as part of the tract by at least one expert. Figure 4 shows the results for 20 subjects. In ILF the median accuracy is comparable



**Fig. 2.** *Top row:* The density of a *single* GP geodesic under the random metric. The density heatmap is projected into axis-aligned slices; the background image is the expected metric trace; and the outline is where at least one expert annotated the CST. *Bottom row:* Standard deviation of  $\Gamma_d$  at the same slices as the top row.



**Fig. 3.** Example shortest paths in the CST and ILF using both Dijkstra’s algorithm on a deterministic metric, and a GP solver with a random Riemannian metric.



**Fig. 4.** Agreement with at least one expert in the Catani atlas as estimated by Dijkstras algorithm and the GP solver.

for Dijkstra and GP paths, but the GP error bars are significantly smaller. For the CST, we observe similar quality results for Dijkstra and GP paths. This is expected, as the ILF is generally considered the harder tract to estimate.

## 5 Discussion and Conclusion

SPT methods are advantageous for tracts that connect two brain regions, e.g. for structural brain networks. However, a long-standing problem is that SPT only finds a single connection with no insight to its uncertainty. This is in contrast to walk-based methods, which are often probabilistic in nature. In this paper, we provide a first fully probabilistic SPT algorithm that models stochastic diffusion tensors and returns a distribution over the shortest path. While uncertainty propagates in probabilistic walk-based methods, the uncertainty of probabilistic SPT only depends on local data uncertainty, not the seed point distance.

Through experiments we visualize the estimated geodesic densities between brain regions, and validate that the estimated geodesic densities are less certain in areas where the estimated diffusion tensor is uncertain. We see that the visualized geodesic densities from the probabilistic SPT yield smoother and more coherent tracts than the corresponding Dijkstra solutions.

While Riemannian SPT only applies to second order tensor models, a Finsler geometry framework has emerged enabling continuous SPT with higher order tensors for HARDI data [1]. Our proposed numerical tools also extend to these models, and can provide a probabilistic interpretation of Finsler models. Further future work includes shape analysis on the resulting estimated geodesic densities, which are suitable for GP-based tract shape analysis [22].

**Acknowledgements.** Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil 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. S.H. is funded in part by the Danish Council for Independent Research (DFF), Natural Sciences. A.F. is funded in part by the DFF, Technology and Production Sciences.

## References

1. Astola, L., Jalba, A., Balmashnova, E., Florack, L.: Finsler streamline tracking with single tensor orientation distribution function for high angular resolution diffusion imaging. *Journal of Mathematical Imaging and Vision* 41(3), 170–181 (2011)
2. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 103(3), 247–254 (1994)
3. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magnetic resonance in medicine* 44(4), 625–632 (2000)
4. Catani, M., de Schotten, M.T.: A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44(8), 1105–1132 (2008)
5. Chkrebtii, O., Campbell, D., Girolami, M., Calderhead, B.: Bayesian uncertainty quantification for differential equations. *arXiv stat.ME 1306.2365* (Jun 2013)

6. Desikan, R.S., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31(3), 968–980 (2006)
7. van Essen, D., et al.: The Human Connectome Project: a data acquisition perspective. *NeuroImage* 62(4), 2222–2231 (2012)
8. van Essen, D., et al.: The WU-Minn Human Connectome Project: an overview. *NeuroImage* 80, 62–79 (2013)
9. Fuster, A., Tristan-Vega, A., Haije, T., Westin, C.F., Florack, L.: A novel Riemannian metric for geodesic tractography in DTI. In: *Computational Diffusion MRI and Brain Connectivity*, pp. 97–104. Math. and Visualization, Springer (2014)
10. Glasser, M., et al.: The minimal preprocessing pipelines for the human connectome project. *NeuroImage* 80, 105–124 (2013)
11. Hao, X., Zygmunt, K., Whitaker, R.T., Fletcher, P.T.: Improved segmentation of white matter tracts with adaptive Riemannian metrics. *Medical Image Analysis* 18(1), 161–175 (2014)
12. Hennig, P., Hauberg, S.: Probabilistic solutions to differential equations and their application to Riemannian statistics. In: *AISTATS*. vol. 33. JMLR, W&CP (2014)
13. LaGatta, T., Wehr, J.: Geodesics of random riemannian metrics. *Communications in Mathematical Physics* 327(1), 181–241 (2014)
14. Lenglet, C., Deriche, R., Faugeras, O.: Inferring white matter geometry from diffusion tensor MRI: Application to connectivity mapping. In: Pajdla, T., Matas, J. (eds.) *Computer Vision – ECCV*. Lecture Notes in Computer Science, vol. 3024, pp. 127–140. Springer (2004)
15. O’Donnell, L., Haker, S., Westin, C.F.: New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space. In: Dohi, T., Kikinis, R. (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI*. pp. 459–466. Lecture Notes in Computer Science, Springer (2002)
16. Parker, G., Haroon, H., Wheeler-Kingshott, C.: A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *J. Magn. Reson. Imaging* 18(2), 242–254 (2003)
17. Rasmussen, C., Williams, C.: *GPs for Machine Learning*. MIT Press (2006)
18. Schober, M., Duvenaud, D., Hennig, P.: Probabilistic ODE solvers with Runge-Kutta means. In: *NIPS* (2014)
19. Schober, M., Kasenburg, N., Feragen, A., Hennig, P., Hauberg, S.: Probabilistic shortest path tractography in DTI using Gaussian Process ODE solvers. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI*. Lecture Notes in Computer Science, vol. 8675, pp. 265–272. Springer (2014)
20. Skilling, J.: *Bayesian solution of ordinary differential equations*. Maximum Entropy and Bayesian Methods, Seattle (1991)
21. Sotiropoulos, S.N., et al.: Effects of image reconstruction on fiber orientation mapping from multichannel diffusion MRI: reducing the noise floor using SENSE. *Magn Reson Med* 70(6), 1682–89 (2013)
22. Wassermann, D., Rathi, Y., Bouix, S., Kubicki, M., Kikinis, R., Shenton, M.E., Westin, C.F.: White matter bundle registration and population analysis based on Gaussian Processes. In: *Proceedings of the 22nd International Conference on Information Processing in Medical Imaging*. pp. 320–332. Springer (2011)
23. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20(1), 45–57 (2001)