Integral Curve Estimation: Methodology and Applications to Diffusion Tensor Imaging

Lyudmila Sakhanenko (MSU)

Problem

Let $G \subset \mathbb{R}^d$ be a bounded open set, $v : G \mapsto \mathbb{R}^d$ be a vector field, $\mathrm{supp}(v) \subset G$, $v \in C^2$. The observations are

$$V_i = v(X_i) + \xi_i, \ i = 1, \ldots, n,$$

where $X_1, ..., X_n$ are iid uniform in G, $\xi_1, ..., \xi_n$ are iid, independent of X's, $\mathbb{E}\xi = 0$ and $\mathrm{Cov}(\xi, \xi) = \Sigma$.
Given $(X_i, V_i), i = 1, ..., n$, estimate $x(t)$ defined as

$$\frac{dx(t)}{dt} = v(x(t)), \ \ t \geq 0, \ x(0) = a \in G,$$

or equivalently $x(t) = a + \int\limits_0^t v(x(s))ds.$

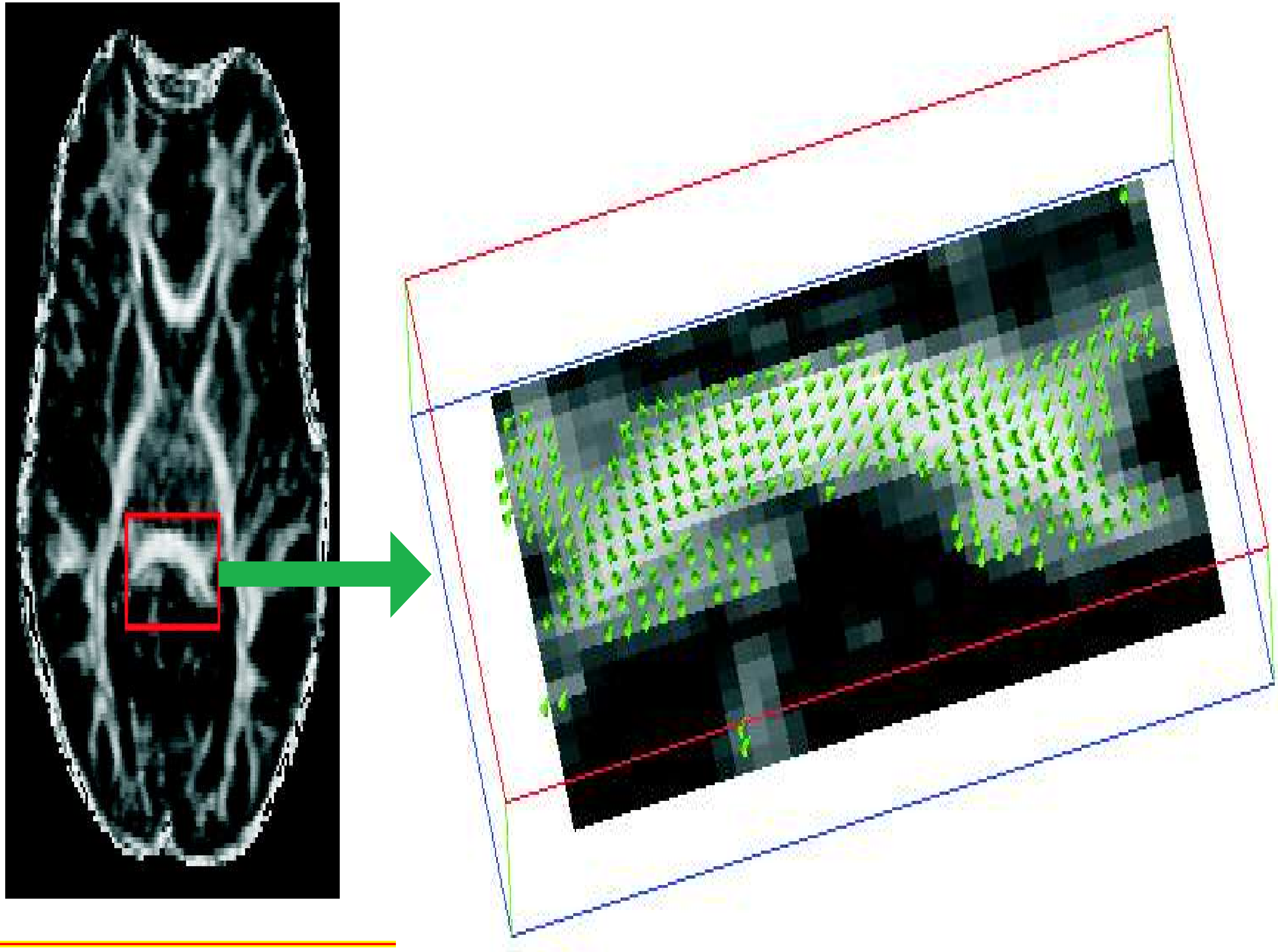
Motivation

Diffusion Tensor Imaging is a brain imaging technique based on measuring the diffusion tensor at discrete locations in the cerebral white matter.

The diffusion tensor data is used

- to estimate the dominant orientations of the diffusion,
- to track white matter fibers from the initial location following these orientations.

3D velocity from DT-MRI data



Diffusion Tensor Imaging

1. Water protons behave as gyroscopes in the presence of a magnetic field. Their motion is constrained in living tissues. The associated diffusion at a given location is characterized by a symmetric positively definite 3×3 diffusion matrix (diffusion tensor).

2. In cerebral white matter, the diffusion is typically anisotropic, whereas it is isotropic in cerebral grey matter.

3. The principal eigenvectors of the diffusion matrix show the dominant directions of the diffusion, which goes along axonal fibers.

4. The fiber tract (axon) can be reconstructed by following the directions of the vectors in small steps.

5. The diffusion tensor field is being measured at a discrete set of locations and each matrix in the field represents an average within a voxel corrupted with noise.

Existing methods

- B-spline smoothing in a least squared sense applied to the tensor field (Basser et al)
- Markov random field models, in particular, the spaghetti plate model (Poupon et al)
- a Monte Carlo approach to constructing of probabilistic connectivity maps, mixture of Gaussians models to handle branching (Parker)
- a bootstrap method of constructing confidence intervals for fiber orientation estimates, cones of uncertainty (Jones)
- PDEs methods (O'Donnell)
- Riemannian geometry methods to find geodesic paths from the tensor field (O'Donnell)
- fluid mechanics model based on the second order non-linear Navier-Stokes equations (Hageman et al)

Our approach

1. Nadaraya-Watson type estimate of v

$$\hat{V}(x) = \hat{V}_n(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) V_i$$

with some kernel K satisfying standard assumptions, in particular, support of K is bounded, $\int_{\mathbb{R}^d} K(x)dx = 1, \int_{\mathbb{R}^d} K(x)xdx = 0, \int_{\mathbb{R}^d} |K(x)||x|^2 dx < \infty$, and with a bandwidth parameter $h = h_n$.

2. A plug-in estimate of the solution $x(t),\ t \geq 0,$

$$\hat{X}(t) = \hat{X}_n(t),\ t \geq 0,$$

the solution of the following Cauchy problem:

$$\frac{d\hat{X}(t)}{dt} = \hat{V}(\hat{X}(t)),\ \ t \geq 0,\ \hat{X}(0) = a \in G,$$

which is equivalent to the integral equation

$$\hat{X}(t) = a + \int_0^t \hat{V}(\hat{X}(s))ds.$$

Theorem 1

Suppose $h_n \to 0$ *and* $nh_n^{d+2} \to \infty$ *as* $n \to \infty$. *Then* $\forall T > 0$

$$\sup_{0 \le t \le T} |\hat{X}_n(t) - x(t)| \to^P 0 \text{ as } n \to \infty.$$

Suppose also $nh_n^{d+3} \to \beta \ge 0$ *as* $n \to \infty$ *and for some* $\gamma = \gamma_T > 0$ *and* $\forall s, t \in [0, T] \left| \frac{1}{t-s} \int_s^t v(x(\lambda)) d\lambda \right| \ge \gamma,\ s \ne t.$

Then the sequence of stochastic processes

$$\sqrt{nh^{d-1}}(\hat{X}_n(t) - x(t)),\ 0 \le t \le T$$

converges weakly in the space $C[0, T]$ *to the Gaussian process* $\xi(t)$ *with mean* $M_\beta(t)$ *and covariance* $C(t_1, t_2)$.

Theorem 1 (cont.)

$$\begin{aligned} M_\beta(t) &= \sqrt{\beta} M(t), \\ M(t) &= \frac{1}{2}\int_0^t U(t,s)\int K(z)\langle v''(x(s))z, z\rangle dz ds, \\ C(t_1,t_2) &= \int\limits_0^{t_1 \wedge t_2} \psi(v(x(s)))U(t_1,s)[\Sigma + \\ & \quad v(x(s))v^*(x(s))]U^*(t_2,s)ds \\ \text{with} \quad \psi(v) &= \int\int K(z)K(z+v\tau)dzd\tau, \end{aligned}$$

and $\frac{dU(t,s)}{dt} = v'(x(t))U(t,s), \;\; U(s,s) = \mathbb{I}.$

SDE for $\xi(t)$

$$d\xi(t) = \frac{\sqrt{\beta}}{2}\int_{\mathbb{R}^d} K(u)\langle v''(x(t))u, u\rangle du dt$$

$$+v'(x(t))\xi(t)dt + \left(\psi(v(x(t)))[\Sigma + v(x(t))v^*(x(t))]\right)^{1/2} dW(t)$$

with the initial condition $\xi(0) = 0$, where $W(t), t \geq 0$, is a standard Brownian motion in $\mathbb{R}^d$.

Intuition

- Problem:

estimate a density function $f(x_1, ..., x_d)$ in $\mathbb{R}^d$

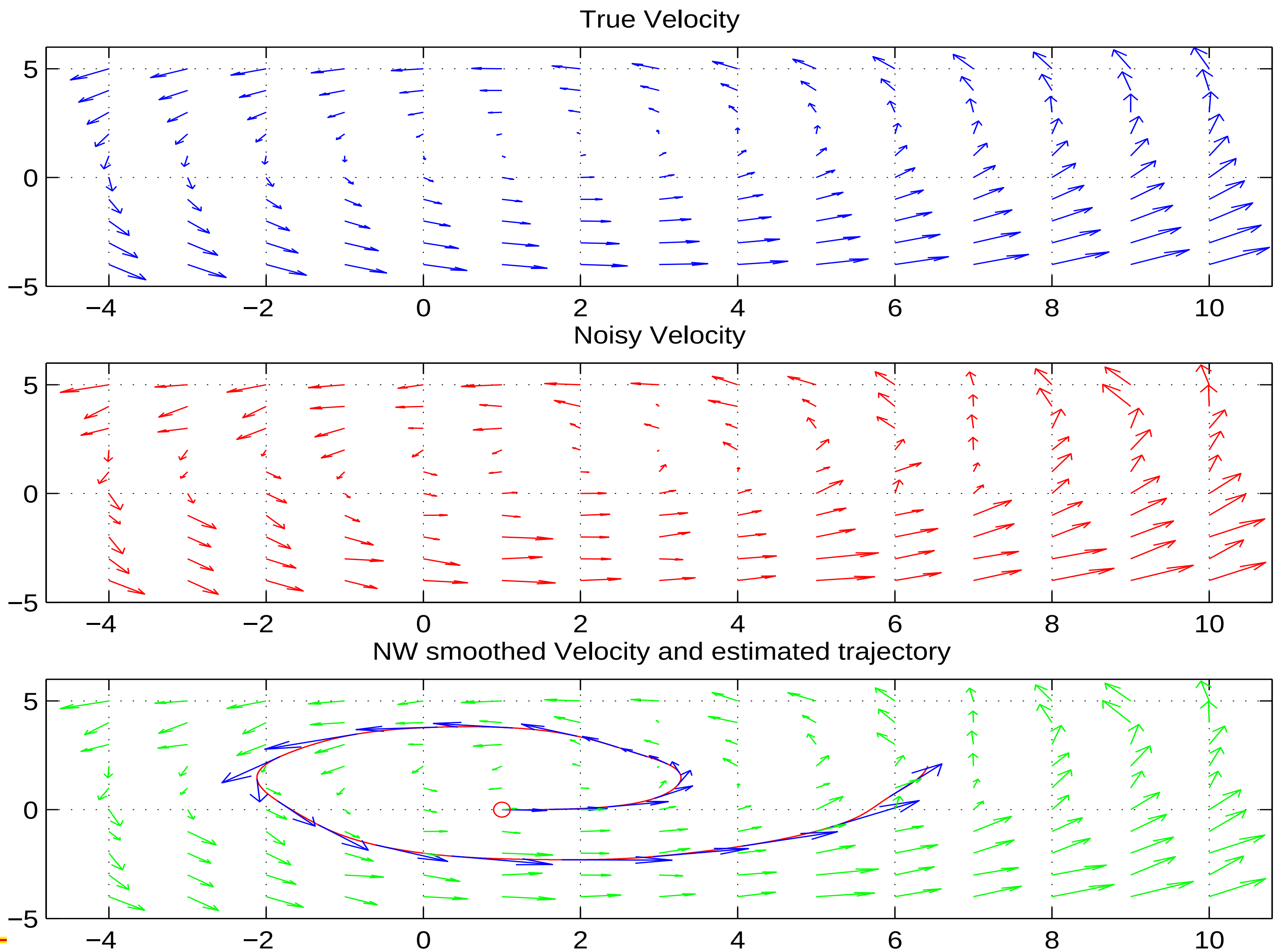
Kernel density estimator of f gives the rate of $\sqrt{nh^d}$

- Problem:

estimate $g(x_1, ..., x_{d-1}) := \int_{\mathbb{R}} f(x_1, ..., x_d) dx_d$

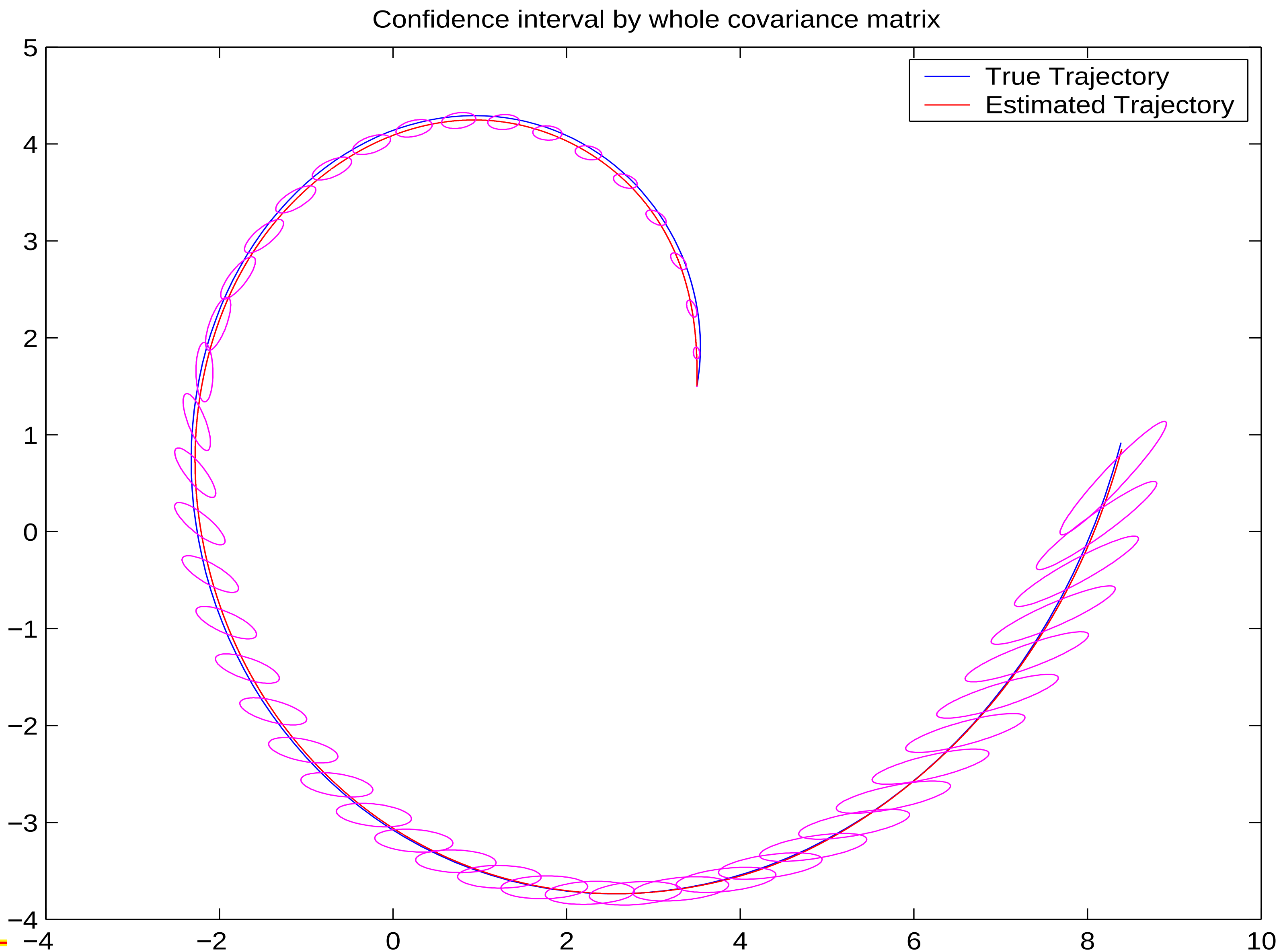
Kernel density estimator of g gives the rate of $\sqrt{nh^{d-1}}$

The spiral


True Velocity
Noisy Velocity
NW smoothed Velocity and estimated trajectory
5
0
−5
−4
−2
0
2
4
6
8
10

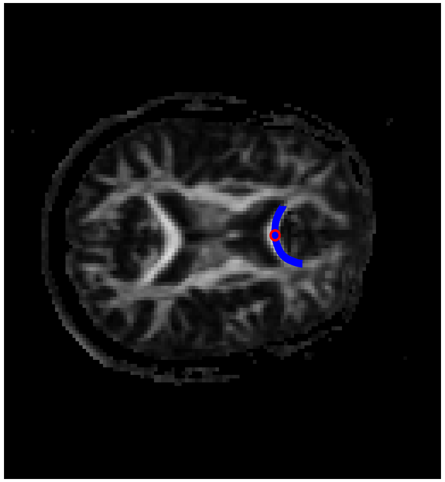
The spiral (cont.)

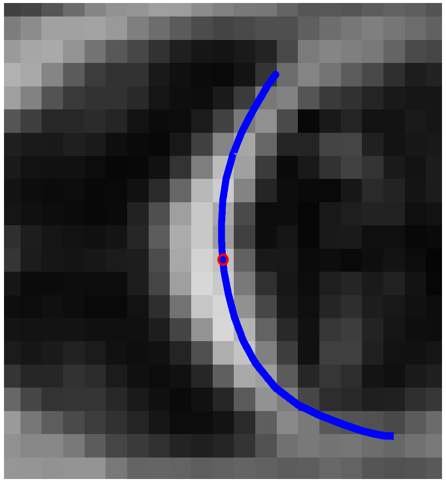
Figure shows the results in the case of spiral integral curves. The true vector field is given by the following formulae: $v_x = -y + 0.2x,\ v_y = x + 0.2y$. The noisy vector field is simulated by adding to the true field independent copies of $3Z$, Z being a 2-dimensional standard normal vector. The estimated trajectory (represented by the red curve) starts at $(1, 0)$ (the small red circle). In this example, $n = 150, h = 0.6, \delta = 0.02$.



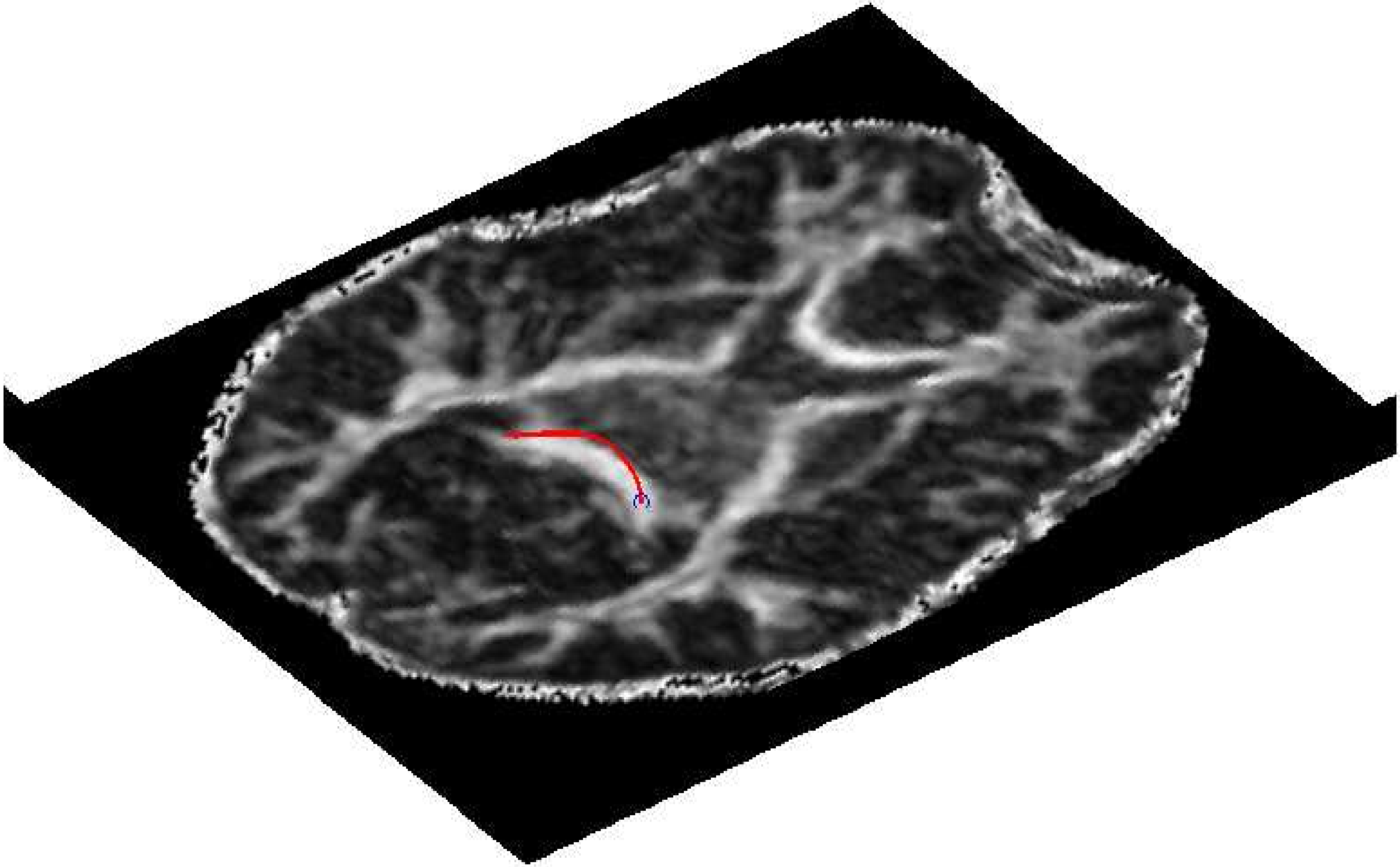
Confidence interval by whole covariance matrix
True Trajectory
Estimated Trajectory
5
4
3
2
1
0
−1
−2
−3
−4
−4
−2
0
2
4
6
8
10

Estimated fiber trajectory

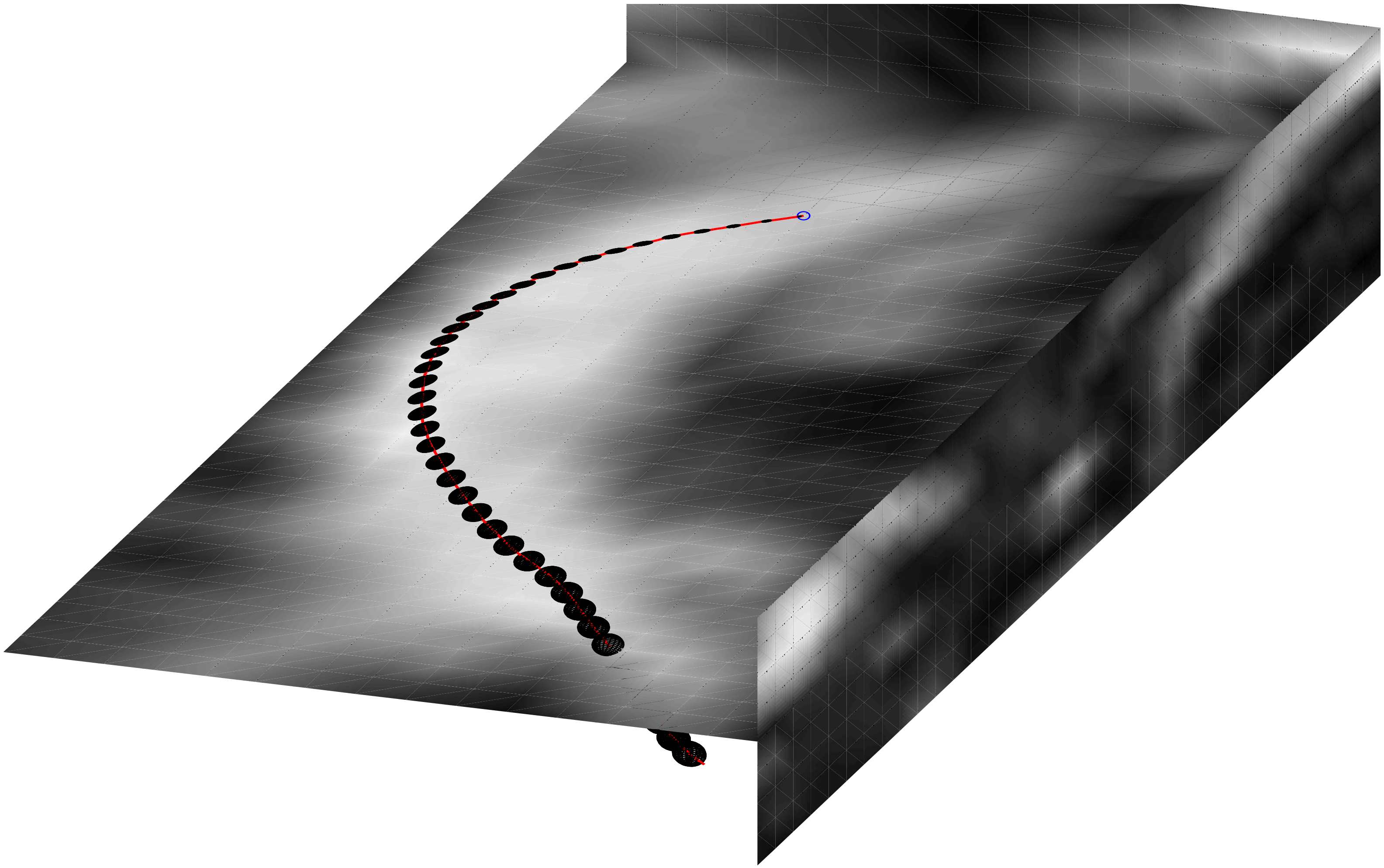




Estimated fiber trajectory



3-D confidence ellipsoids



Does the fiber reach the region?

Let Γ be a subset of G, $d(x, \Gamma)$ be a distance from x to Γ, ψ be a monotone function on $\mathbb{R}_+$. Define $\varphi(x) := \psi(d(x, \Gamma))$ and

$$M := \{\tau \in [0, T] : \varphi(x(\tau)) = \inf_{0 \le t \le T} \varphi(x(t))\}.$$

Goal: study convergence in distribution (subject to a proper normalization) of

$$\inf_{t \in [0,T]} \varphi(\hat{X}(t)) - \inf_{t \in [0,T]} \varphi(x(t)).$$

Theorem 2

Suppose $\varphi \in C^1(G)$ *and the conditions of Theorem 1 hold. Then the sequence of random variables*

$$\sqrt{nh^{d-1}}\left[\inf_{t\in[0,T]} \varphi(\hat{X}(t)) - \inf_{t\in[0,T]} \varphi(x(t))\right]$$

converges in distribution to

$$\inf_{\tau\in M} \xi(\tau)^* \varphi'(x(\tau)).$$

In particular, if M *consists only of one point* $\tau \in (0,T)$, *then the above sequence is asymptotically normal with mean* $M_\beta(\tau)$ *and variance* $\sigma^2 = (\varphi'(x(\tau)))^* C(\tau,\tau)\varphi'(x(\tau))$.

Theorem 2 (cont.)

Suppose now $\varphi \in C^2(G)$*. If* $\forall \tau \in M$, $\varphi'(x(\tau)) = 0$ *and* $\varphi''(x(\tau))(v(x(\tau)), v(x(\tau))) > 0$, *then the sequence of random variables*

$$nh^{d-1}\left[\inf_{t\in[0,T]} \varphi(\hat{X}(t)) - \inf_{t\in[0,T]} \varphi(x(t))\right]$$

converges in distribution to

$$\frac{1}{2}\inf_{\tau\in M}\left[\varphi''(x(\tau))(\xi(\tau),\xi(\tau)) - \frac{\Big(\varphi''(x(\tau))(v(x(\tau)),\xi(\tau))\Big)^2}{\varphi''(x(\tau))(v(x(\tau)),v(x(\tau)))}\right].$$

If M *consists only of one point* τ, *then the limit becomes*

$$\frac{1}{2}\left[\varphi''(x(\tau))(Z,Z) - \frac{\Big(\varphi''(x(\tau))(v(x(\tau)),Z)\Big)^2}{\varphi''(x(\tau))(v(x(\tau)),v(x(\tau)))}\right],$$

Theorem 2 (cont.)

where Z *is a normal random vector in* $\mathbb{R}^d$ *with mean* $M_\beta(\tau)$ *and covariance* $C(\tau,\tau)$.

On the other hand, if for all $u \in \mathbb{R}^d$

$$\varphi''(x(\tau))(v(x(\tau)),u) = 0,$$

then the distributional limit of the sequence

$$nh^{d-1}\left[\inf_{t\in[0,T]} \varphi(\hat{X}(t)) - \inf_{t\in[0,T]} \varphi(x(t))\right]$$

is

$$\frac{1}{2}\inf_{\tau\in M} \varphi''(x(\tau))(\xi(\tau),\xi(\tau)),$$

which in the unique minimum case becomes $\frac{1}{2}\varphi''(x(\tau))(Z,Z)$.

Example

Let $b \in G$ and let

$$\varphi(x) := |x - b|^2.$$

Then

$$\varphi'(x) = 2(x - b),$$

$$\varphi''(x) = 2\mathbb{I}.$$

The asymptotic behavior of

$$\inf_{t \in [0,T]} |\hat{X}(t) - b|^2$$

is described by the following corollary.

Corollary

Suppose that for some $\tau \in (0, T)$

$$\inf_{0 \le t \le T} |x(t) - b|^2 = |x(\tau) - b|^2,$$

and, moreover, suppose that τ *is the only point where the infimum is attained. Suppose also the conditions of Theorem 1 hold.*
1. If $x(\tau) \neq b$, *then the sequence*

$$\sqrt{nh^{d-1}} \left[\inf_{0 \le t \le T} |\hat{X}(t) - b|^2 - \inf_{0 \le t \le T} |x(t) - b|^2 \right]$$

is asymptotically normal with mean $2M_\beta(\tau)^*(x(\tau) - b)$ *and variance* $\sigma^2 = 4(x(\tau) - b)^* C(\tau, \tau)(x(\tau) - b)$.

Corollary (cont.)

2. If $x(\tau) = b$*, then the sequence*

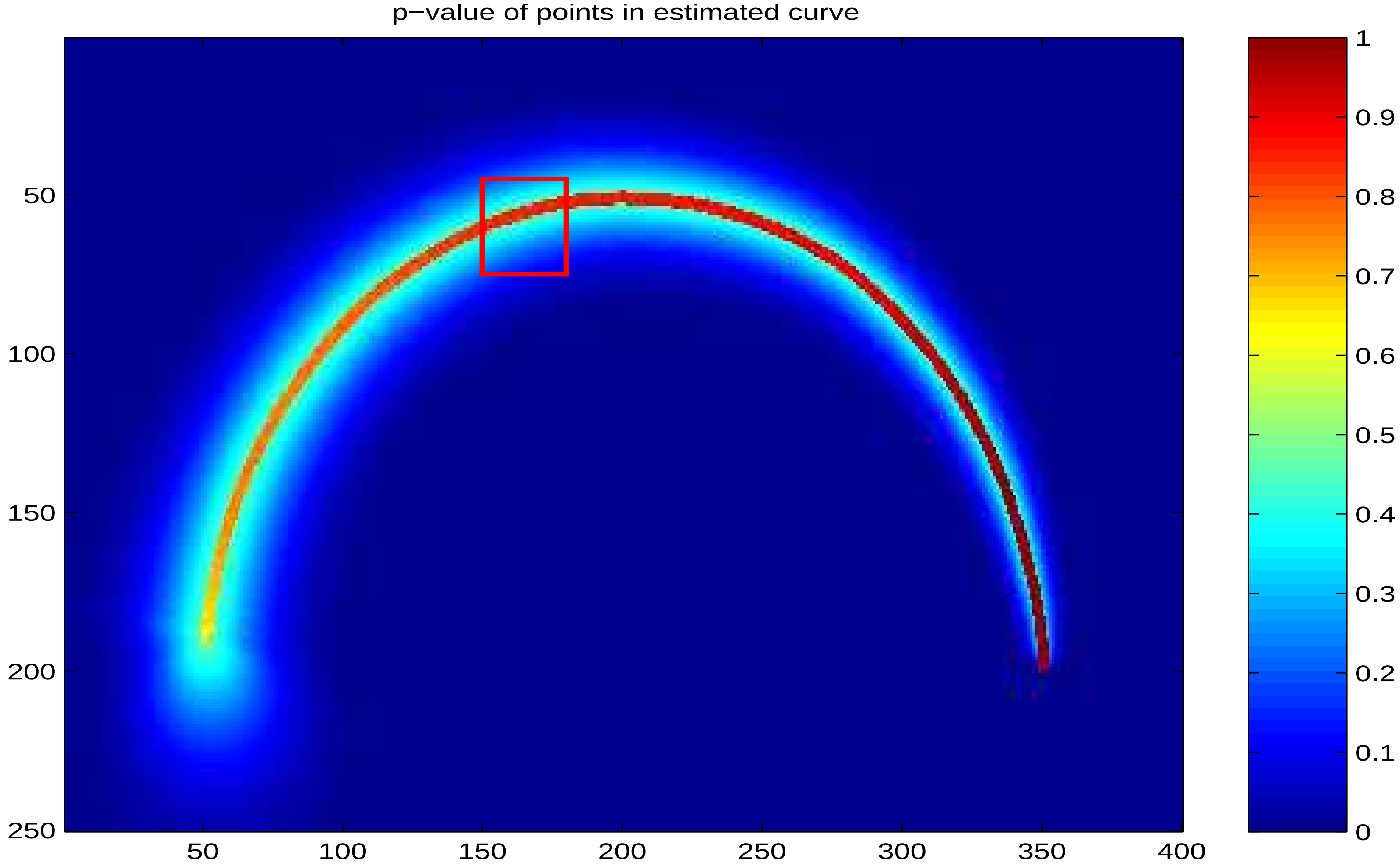
$$nh^{d-1} \inf_{0 \le t \le T} |\hat{X}(t) - b|^2$$

converges in distribution to the random variable

$$|Z|^2 - \frac{\left(v(x(\tau))^* Z\right)^2}{|v(x(\tau))|^2},$$

where Z *is a normal random vector in* $\mathbb{R}^d$ *with mean* $M_\beta(\tau)$ *and covariance* $C(\tau, \tau)$.

P-value map



H_0 : the true integral curve passes through the point

Optimal rates

Fix $T > 0$. Let $\mathcal{C}^2(a, G, T)$ be the class of vector fields satisfying conditions
$\mathrm{supp}(v) \subset G$,
there exists $\gamma > 0$ such that $\left|\frac{1}{t-s}\int_s^t v(x(\lambda))d\lambda\right| \geq \gamma$ for all $0 \leq s < t \leq T$,
that are 0 outside of G and not 0 at a.
Let $\mathcal{W}$ be the class of all non-trivial even functions $w : \mathbb{R} \mapsto \mathbb{R}$ that are non-decreasing on $\mathbb{R}_+$ and 0 at 0. (For example, $u^2, |u|, I(|u| > c)$.)
Let $\mathcal{E}_n(T)$ denote the class of all possible estimators of the integral curve $x(t), t \in [0, T]$, based on the data $(X_i, V_i), i = 1, ..., n$.

Theorem 3

For any point $a \in G$*, any* $t_0 \in (0,T]$*, any function* $w \in \mathcal{W}$*, any index* $l = 1, ..., d$ *we have*

$$\liminf_{n\to\infty} \inf_{\hat{X}_n \in \mathcal{E}_n(T)} \sup_{v \in \mathcal{C}^2(a,G,T)} \mathbb{E} w\left(n^{2/(d+3)} |\hat{X}_{n,l}(t_0) - x_l(t_0)| \right) > 0.$$

Theorem 4

For a fixed $T > 0$ let $\mathcal{M}_n(T)$ denote the class of all possible estimators of $\inf_{t\in[0,T]} \varphi(x(t))$ based on the data $(X_i, V_i), i = 1, ..., n$.

For any closed subset $\Gamma \subset G$ and any point $a \in G \backslash \Gamma$ such that φ is continuously differentiable around a and $\nabla\varphi(a) \neq 0$, and for any function $w \in \mathcal{W}$ and any $T > 0$ we have

$$\liminf_{n\to\infty} \inf_{\hat{M}_n \in \mathcal{M}_n(T)} \sup_{v \in \mathcal{C}^2(a,G,T)} \mathbb{E} w\left(n^{2/(d+3)} \left| \hat{M}_n - \inf_{t\in[0,T]} \varphi(x(t)) \right| \right) > 0.$$

Theorem 5

For a vector function $y(t), t \in [0,T]$, denote

$$\|y\|_{p,T} := \left(\int_0^T \sum_{i=1}^{d} |y_i(t)|^p dt \right)^{1/p}.$$

For any point $a \in G$ and any function $w \in \mathcal{W}$ we have

$$\liminf_{n\to\infty} \inf_{\hat{X}_n \in \mathcal{E}_n(T)} \sup_{v \in \mathcal{C}^2(a,G,T)} \mathbb{E} w\left(n^{2/(d+3)} \|\hat{X}_n - x\|_{p,T} \right) > 0.$$

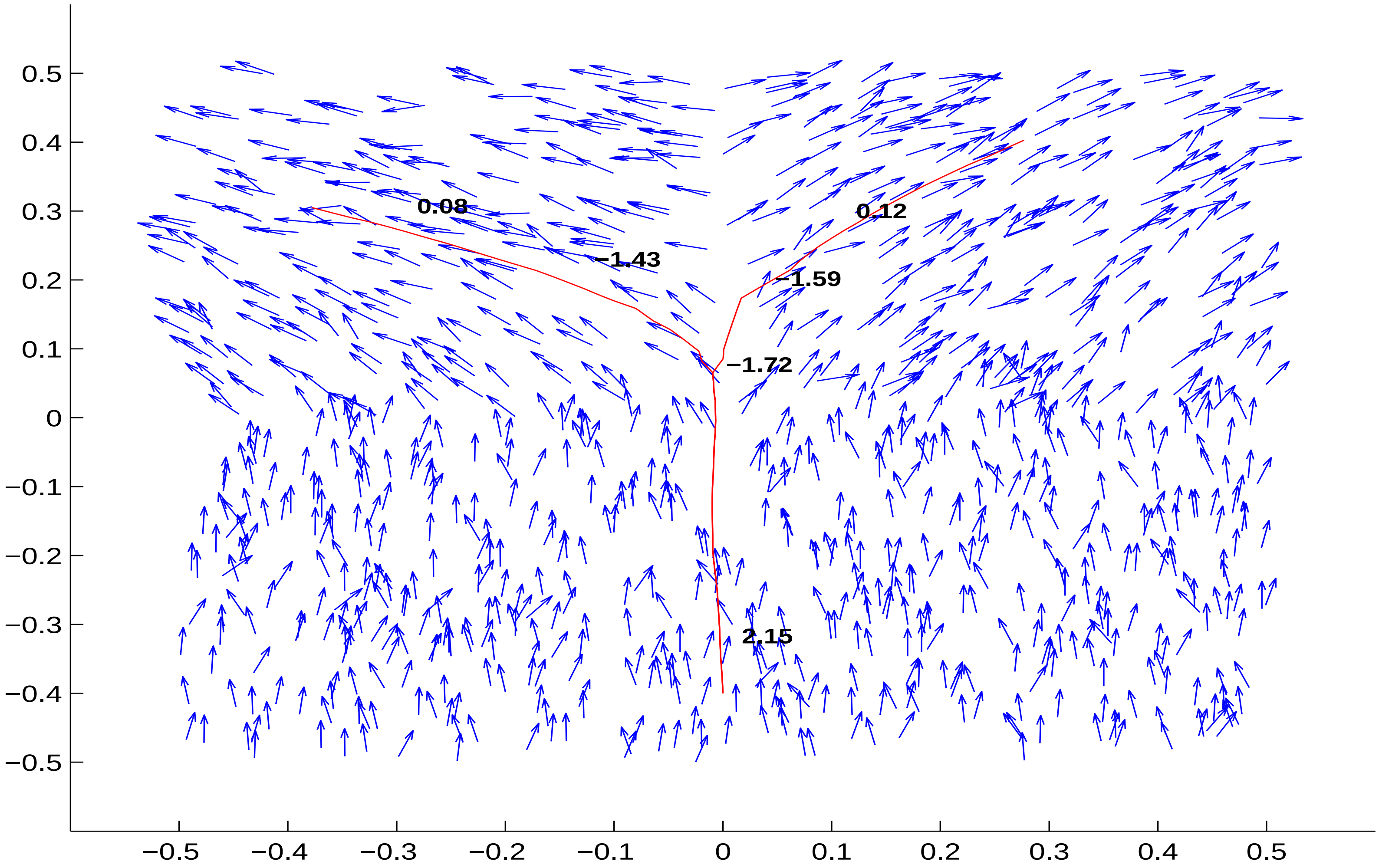
Branching: Heuristics

Usually, in DTI $|v(x)| = 1$. If the vector field is smooth around a given point $x \in G$, then $\hat{V}(x)$, as a weighted local average of unit vectors $v(X_i)$ of approximately the same direction plus a small noise, should have the norm close to 1 (for large enough n).
On the other hand, if two or more trajectories intersect at the point x, then $\hat{V}(x)$ becomes the weighted local average of unit vectors of several different directions and, as a result, $|\hat{V}(x)|$ will be, with a high probability, significantly smaller than 1. The statistic

$$\nu_n := \sqrt{nh_n^d}\frac{|\hat{V}(\hat{X}(t))|^2 - 1}{\hat{\sigma}}$$

becomes asymptotically standard normal.

Branching: Simulation



Ideas of the proof of Theorem 1

1. $\hat{V}, \hat{V}'$ are consistent estimates of v, v' uniformly in $\mathbb{R}^d$.

2. Let $y(t) := \hat{X}(t) - x(t) = \int_0^t [\hat{V}(\hat{X}(s)) - v(x(s))]ds =$

$\int_0^t (\hat{V} - v)(\hat{X}(s))ds + \int_0^t [v(\hat{X}(s)) - v(x(s))]ds.$

Then $|y(t)| \le T \sup_{x \in \mathbb{R}^d} |\hat{V}(x) - v(x)| + L \int_0^t |y(s)|ds.$

By Gronwall–Bellman inequality, for all $t \in [0, T]$

$$|y(t)| \le T \sup_{x \in \mathbb{R}^d} |\hat{V}(x) - v(x)| e^{Lt}.$$

3. $y(t) = \int_0^t (\hat{V} - v)(x(s))ds + \int_0^t v'(x(s))y(s)ds + R(t)$

$$\sup_{0 \le t \le T} |R(t)| = o_p\left(\sup_{0 \le t \le T} |y(t)|\right).$$

4. Denote $z(t) := \int_0^t (\hat{V} - v)(x(s))ds + \int_0^t v'(x(s))z(s)ds.$

Equivalently, $\frac{dz(t)}{dt} = \hat{V}(x(t)) - v(x(t)) + v'(x(t))z(t),\ z(0) = 0.$

Then $z(t) = \int_0^t U(t,s)[\hat{V}(x(s)) - v(x(s))]ds$ and

$$
\begin{aligned}
z(t) &= \frac{1}{nh^d} \sum_{i=1}^{n} \int I_{[0,t]}(s) U(t,s) K(\frac{(x(s) - X_i)}{h}) ds (v(X_i) + \xi_i) \\
&- \int I_{[0,t]}(s) U(t,s) v(x(s)) ds
\end{aligned}
$$

5. Denote $\delta(t) := y(t) - z(t)$. By Gronwall-Bellman inequality and the computation of the mean and the covariance function of $z(t)$

$$\sup_{0 \le t \le T} |\delta(t)| = o_p\left(\int_0^T |z(t)| dt\right) = o_p\left(\frac{1}{\sqrt{nh^{d-1}}}\right).$$

6. Asymptotic normality of $\sqrt{nh^{d-1}} z(t)$ follows from Lyapunov's conditions of the CLT and the asymptotic equicontinuity condition.

Ideas of the proof of Theorems 3-4

1. A small parametric subclass of $\mathcal{C}^2(a, G, T)$ is constructed.
2. Lower bounds for the difference between the parameter and its estimate are obtained from results for *locally asymptotically normal* families of distributions (key point).
3. The deviation between a curve estimate and the true curve is written as the difference between this parameter and its estimate.

Future work

- Fixed design
- Non-parametric regression vs additive noise (since usually $|V_i| = 1$)
- Smooth the tensor, then obtain principle eigenvectors
- Branching of fiber paths

Thank You!