Nonparametric Inference for Geometric Objects

Wolfgang Polonik

Department of Statistics, UC Davis

Van Dantzig Seminar, University of Leiden, The Netherlands, Oct. 7, 2015
Outline:

- inference for geometric features/objects - overview
- distribution theory for filament estimation
- suprema of gaussian processes on growing manifolds
Inference for geometric objects
Inference for geometric objects

- Estimation of integral curves
Inference for geometric objects

- Estimation of integral curves
- Estimation of level sets
Inference for geometric objects

- Estimation of integral curves
- Estimation of level sets
- Inference for modes / modal clustering
Inference for geometric objects

- Estimation of integral curves
- Estimation of level sets
- Inference for modes / modal clustering
- Estimation and inference for persistent homology
  (topological data analysis)
Inference for geometric objects

- Estimation of integral curves
- Estimation of level sets
- Inference for modes / modal clustering
- Estimation and inference for persistent homology
  (topological data analysis)
- Filament estimation
Estimation of integral curves

Given $v : \mathbb{R}^d \to \mathbb{R}^d$ and starting point $x_0$, the integral curve $X : [0, T] \to \mathbb{R}^d$ is solution to
\[ \frac{d}{dt}X(t) = v(X(t)), \quad X(0) = x_0. \]

Estimation (Koltchinskii et al. 2007):

Model:
\[ V_i = v(X_i) + \epsilon_i, \quad \epsilon_i \text{ iid., } X_i \text{ iid, uniform on } G, \text{ indep. of } \epsilon_i \]

Applications in medical imaging (DTI); filament estimation; etc.

Consider \( \hat{V}(x) = \frac{1}{nh} \sum_{i=1}^{n} K(X_i - x) v_i \) and estimate $X(t)$ via
\[ \frac{d}{dt} \hat{X}(t) = \hat{V}(\hat{X}(t)), \quad \hat{X}(0) = x_0. \]
Estimation of integral curves

Given $\nu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and starting point $x_0$
Estimation of integral curves

Given \( \nu : \mathbb{R}^d \to \mathbb{R}^d \) and starting point \( x_0 \)

integral curve \( \mathcal{X} : [0, T] \to \mathbb{R}^d \) is solution to

\[
\frac{d}{dt} \mathcal{X}(t) = \nu(\mathcal{X}(t)), \quad \mathcal{X}(0) = x_0.
\]
Estimation of integral curves

Given \( \nu : \mathbb{R}^d \rightarrow \mathbb{R}^d \) and starting point \( x_0 \)

integral curve \( \mathcal{X} : [0, T] \rightarrow \mathbb{R}^d \) is solution to

\[
\frac{d}{dt} \mathcal{X}(t) = \nu(\mathcal{X}(t)), \quad \mathcal{X}(0) = x_0.
\]

Estimation (Koltchinskii et al. 2007):

Model: \( V_i = \nu(X_i) + \epsilon_i, \ \epsilon_i \text{ iid.}, \ X_i \text{ iid, uniform on } G, \text{ indep. of } \epsilon_i \)
Estimation of integral curves

Given \( v : \mathbb{R}^d \to \mathbb{R}^d \) and starting point \( x_0 \)

integral curve \( \mathcal{X} : [0, T] \to \mathbb{R}^d \) is solution to

\[
\frac{d}{dt} \mathcal{X}(t) = v(\mathcal{X}(t)), \quad \mathcal{X}(0) = x_0.
\]

Estimation (Koltchinskii et al. 2007):

Model: \( V_i = v(X_i) + \epsilon_i, \; \epsilon_i \text{ iid.}, \; X_i \text{ iid, uniform on } G, \text{ indep. of } \epsilon_i \)

Applications in medical imaging (DTI); filament estimation; etc.
Estimation of integral curves

Given $\nu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and starting point $x_0$

integral curve $\mathcal{X} : [0, T] \rightarrow \mathbb{R}^d$ is solution to

$$\frac{d}{dt} \mathcal{X}(t) = \nu(\mathcal{X}(t)), \quad \mathcal{X}(0) = x_0.$$

Estimation (Koltchinskii et al. 2007):

Model: $V_i = \nu(X_i) + \epsilon_i$, $\epsilon_i$ iid., $X_i$ iid, uniform on $G$, indep. of $\epsilon_i$

Applications in medical imaging (DTI); filament estimation; etc.

Consider $\widehat{V}(x) = \frac{1}{nh^d} \sum_{i=1}^n K \left( \frac{X_i - x}{h} \right) V_i$ and estimate $\mathcal{X}(t)$ via

$$\frac{d}{dt} \widehat{\mathcal{X}}(t) = \widehat{V}(\widehat{\mathcal{X}}(t)), \quad \widehat{\mathcal{X}}(0) = x_0.$$
Estimation of integral curves

Koltchinskii et al. (2007) show that under appropriate assumptions
\[ \sqrt{n} h^{d-1} (\hat{X}(t) - x(t)) \rightarrow_D G(t), \]
where \( T > 0, \) \( \{G(t), 0 \leq t \leq T\} \) mean zero Gaussian process.

Heuristics underlying the derivation of the rate:
• Integral curve: \( X(t) = x_0 + \int_0^t V(X(s)) \, ds; \)
• estimated integral curve: \( \hat{X}(t) = x_0 + \int_0^t \hat{V}(\hat{X}(s)) \, ds; \)
\[ \Rightarrow \hat{X}(t) - X(t) = \int_0^t [\hat{V}(\hat{X}(s)) - V(X(s))] \, ds. \]

Rate of convergence of
\[ \hat{V}(\hat{X}(s)) - V(X(s)) = O_p\left(\sqrt{nh^{d-1}}\right); \]
integration \( \Rightarrow \) gain of one power of \( h. \)
Koltchinskii et al. (2007) show that under appropriate assumptions

\[ \sqrt{nh^{d-1}} \left( \hat{\mathcal{X}}(t) - x(t) \right) \to \mathcal{D} \ G(t), \quad 0 \leq t \leq T, \]

where \( T > 0, \{G(t), 0 \leq t \leq T\} \) mean zero Gaussian process.
Estimation of integral curves

Koltchinskii et al. (2007) show that under appropriate assumptions

$$\sqrt{nh^{d-1}} \left( \hat{X}(t) - x(t) \right) \to_D G(t), \quad 0 \leq t \leq T,$$

where $T > 0$, $\{G(t), 0 \leq t \leq T\}$ mean zero Gaussian process.

Heuristics underlying the derivation of the rate:

- Integral curve: $\mathcal{X}(t) = x_0 + \int_0^t V(\mathcal{X}(s)) \, ds$;
- estimated integral curve: $\hat{\mathcal{X}}(t) = x_0 + \int_0^t \hat{V}(\hat{\mathcal{X}}(s)) \, ds$;
Estimation of integral curves

Koltchinskii et al. (2007) show that under appropriate assumptions

$$\sqrt{nh^{d-1}} \left( \hat{X}(t) - x(t) \right) \rightarrow_{D} G(t), \quad 0 \leq t \leq T,$$

where $T > 0$, $\{G(t), 0 \leq t \leq T\}$ mean zero Gaussian process.

**Heuristics** underlying the derivation of the rate:

- Integral curve: $\mathcal{X}(t) = x_0 + \int_0^t V(\mathcal{X}(s)) \, ds$;
- estimated integral curve: $\hat{\mathcal{X}}(t) = x_0 + \int_0^t \hat{V}(\hat{\mathcal{X}}(s)) \, ds$;

\[ \Rightarrow \quad \hat{\mathcal{X}}(t) - \mathcal{X}(t) = \int_0^t [\hat{V}(\hat{\mathcal{X}}(s)) - V(\mathcal{X}(s))] \, ds \]
Estimation of integral curves

Koltchinskii et al. (2007) show that under appropriate assumptions

$$\sqrt{nh^{d-1}} \left( \hat{X}(t) - x(t) \right) \rightarrow_{D} G(t), \quad 0 \leq t \leq T,$$

where $T > 0, \{G(t), 0 \leq t \leq T\}$ mean zero Gaussian process.

Heuristics underlying the derivation of the rate:

- Integral curve: $X(t) = x_0 + \int_0^t V(X(s)) \, ds$;
- estimated integral curve: $\hat{X}(t) = x_0 + \int_0^t \hat{V}(\hat{X}(s)) \, ds$;

$$\Rightarrow \hat{X}(t) - X(t) = \int_0^t \left[ \hat{V}(\hat{X}(s)) - V(X(s)) \right] \, ds.$$

Rate of convergence of $\hat{V}(\hat{X}(s)) - V(X(s)) = O_P((nh^d)^{-1})$;
Koltchinskii et al. (2007) show that under appropriate assumptions

\[ \sqrt{nh^{d-1}} \left( \hat{X}(t) - x(t) \right) \to_{\mathcal{D}} G(t), \quad 0 \leq t \leq T, \]

where \( T > 0, \{ G(t), 0 \leq t \leq T \} \) mean zero Gaussian process.

**Heuristics** underlying the derivation of the rate:

- Integral curve: \( \mathcal{X}(t) = x_0 + \int_0^t V(\mathcal{X}(s)) \, ds \);
- estimated integral curve: \( \hat{\mathcal{X}}(t) = x_0 + \int_0^t \hat{V}(\hat{\mathcal{X}}(s)) \, ds \);

\[ \leadsto \hat{\mathcal{X}}(t) - \mathcal{X}(t) = \int_0^t [\hat{V}(\hat{\mathcal{X}}(s)) - V(\mathcal{X}(s))] \, ds \]

Rate of convergence of \( \hat{V}(\hat{\mathcal{X}}(s)) - V(\mathcal{X}(s)) = O_P((nh^d)^{-1}) \);

integration \( \leadsto \) gain of one power of \( h \).
Estimation of integral curves

Note also that

\[ \hat{X}(t) - x(t) = \int_0^t \left[ \hat{V}(\hat{X}(s)) - V(x(s)) \right] ds \]

\[ = \int_0^t (\hat{V} - V)(x(s)) ds + \int_0^t v'(x(s))(\hat{X}(s) - x(s)) ds + r_n \]
Estimation of integral curves

Note also that

\[ \hat{X}(t) - x(t) = \int_0^t \left[ \hat{V}(\hat{X}(s)) - V(x(s)) \right] ds \]

\[ = \int_0^t (\hat{V} - V)(x(s)) ds + \int_0^t v'(x(s))(\hat{X}(s) - x(s)) ds + r_n \]

This indicates that process \( \hat{X}(t) - x(t) \) appropriately normalized is closely related to a solution to stochastic differential equation.
Estimation of integral curves

Note also that

\[ \hat{X}(t) - x(t) = \int_0^t \left[ \hat{V}(\hat{X}(s)) - V(x(s)) \right] ds \]

\[ = \int_0^t (\hat{V} - V)(x(s)) ds + \int_0^t v'(x(s))(\hat{X}(s) - x(s)) ds + r_n \]

This indicates that process \( \hat{X}(t) - x(t) \) appropriately normalized is closely related to a solution to stochastic differential equation.

Integral curves driven by second eigenvector of Hessian

Qiao and WP (2015); dimension $d = 2$. 
Integral curves driven by second eigenvector of Hessian

Qiao and WP (2015); dimension $d = 2$.

driving vector field: $v(x) = \text{second eigenvector of Hessian}$. 
Integral curves driven by second eigenvector of Hessian

Qiao and WP (2015); dimension $d = 2$.

driving vector field: $v(x) = \text{second eigenvector of Hessian}$.

Motivation: Filament (ridge line) estimation.

More later.
Estimation of level sets

Level sets of a function $f: \mathbb{R}^d \to \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}([\lambda, \infty]) .$$

Note: regularity $\Rightarrow$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!

Direct estimates:
- Volume (length) of MV-sets: generalized quantiles (Grübel, 1988; Einmahl and Mason, 1992; WP 1997)

Plug-in approach via kernel density estimation:
Estimation of level sets

Level sets of a function \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) are given by

\[
\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].
\]
Estimation of level sets

Level sets of a function $f : \mathbb{R}^d \to \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{x \in \mathbb{R}^d : f(x) \geq \lambda\} = f^{-1}[\lambda, \infty].$$

Note: regularity $\leadsto$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!
Estimation of level sets

Level sets of a function \( f : \mathbb{R}^d \to \mathbb{R} \) are given by

\[
\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].
\]

Note: regularity \( \rightsquigarrow \) boundaries of level sets \( f^{-1}(\lambda) \) are integral curves!

- direct estimates
- plug-in approach via kernel density estimation
Estimation of level sets

Level sets of a function $f : \mathbb{R}^d \to \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].$$

Note: regularity $\leadsto$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!

- direct estimates
  - excess mass approach

- plug-in approach via kernel density estimation
Estimation of level sets

Level sets of a function \( f : \mathbb{R}^d \to \mathbb{R} \) are given by

\[
\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].
\]

**Note:** regularity \( \rightsquigarrow \) boundaries of level sets \( f^{-1}(\lambda) \) are integral curves!

- direct estimates
  - excess mass approach
- minimum volume sets:
  - plug-in approach via kernel density estimation
Estimation of level sets

Level sets of a function $f : \mathbb{R}^d \to \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].$$

Note: regularity $\Rightarrow$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!

- **direct estimates**
  - minimum volume sets:
    -
    -
    -
- **plug-in approach via kernel density estimation**
Estimation of level sets

Level sets of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].$$

**Note:** regularity $\rightsquigarrow$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!

- direct estimates
  - minimum volume sets:
    - classical concept; shorth (Lientz, 1970, Andrews et al. 1972)
    - 
    - 
  - plug-in approach via kernel density estimation
Estimation of level sets

Level sets of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ are given by

$$\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].$$

Note: regularity $\nabla$ boundaries of level sets $f^{-1}(\lambda)$ are integral curves!

- direct estimates
  - minimum volume sets:
    - classical concept; shorth (Lientz, 1970, Andrews et al. 1972)
- plug-in approach via kernel density estimation
Estimation of level sets

Level sets of a function \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) are given by

\[
\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].
\]

Note: regularity \( \mapsto \) boundaries of level sets \( f^{-1}(\lambda) \) are integral curves!

- **direct estimates**
  - minimum volume sets:
    - classical concept; shorth (Lientz, 1970, Andrews et al. 1972)
    - volume (length) of MV-sets: generalized quantiles: Grübel (1988); Einmahl and Mason (1992); WP (1997)

- **plug-in approach via kernel density estimation**
Overview | Integral curves | Level set estimation | Inference for modes / modal clustering | Filament

**Estimation of level sets**

Level sets of a function \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) are given by

\[
\Gamma_f(\lambda) = \{ x \in \mathbb{R}^d : f(x) \geq \lambda \} = f^{-1}[\lambda, \infty].
\]

**Note:** regularity \( \Rightarrow \) boundaries of level sets \( f^{-1}(\lambda) \) are integral curves!

- **direct estimates**
  - minimum volume sets:
    - classical concept; shorth (Lientz, 1970, Andrews et al. 1972)
Confidence regions for density level sets
Confidence regions for density level sets

\(X_1, \ldots, X_n \sim f.\) Fix \(\lambda > 0\) and \(\gamma \in [0, 1].\)

**Goal:** Find region \(\hat{C}_n\) with \(P(f^{-1}(\lambda) \subset \hat{C}_n) \to \gamma.\)
$X_1, \ldots, X_n \sim f$. Fix $\lambda > 0$ and $\gamma \in [0, 1]$.

**GOAL:** Find region $\hat{C}_n$ with $P(f^{-1}(\lambda) \subset \hat{C}_n) \to \gamma$.

Two different approaches in literature, based on

- vertical variation
- horizontal variation
$X_1, \ldots, X_n \sim f$. Fix $\lambda > 0$ and $\gamma \in [0, 1]$.

**Goal:** Find region $\hat{C}_n$ with $P(f^{-1}(\lambda) \subset \hat{C}_n) \rightarrow \gamma$.

Two different approaches in literature, based on

- vertical variation
- horizontal variation

Both approaches are based on kernel density estimation:

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

$$\Gamma_{\hat{f}}(\lambda) = \{x \in \mathbb{R}^d : \hat{f}_n(x) \geq \lambda\}.$$
Vertical variation

Construct confidence region of the form

$$\hat{C}_n = \Gamma_{bf}(\lambda - \beta n) \setminus \Gamma_{bf}(\lambda + \beta n) = \hat{f}^{-1}(\lambda) - 1_n \left[ \lambda - \beta n, \lambda + \beta n \right].$$

Question:
How to find an appropriate value of $\beta n$?

Idea:
Use $\gamma$-quantile of distribution of $\sup x \in f^{-1}(\lambda) | \hat{f}_n(x) - f(x)$, because $f^{-1}(\lambda) \subset \hat{f}^{-1}_n \left[ \lambda - \beta n, \lambda + \beta n \right] \iff -\beta n \leq \hat{f}_n(x) - \lambda \leq \beta n$ for all $x \in f^{-1}(\lambda)$.
Vertical variation

Construct confidence region of the form

\[ \hat{C}_n = \Gamma_{\hat{f}}(\lambda - \beta_n) \setminus \Gamma_{\hat{f}}(\lambda + \beta_n) = \hat{f}_n^{-1}[\lambda - \beta_n, \lambda + \beta_n]. \]
Vertical variation

Construct confidence region of the form

\[ \hat{C}_n = \Gamma_{\hat{f}}(\lambda - \beta_n) \setminus \Gamma_{\hat{f}}(\lambda + \beta_n) = \hat{f}_n^{-1}[\lambda - \beta_n, \lambda + \beta_n]. \]

**Question:** How to find an appropriate value of \( \beta_n \)?

**Idea:** Use \( \gamma \)-quantile of distribution of \( \sup_{x \in f^{-1}(\lambda)} |\hat{f}_n(x) - f(x)| \),

Vertical variation

Construct confidence region of the form

\[ \hat{C}_n = \Gamma(\hat{f}(\lambda - \beta_n)) \setminus \Gamma(\hat{f}(\lambda + \beta_n)) = \hat{f}_n^{-1}[\lambda - \beta_n, \lambda + \beta_n]. \]

**Question:** How to find an appropriate value of \( \beta_n \)?

**Idea:** Use \( \gamma \)-quantile of distribution of \( \sup_{x \in f^{-1}(\lambda)} |\hat{f}_n(x) - f(x)| \), because

\[ f^{-1}(\lambda) \subset \hat{f}_n^{-1}[\lambda - \beta_n, \lambda + \beta_n] \quad \Leftrightarrow \quad -\beta_n \leq \hat{f}_n(x) - \lambda \leq \beta_n \quad \text{for all} \ x \in f^{-1}(\lambda) \]
One might consider two approximations of distribution of $\sup_{x \in f^{-1}(\lambda)} |\hat{f}_n(x) - f(x)|$:

- bootstrap
- large sample

(cf. Qiao and WP, 2015).
Vertical variation

One might consider two approximations of distribution of
\( \sup_{x \in f^{-1}(\lambda)} |\hat{f}_n(x) - f(x)|: \)

- bootstrap
- large sample

(cf. Qiao and WP, 2015).

Mammen and WP (2013) use related approach and construct bootstrap approximation of \( \sup_{x \in f^{-1}[\lambda - b_n, \lambda + b_n]} |\hat{f}_n(x) - f(x)|, \) for appropriately chosen sequence \( b_n \to 0. \)
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point $x \in f^{-1}(\lambda)$,

$$\frac{|\hat{f}_n(x) - f(x)|}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad} f(x)\|,$$
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point $x \in f^{-1}(\lambda)$,

$$\frac{|\hat{f}_n(x) - f(x)|}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad} f(x)\|,$$

where $d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y)$.
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point \( x \in f^{-1}(\lambda) \),

\[
\frac{|\hat{f}_n(x) - f(x)|}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad} f(x)\|,
\]

where \( d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y) \). In other words,

\[
\frac{|\hat{f}_n(x) - f(x)|}{\|\text{grad} f(x)\|} \approx d(x, \hat{f}_n^{-1}(\lambda))
\]
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point \( x \in f^{-1}(\lambda) \),

\[
\frac{\hat{f}_n(x) - f(x)}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad} f(x)\|, 
\]

where \( d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y) \). In other words,

\[
\frac{\hat{f}_n(x) - f(x)}{\|\text{grad} f(x)\|} \approx d(x, \hat{f}_n^{-1}(\lambda))
\]

Uniform control of \( \frac{\hat{f}_n(x) - f(x)}{\|\text{grad} f(x)\|} \)
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point $x \in f^{-1}(\lambda)$,

$$\frac{\hat{f}_n(x) - f(x)}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad}f(x)\|,$$

where $d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y)$. In other words,

$$\frac{\hat{f}_n(x) - f(x)}{\|\text{grad}f(x)\|} \approx d(x, \hat{f}_n^{-1}(\lambda))$$

Uniform control of $\frac{\hat{f}_n(x) - f(x)}{\|\text{grad}f(x)\|} \rightsquigarrow$ control of Hausdorff distance
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point $x \in f^{-1}(\lambda)$,

$$\frac{\hat{f}_n(x) - f(x)}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \|\text{grad}f(x)\|,$$

where $d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y)$. In other words,

$$\frac{\hat{f}_n(x) - f(x)}{\|\text{grad}f(x)\|} \approx d(x, \hat{f}_n^{-1}(\lambda))$$

Uniform control of $\frac{\hat{f}_n(x) - f(x)}{\|\text{grad}f(x)\|} \leadsto$ control of Hausdorff distance $\leadsto$ confidence regions by using quantiles of Hausdorff distance.
Horizontal variation

Chen et al. (2015a), Qiao and WP (2015)

Simple relation: At a given point $x \in f^{-1}(\lambda)$,

$$\frac{\left| \hat{f}_n(x) - f(x) \right|}{d(x, \hat{f}_n^{-1}(\lambda))} \approx \| \text{grad} f(x) \|,$$

where $d(x, \hat{f}_n^{-1}(\lambda)) = \inf_{y \in \hat{f}_n^{-1}(\lambda)} d(x, y)$. In other words,

$$\frac{\left| \hat{f}_n(x) - f(x) \right|}{\| \text{grad} f(x) \|} \approx d(x, \hat{f}_n^{-1}(\lambda))$$

Uniform control of $\frac{\left| \hat{f}_n(x) - f(x) \right|}{\| \text{grad} f(x) \|}$ $\leadsto$ control of Hausdorff distance $\leadsto$ confidence regions by using quantiles of Hausdorff distance

$$d_H(f^{-1}(\lambda), \hat{f}_n^{-1}(\lambda)) = \max \left[ \sup_{x \in f^{-1}(\lambda)} d(x, \hat{f}_n^{-1}(\lambda)), \sup_{x \in \hat{f}_n^{-1}(\lambda)} d(x, f^{-1}(\lambda)) \right].$$
Inference for modes / modal clustering

- (local) level sets $\Rightarrow$ modal regions
- geometric properties of level sets $\Rightarrow$ number of modes
- geometric properties of level sets $\Rightarrow$ capture features of density $\Rightarrow$ visualization (level set tree)
- excess mass approach, Hartigan's dip $\Rightarrow$ testing for modes
- integral curves driven by gradient fields $\Rightarrow$ modal clustering
- existence of antimodes $\Rightarrow$ testing for modes
Inference for modes / modal clustering

- (local) level sets $\leadsto$ modal regions
Inference for modes / modal clustering

- (local) level sets \( \rightsquigarrow \) modal regions
- geometric properties of level sets \( \rightsquigarrow \) number of modes
Inference for modes / modal clustering

- (local) level sets \( \rightsquigarrow \) modal regions
- geometric properties of level sets \( \rightsquigarrow \) number of modes
- geometric properties of level sets
  \( \rightsquigarrow \) capture features of density \( \rightsquigarrow \) visualization (level set tree)

Inference for modes / modal clustering

- (local) level sets $\leadsto$ modal regions
- geometric properties of level sets $\leadsto$ number of modes
- geometric properties of level sets
  $\leadsto$ capture features of density $\leadsto$ visualization (level set tree)
- excess mass approach, Hartigan’s dip $\leadsto$ testing for modes
Inference for modes / modal clustering

- (local) level sets $\rightsquigarrow$ modal regions
- geometric properties of level sets $\rightsquigarrow$ number of modes
- geometric properties of level sets $\rightsquigarrow$ capture features of density $\rightsquigarrow$ visualization (level set tree)
- excess mass approach, Hartigan’s dip $\rightsquigarrow$ testing for modes
- integral curves driven by gradient fields $\rightsquigarrow$ modal clustering
Inference for modes / modal clustering

- (local) level sets $\leadsto$ modal regions
- geometric properties of level sets $\leadsto$ number of modes
- geometric properties of level sets  
  $\leadsto$ capture features of density $\leadsto$ visualization (level set tree)
- excess mass approach, Hartigan’s dip $\leadsto$ testing for modes
- integral curves driven by gradient fields $\leadsto$ modal clustering
- existence of antimodes $\leadsto$ testing for modes

Estimation and inference for persistent homology: TDA
Estimation and inference for persistent homology: TDA

**Target:** topological properties of supports and more general of level sets (Bobrowski et al. 2015); measured by ranks of homology groups
Estimation and inference for persistent homology: TDA

**Target:** topological properties of supports and more general of level sets (Bobrowski et al. 2015); measured by ranks of homology groups

Estimate homologies of a filtration based on simplicial complexes built on data (filtration based on level sets); Betti numbers (often: \( \beta_0 \) - number of connected components)
**Target:** topological properties of supports and more general of level sets (Bobrowski et al. 2015); measured by ranks of homology groups

Estimate homologies of a filtration based on simplicial complexes built on data (filtration based on level sets); Betti numbers (often: $\beta_0$ - number of connected components)

Distinguish between signal and noise by using *persistency*. 
Target: topological properties of supports and more general of level sets (Bobrowski et al. 2015); measured by ranks of homology groups.

Estimate homologies of a filtration based on simplicial complexes built on data (filtration based on level sets); Betti numbers (often: $\beta_0$ - number of connected components).

Distinguish between signal and noise by using persistency.

Bubenik and Kim (2006); Balakrishnan et al. (2011, 2013); Chazal et al. (2014a,b), Fasy et al. (2013); Bauer et al. (2014), Bobrowski et al. (2015), Boissonat et al. (2015), ...
Filament or ridge line estimation

A point is said to be a ridge point or a filament point if \( \lambda_2 < 0 \) where \( \lambda_1 > \lambda_2 \) are the two eigenvalues of the Hessian \( H(x) \).

A filament consists of filament points and is an integral curve of the gradient.

Let \( V(x) \) denote the second eigenvector of Hessian \( H(x) \).

On the filament, either \( \nabla f = 0 \) or \( \nabla f \parallel V \perp \), i.e. \( \langle \nabla f, V \rangle = 0 \).
Filament or ridge line estimation

- What is a filament?

**Definition:** A point is said to be a ridge point or a filament point if $\lambda_2 < 0$ and $\nabla f = \lambda_1 \nabla f$ where $\lambda_1 > \lambda_2$ are the two eigenvalues of the Hessian $H(x)$.

A filament consists of filament points and is an integral curve of the gradient. Let $V(x)$ denote the second eigenvector of Hessian $H$. On the filament, either $\nabla f = 0$ or $\nabla f \parallel V \perp$, i.e. $\langle \nabla f, V \rangle = 0$. 
What is a filament?

**Definition:** A point is said to be a ridge point or a *filament point* if

\[ \lambda_2 < 0 \]

\[ H \nabla f = \lambda_1 \nabla f \]

where \( \lambda_1 > \lambda_2 \) are the two eigenvalues of the Hessian \( H(x) \).

A *filament* consists of filament points and is an integral curve of the gradient.
Filament or ridge line estimation

• What is a filament?

**Definition:** A point is said to be a ridge point or a *filament point* if

\[
\lambda_2 < 0 \\
H \nabla f = \lambda_1 \nabla f
\]

where \( \lambda_1 > \lambda_2 \) are the two eigenvalues of the Hessian \( H(x) \).

A *filament* consists of filament points and is an integral curve of the gradient.

Let \( V(x) \) denote *second* eigenvector of Hessian \( H \).

On the filament, either \( \nabla f = 0 \) or \( \nabla f \parallel V^\perp \), i.e. \( \langle \nabla f, V \rangle = 0 \).
From Chen et al. (2014).
Geometric idea

\[ \langle \nabla f(x), V(x) \rangle \]

and

\[ V(x)^T \nabla^2 f(x) V(x) = \lambda_2(x) \| V(x) \|^2 \]

are first and second order directional derivative of \( f(x) \) along \( V(x) \). Thus filament points are local mode of \( f(x) \) along the direction \( V(x) \).

**Geometric idea:** Consider vector field generated by the second eigenvectors \( V(x) \) of the Hessian \( H \) of \( f \).

- A ridge point corresponds to a local mode of \( f \) along the path of the corresponding integral curve for the vector field generated by \( V(x) \).
Overview

Integral curves

Level set estimation

Inference for modes / modal clustering

Application areas

seismology: analysis of fault lines
analysing road or river networks
cosmology: cosmic web
medical imaging: e.g. blood vessels network

Nonparametric Inference for Geometric Objects
Application areas

- seismology: analysis of fault lines
- analysing road or river networks
- cosmology: cosmic web
- medical imaging: e.g. blood vessels network
Related literature

- Minimum spanning tree, Barrow et al. (1985)
- Candy model, Stoica et al. (2005)
- Principal curves; Hastie and Stuetzle (1989), Kegl et al. (2000), Sandilya and Kulkarni (2002), and Smola et al. (2001)
- Local principal curve; Einbeck, Tutz and Evers (2005), Einbeck, Evers, and Bailer-Jones (2007)
- Skeleton; Novikov et al. (2006)
- Nonparametric penalized maximum likelihood; Tibshirani (1992)
- Beamlets; Donoho and Huo (2002), Arias-Castro et al. (2006)
- Feature detection in point clouds (Engineering/CS): e.g. Weber et al. (2006), Daniels et al. (2007) . . .
Related other concepts

Conceptually related to other statistical concepts:

- mode hunting
Related other concepts

Conceptually related to other statistical concepts:

- mode hunting
- integral curve estimation
Related other concepts

Conceptually related to other statistical concepts:

- mode hunting
- integral curve estimation
- tracking fault lines (Hall and Rau, 2000);
Related other concepts

Conceptually related to other statistical concepts:

- mode hunting
- integral curve estimation
- tracking fault lines (Hall and Rau, 2000);
- principal curves (Hastie and Stuetzle, 1989, Sandilya and Kukarni, 2002);
Related other concepts

Conceptually related to other statistical concepts:

- mode hunting
- integral curve estimation
- tracking fault lines (Hall and Rau, 2000);
- principal curves (Hastie and Stuetzle, 1989, Sandilya and Kukarni, 2002);
- beamlets, curvelets, ridgelets . . . (Candés 1999; Candés and Donoho, 1999; Donoho and Huo, 2002).
Statistical literature

Above literature: No statistical quantifications.
Statistical literature

Above literature: No statistical quantifications.

Statistical literature:

- Cheng, Hall and Hartigan (2004);
- Arias-Castro, Donoho, and Huo (2006);
- Genovese et al. (2009, 2012, 2014);
- Chen et al. (2013, 2014)
- Qiao and WP (2015)
Genovese et al. (2009): Path density

- $\mathcal{X}_{x_0}(t)$ integral curve of gradient field; starting at $x_0$

$\mathcal{V}(A) = \{x_0 : \mathcal{X}_{x_0} \cap A \neq \emptyset\}$
Genovese et al. (2009): Path density

• $\mathcal{X}_{x_0}(t)$ integral curve of gradient field; starting at $x_0$

$\mathcal{V}(A) = \{x_0 : \mathcal{X}_{x_0} \cap A \neq \emptyset\}$ (purple area)
Genovese et al. (2009): Path density

- $\mathcal{X}_{x_0}(t)$ integral curve of gradient field; starting at $x_0$
- $\mathcal{V}(A) = \{x_0 : \mathcal{X}_{x_0} \cap A \neq \emptyset\}$ (purple area)

- Path measure $\pi(A) = \int_{\mathcal{V}(A)} g(x) dx$

Nonparametric Inference for Geometric Objects
Genovese et al. (2009): Path density

• $\mathcal{X}_{x_0}(t)$ integral curve of gradient field; starting at $x_0$

$\mathcal{V}(A) = \{x_0 : \mathcal{X}_{x_0} \cap A \neq \emptyset\}$ (purple area)

• Path measure $\pi(A) = \int_{\mathcal{V}(A)} g(x) dx$

• Path density $p$:

$$p(x) = \lim_{r \to 0} \frac{\pi(B(x, r))}{r} = \begin{cases} \infty & \text{for } x \text{ on filament} \\ < \infty & \text{for } x \text{ off filament} \end{cases}$$
Genovese et al. (2009): Path density

- $\mathcal{X}_{x_0}(t)$ integral curve of gradient field; starting at $x_0$
  
  $\mathcal{V}(A) = \{x_0 : \mathcal{X}_{x_0} \cap A \neq \emptyset\}$ (purple area)

- Path measure $\pi(A) = \int_{\mathcal{V}(A)} g(x) \, dx$

- Path density $p$:
  
  $p(x) = \lim_{r \to 0} \frac{\pi(B(x, r))}{r} = \begin{cases} \infty & \text{for } x \text{ on filament} \\ < \infty & \text{for } x \text{ off filament} \end{cases}$

- Consider level set of estimated path density as ‘estimator’.
Path density

Galaxy distribution in a slice

Data source: www.mpa-garching.mpg.de
A different model

Filament

Genovese et al. (2012a) consider the model

\[ Y_i = f(U_i) + \epsilon_i \]

with \( U_i \) drawn from a distribution on \([0, 1]\) \( \epsilon_i \) independent such that support(\( Y \)) = \( M \oplus \sigma \). Minimax rates for estimating the filament using Hausdorff distance are derived in this model.

Genovese et al. (2012b) consider the medial axis of the level set to estimate the filament.
A different model

Filament: $\mathcal{M} = \{f(x) : x \in [0, 1]\} \subset \mathbb{R}^d$. 

Genovese et al. (2012a) consider the model $Y_i = f(U_i) + \epsilon$ with $U_i$ drawn from a distribution on $[0, 1]$ and $\epsilon$ independent such that $\text{support}(Y) = \mathcal{M} \oplus \sigma$. Minimax rates for estimating the filament $f$ using Hausdorff distance are derived in this model. Genovese et al. (2012b) consider the medial axis of the level set to estimate the filament.
A different model

Filament: $\mathcal{M} = \{f(x) : x \in [0, 1]\} \subset \mathbb{R}^d$. Genovese et al. (2012a) consider the model

$$Y_i = f(U_i) + \epsilon$$
A different model

Filament: \( \mathcal{M} = \{ f(x) : x \in [0, 1] \} \subset \mathbb{R}^d \). Genovese et al. (2012a) consider the model

\[
Y_i = f(U_i) + \epsilon
\]

with

- \( U_i \) drawn from a distribution on [0, 1]
A different model

Filament: \( \mathcal{M} = \{ f(x) : x \in [0, 1] \} \subset \mathbb{R}^d \). Genovese et al. (2012a) consider the model

\[ Y_i = f(U_i) + \epsilon \]

with

- \( U_i \) drawn from a distribution on \([0, 1]\)
- \( \epsilon_i \) independent such that \( \text{support}(Y) = \mathcal{M} \oplus \sigma \)
A different model

Filament: $\mathcal{M} = \{f(x) : x \in [0, 1]\} \subset \mathbb{R}^d$. Genovese et al. (2012a) consider the model

$$Y_i = f(U_i) + \epsilon$$

with

- $U_i$ drawn from a distribution on $[0, 1]$
- $\epsilon_i$ independent such that $\text{support}(Y) = \mathcal{M} \oplus \sigma$

Minimax rates for estimating the filament $f$ using Hausdorff distance are derived in this model.
A different model

Filament: $\mathcal{M} = \{f(x) : x \in [0, 1]\} \subset \mathbb{R}^d$. Genovese et al. (2012a) consider the model

$$Y_i = f(U_i) + \epsilon$$

with

- $U_i$ drawn from a distribution on $[0, 1]$
- $\epsilon_i$ independent such that $\text{support}(Y) = \mathcal{M} \oplus \sigma$

Minimax rates for estimating the filament $f$ using Hausdorff distance are derived in this model.

Genovese et al. (2012b) consider the medial axis of the level set to estimate the filament.
3. Some Background on Geometry.

3.1. Thickness and the Medial Axis.

Let \( S \subset \mathbb{R}^2 \) be a compact set. A ball \( B \subset S \) is called medial if

1. \( \text{interior}(B) \cap \partial S = \emptyset \)
2. \( B \cap \partial S \) contains at least 2 points.

The medial axis \( M \equiv M(S) \), shown in Figure 3, is the closure of the set

\[
\{ x \in S : B(x, r) \text{ is medial for some } r > 0 \}
\]

See Dey (2006) and references therein for more information about the properties of the medial axis.

Now we relate the filament to the medial axis. For any three distinct points \( x, y, z \) on \( \Gamma_f \) let \( r(x, y, z) \) be the radius of the circle passing through the three points. Define the thickness of the curve \( \Gamma_f \) (Gonzalez and Maddocks, 1999) denoted \( \Delta \equiv \Delta(f) \) by

\[
\Delta \equiv \Delta(f) = \min_{x,y,z} r(x, y, z)
\]

where the minimum is over all triples of distinct points on \( \Gamma_f \). \( \Delta \) is also called the minimum global radius of curvature, and the normal injectivity radius.
Distribution theory for filament estimation

Qiao and WP (2015)

\[ d = 2 \]
Ridge estimation via bump hunting

We now consider filament estimation based on iid observations from a density $f$ assuming the existence of a ridge line. Recall
We now consider filament estimation based on iid observations from a density $f$ assuming the existence of a ridge line. Recall

**Definition:** A point is said to be a ridge point or a *filament point* if

$$\lambda_2 < 0$$

$$H \nabla f = \lambda_1 \nabla f$$

where $\lambda_1 > \lambda_2$ are the two eigenvalues of the Hessian $H(x)$. $V(x)$ denotes *second* eigenvector of Hessian $H$.

- On the filament, either $\nabla f = 0$ or $\nabla f \parallel V^\perp$, i.e. $\langle \nabla f, V \rangle = 0$.
- Filament points are local mode of $f(x)$ along the direction $V(x)$. 

**Ridge estimation via bump hunting**
Geometric idea

\[ \langle \nabla f(x), V(x) \rangle \]

and

\[ V(x)^T \nabla^2 f(x) V(x) = \lambda_2(x) \| V(x) \|^2 \]

are first and second order directional derivative of \( f(x) \) along \( V(x) \). Thus filament points are local mode of \( f(x) \) along the direction \( V(x) \).

**Geometric idea**: Consider vector field generated by the second eigenvectors \( V(x) \) of the Hessian \( H \) of \( f \).

- A ridge point corresponds to a local mode of \( f \) along the path of the corresponding integral curve for the vector field generated by \( V(x) \).
Geometric idea

\[ \langle \nabla f(x), V(x) \rangle \]

and

\[ V(x)^T \nabla^2 f(x) V(x) = \lambda_2(x) \| V(x) \|^2 \]

are first and second order directional derivative of \( f(x) \) along \( V(x) \). Thus filament points are local mode of \( f(x) \) along the direction \( V(x) \).

**Geometric idea**: Consider vector field generated by the second eigenvectors \( V(x) \) of the Hessian \( H \) of \( f \).

- A ridge point corresponds to a local mode of \( f \) along the path of the corresponding integral curve for the vector field generated by \( V(x) \).

Same idea is used in Chen et al. (2015c).
Some notation

Hessian

\[ H = H(x) = \begin{pmatrix} f_{11}(x) & f_{12}(x) \\ f_{12}(x) & f_{22}(x) \end{pmatrix} \]

Let

\[ V = \begin{pmatrix} f_{11} - f_{22} + f_{12} - \sqrt{(f_{22} - f_{11})^2 + 4f_{12}^2} \\ \frac{1}{2}(f_{22} - f_{11} + f_{12} - 4\sqrt{(f_{22} - f_{11})^2 + 4f_{12}^2}) \end{pmatrix}. \]

then \( V(x) \) is eigenvectors for \( \lambda_2(x) \).
Use kernel density estimator based on $X_1, X_2, \cdots, X_n \overset{iid}{\sim} f$

$$
\hat{f}(x) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right).
$$

The kernel estimator of Hessian is

$$
\hat{H}(x) = \begin{pmatrix}
\hat{f}_{11}(x) & \hat{f}_{12}(x) \\
\hat{f}_{12}(x) & \hat{f}_{22}(x)
\end{pmatrix}
$$

$$
= \frac{1}{nh^4} \sum_{i=1}^{n} \begin{pmatrix}
K_{11}\left(\frac{x-X_i}{h}\right) & K_{12}\left(\frac{x-X_i}{h}\right) \\
K_{12}\left(\frac{x-X_i}{h}\right) & K_{22}\left(\frac{x-X_i}{h}\right)
\end{pmatrix}
$$

with second eigenvalue $\hat{\lambda}_2$ corresponding second eigenvector $\hat{V}(x)$. 
More notation

For each \( x_0 \in \mathcal{G} \) let

- \( \mathcal{X}_{x_0}(t), t \in [0, T] \) integral curve corresponding to vector field \( V(x) \) starting at \( x_0 \);

- \( \theta_{x_0} = \arg \max_{t \in [0, T]} f(\mathcal{X}_{x_0}(t)), \) i.e. \( \mathcal{X}_{x_0}(\theta_{x_0}) \) lies on filament.
More notation

For each $x_0 \in G$ let

- $\mathcal{X}_{x_0}(t), t \in [0, T]$ integral curve corresponding to vector field $V(x)$ starting at $x_0$;
- $\theta_{x_0} = \arg \max_{t \in [0, T]} f(\mathcal{X}_{x_0}(t))$, i.e. $\mathcal{X}_{x_0}(\theta_{x_0})$ lies on filament.

- $\hat{\mathcal{X}}_{x_0}(t), t \in [0, T]$ integral curve corresponding to vector field $\hat{V}(x)$ starting at $x_0$
- $\hat{\theta}_{x_0} = \arg \max_{t \in [0, T]} f(\hat{\mathcal{X}}_{x_0}(t))$, i.e. $\hat{\mathcal{X}}_{x_0}(\hat{\theta}_{x_0})$ lies on filament.
Overview
Integral curves
Level set estimation
Inference for modes / modal clustering

Filament
Nonparametric Inference for Geometric Objects
Mathematical problems

*Integral curve estimation:*

Find asymptotic distribution of (appropriately normalized)
- $\hat{X}_{x_0}(t) - X_{x_0}(t)$.
Mathematical problems

**Integral curve estimation:**

Find asymptotic distribution of (appropriately normalized)

- $\hat{X}_{x_0}(t) - X_{x_0}(t)$.

**Filament estimation:**

Find asymptotic distribution of (appropriately normalized)

- $\hat{X}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0})$. 
Mathematical problems

**Integral curve estimation:**

Find asymptotic distribution of (appropriately normalized)

- $\hat{\mathcal{X}}_{x_0}(t) - X_{x_0}(t)$.

**Filament estimation:**

Find asymptotic distribution of (appropriately normalized)

- $\hat{\mathcal{X}}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0})$.

- $\sup_{x_0 \in G} |\hat{\mathcal{X}}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0})|$, $G \subset \mathbb{R}^2$, compact.
Mathematical problems

**Integral curve estimation:**

Find asymptotic distribution of (appropriately normalized)

- \( \hat{X}_{x_0}(t) - X_{x_0}(t) \).

**Filament estimation:**

Find asymptotic distribution of (appropriately normalized)

- \( \hat{X}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0}) \).

- \( \sup_{x_0 \in G} |\hat{X}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0})|, \ G \subset \mathbb{R}^2, \text{ compact.} \)

\( \sim \) involves finding limit of the distribution of the supremum over (increasing) manifolds of a sequence non-stationary Gaussian process.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$. 
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.

(F1) $f$ is four times continuously differentiable.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.

(F1) $f$ is four times continuously differentiable.

(F2) Eigenvalues of Hessian are different.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.

(F1) $f$ is four times continuously differentiable.

(F2) Eigenvalues of Hessian are different.

(F3) Norm of second eigenvectors $\|V(x)\|$ of Hessian is bounded away from zero.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.

(F1) $f$ is four times continuously differentiable.

(F2) Eigenvalues of Hessian are different.

(F3) Norm of second eigenvectors $\|V(x)\|$ of Hessian is bounded away from zero.

(F4) For each $x \in \mathcal{L}$, $V(x)$ is not orthogonal to the normal direction to the filament.
Estimation of integral curves: Assumptions

Let $\mathcal{L}$ denote the ‘true’ filament lying in a set $\mathcal{H} \subset \mathbb{R}^2$.

(A1) $\mathcal{L}$ is a compact smooth filament with bounded curvature.

(F1) $f$ is four times continuously differentiable.

(F2) Eigenvalues of Hessian are different.

(F3) Norm of second eigenvectors $\|V(x)\|$ of Hessian is bounded away from zero.

(F4) For each $x \in \mathcal{L}$, $V(x)$ is not orthogonal to the normal direction to the filament.

(F5) $\{x : \lambda_2(x) = 0, \langle \nabla f(x), V(x) \rangle = 0\} = \emptyset$
Estimation of integral curves: Assumptions

(K1) The kernel $K$ is a symmetric probability density function with support $\{x : \|x\| < 1\}$. All of its first to fourth order partial derivatives are bounded and $\int_{\mathbb{R}^2} K(x)xx^T dx = \mu_2(K)\text{Id}$ with $\mu_2(K) < \infty$.

(K2) $R(d^2K) < \infty$, where for any function $g : \mathbb{R}^2 \mapsto \mathbb{R}^3$, $R(g) \equiv \int_{\mathbb{R}^2} g(x)g(x)^T dx$.

(K3) $\int [K^{(3,0)}(z)]^2 dz \neq \int [K^{(1,2)}(z)]^2 dz$.

(H1) As $n \to \infty$, $h_n \downarrow 0$, $nh_n^8/(\log n)^3 \to \infty$, $nh_n^9 \to \beta$, $\beta \geq 0$. 
Theorem

Under above assumptions and for each $T > 0$, the sequence of stochastic process

$$\sqrt{n}h^5 (\hat{X}_{x_0}(t) - X_{x_0}(t)), \quad 0 \leq t \leq T$$

converges weakly in $C([0, \ T], \mathbb{R}^2)$ to a Gaussian process as $n \to \infty$.

The proof is an adaptation of Koltchinskii et al. (2007).
Estimation of integral curves

**Theorem**

*Under above assumptions and for each $T > 0$, the sequence of stochastic process*

$$\sqrt{nh^5} (\hat{X}_{x_0}(t) - X_{x_0}(t)), \quad 0 \leq t \leq T$$

*converges weakly in $C([0, T], \mathbb{R}^2)$ to a Gaussian process as $n \to \infty$.*

The proof is an adaptation of Koltchinskii et al. (2007).

**Theorem**

*Under above assumptions, for each $T > 0$ as $n \to \infty$,*

$$\sup_{x_0 \in G, t \in [0, T]} \| \hat{X}_{x_0}(t) - X_{x_0}(t) \| = O_p \left( \frac{\log n}{\sqrt{nh^5}} \right)$$
Some heuristics

- Estimating 1st derivatives: rate $O_P(1/\sqrt{nh^{d+2}}) = O_P(\sqrt{1/nh^4})$. 


- Integral curves: $X_{x_0}(t) = x_0 + \int_0^t V(X_{x_0}(s)) \, ds$; one-dim. integral of function of second derivatives $\Rightarrow$ gain one power of $h$: $O_P(1/\sqrt{nh^5})$.

- Omitting index $x_0$:
  
- $\hat{X}^{(\hat{\theta})} - X^{(\theta)} = \left[ \hat{X}^{(\hat{\theta})} - X^{(\hat{\theta})} \right] + \left[ X^{(\hat{\theta})} - X^{(\theta)} \right] = O_P(1/\sqrt{nh^5}) + O_P \left( V(X^{(\theta)})(b_\theta - \theta) \right)$.

- $\hat{\theta} - \theta = O_P(1/\sqrt{nh^6})$ if $\nabla f(X^{(\theta)}) \neq 0$, and $\hat{\theta} - \theta = O_P(1/\sqrt{nh^5})$ if $\nabla f(X^{(\theta)}) = 0$.
Some heuristics

- Estimating 1st derivatives: rate $O_P(1/\sqrt{nh^{d+2}}) = O_P(\sqrt{1/nh^4})$.
Some heuristics

- Estimating 1st derivatives: rate $O_P(1/\sqrt{nh^{d+2}}) = O_P(\sqrt{1/nh^4})$.
- Integral curves: $\mathcal{X}_{x_0}(t) = x_0 + \int_0^t V(\mathcal{X}_{x_0}(s)) \, ds$; one-dim. integral of function of second derivatives
  $\rightsquigarrow$ gain one power of $h$: $O_P(1/\sqrt{nh^5})$
Some heuristics

- Estimating 1st derivatives: rate $O_P(1/\sqrt{nh^{d+2}}) = O_P(\sqrt{1/nh^4})$.
- Integral curves: $\mathcal{X}_{x_0}(t) = x_0 + \int_0^t V(\mathcal{X}_{x_0}(s)) \, ds$; one-dim. integral of function of second derivatives $\leadsto$ gain one power of $h$: $O_P(1/\sqrt{nh^5})$

Omitting index $x_0$:

- $\hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) = \left[ \hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\hat{\theta}) \right] + \left[ \mathcal{X}(\hat{\theta}) - \mathcal{X}(\theta) \right]$.  

  $\underbrace{O_P(1/\sqrt{nh^5})}_{\text{approx.}} + \underbrace{O_P\left( V(\mathcal{X}(\theta))(\hat{\theta}-\theta) \right)}_{\text{approx.}}$
Some heuristics

- Estimating 1st derivatives: rate $O_P\left(1/\sqrt{nh^{d+2}}\right) = O_P\left(\sqrt{1/nh^4}\right)$.
- Estimating 2nd derivatives: rate $O_P\left(1/\sqrt{nh^{d+4}}\right) = O_P\left(1/\sqrt{nh^6}\right)$.
- Integral curves: $\mathcal{X}_{x_0}(t) = x_0 + \int_0^t \nabla(\mathcal{X}_{x_0}(s)) \, ds$;
  one-dim. integral of function of second derivatives
  $\leadsto$ gain one power of $h$: $O_P\left(1/\sqrt{nh^5}\right)$

Omitting index $x_0$:

- $\hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) = \underbrace{\hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\hat{\theta})}_{O_P\left(1/\sqrt{nh^5}\right)} + \underbrace{\mathcal{X}(\hat{\theta}) - \mathcal{X}(\theta)\right)}_{O_P\left(\nabla(\mathcal{X}(\theta))(\hat{\theta} - \theta)\right)}$

- $\hat{\theta} - \theta = O_P\left(1/\sqrt{nh^6}\right)$ if $\nabla f(\mathcal{X}(\theta)) \neq 0$, and
  $\hat{\theta} - \theta = O_P\left(1/\sqrt{nh^5}\right)$ if $\nabla f(\mathcal{X}(\theta)) = 0$
A heuristic argument for why estimation of filaments is easier when \( \nabla f(x) = 0 \) at the filament:
A heuristic argument for why estimation of filaments is easier when $\nabla f(x) = 0$ at the filament:

Recall: on filament $H(x) \nabla f(x) = \lambda_1(x) \nabla f(x)$.

Thus, when replacing $H$ and $f$ by their estimates, then, if $\nabla f(x) = 0$, this equality holds approximately if we can estimate first derivatives well. The estimation of second derivatives is not too important. Thus the rates are driven by how well we can estimate first derivatives as opposed to second derivatives, and the former is easier (faster rates).
Filament estimation: more assumptions

(F6) For any $x_0 \in \mathcal{H}$ with $x_0 \prec \mathcal{L}$, $\theta_{x_0}$ exists and
\[\sup_{x_0 \in \mathcal{H}, x_0 \prec \mathcal{L}} T_{x_0} < \infty.\]

(F7) $\nabla \langle \nabla f(x) V(x) \rangle \neq 0$ for all $x \in \mathcal{L}$

(F8) $\{x \in \mathcal{H} : \lambda_2(x) = 0, \nabla f(x) V(x) = 0\} = \emptyset$. 
Theorem

Assume that above assumptions hold, \( nh^9 \to \beta \geq 0, h_n \to 0 \). Then for any fixed starting point \( x_0 \):

(a) \[ \sqrt{nh^6} \langle V(\hat{\mathcal{X}}(\theta)), \hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) \rangle \to_{\mathcal{D}} N(0, \sigma_1^2) \],

(b) If \( \nabla f(\mathcal{X}(\theta)) = 0 \), then \( \sqrt{nh^5} \langle V(\mathcal{X}(\theta)), \hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) \rangle \to_{\mathcal{D}} N(\mu_1, \sigma_2^2) \),

\( \sqrt{nh^5} \langle V(\mathcal{X}(\theta)) \perp \mathcal{X}(\theta) \rangle \to_{\mathcal{D}} N(\mu_2, \sigma_3^2) \).
Filament estimation: Pointwise convergence

**Theorem**

Assume that above assumptions hold, \( nh^9 \to \beta \geq 0, \ h_n \to 0 \). Then for any fixed starting point \( x_0 \):

\[(a) \quad \sqrt{nh^6} \left\langle V(\mathcal{X}(\theta)), \ \hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) \right\rangle \to_{\mathcal{D}} N(0, \sigma_1^2),\]

\[\sqrt{nh^5} \left\langle V(\mathcal{X}(\theta))^\perp, \ \hat{\mathcal{X}}(\hat{\theta}) - \mathcal{X}(\theta) \right\rangle \to_{\mathcal{D}} N(0, \sigma_2^2).\]
Theorem

Assume that above assumptions hold, \( nh^9 \to \beta \geq 0, h_n \to 0 \). Then for any fixed starting point \( x_0 \):

(a) \[
\sqrt{nh^6} \left\langle \nabla (\hat{X}(\theta)), \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D \mathcal{N}(0, \sigma_1^2),
\]
\[
\sqrt{nh^5} \left\langle \nabla (\hat{X}(\theta)) \perp, \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D \mathcal{N}(0, \sigma_2^2).
\]

(b) If \( \nabla f(X(\theta)) = 0 \), then
\[
\sqrt{nh^5} \left\langle \nabla (X(\theta)), \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D \mathcal{N}(\mu_1, \sigma_3^2),
\]
Theorem

Assume that above assumptions hold, $nh^9 \to \beta \geq 0$, $h_n \to 0$. Then for any fixed starting point $x_0$:

(a) $\sqrt{nh^6} \left\langle V(X(\theta)), \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D N(0, \sigma_1^2)$,

$\sqrt{nh^5} \left\langle V(X(\theta)) \perp, \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D N(0, \sigma_2^2)$.

(b) If $\nabla f(X(\theta)) = 0$, then

$\sqrt{nh^5} \left\langle V(X(\theta)), \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D N(\mu_1, \sigma_3^2)$,

$\sqrt{nh^5} \left\langle V(X(\theta)) \perp, \hat{X}(\hat{\theta}) - X(\theta) \right\rangle \to_D N(\mu_2, \sigma_4^2)$. 

Filament estimation: Pointwise convergence
Theorem

Assume that above assumptions hold, $nh^9 \to \beta \geq 0$, $h \to 0$. Then for any fixed starting point $x_0$

$$\sqrt{nh^6}[\hat{X}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0})] \to Z(X_{x_0}(\theta_{x_0})) V(X_{x_0}(\theta_{x_0})),$$

where $Z(X_{x_0}(\theta_{x_0}))$ is a mean zero normal random variable.
Theorem

Suppose that the assumptions of the above Theorem hold, and in addition assume that $\nabla f(\chi_{x_0}(\theta_{x_0})) = 0$. Then there exists $\mu(x_0) \in \mathbb{R}^2$ and $\Sigma(x_0) \in \mathbb{R}^{2 \times 2}$ such that

$$\sqrt{nh^5} [\hat{\chi}_{x_0}(\hat{\theta}_{x_0}) - \chi_{x_0}(\theta_{x_0})] \to \mathcal{N}(\mu(x_0), \Sigma(x_0)).$$
Theorem

Under the above assumptions there exists a constant $c > 0$ and a function $b(x)$, both depending on $f$ and the kernel $K$, such that for any fixed $z$, we have

$$\lim_{n \to \infty} P \left( \sup_{x_0 \in \mathcal{G}} \left\| b(\mathcal{X}_{x_0}(\theta_{x_0})) \sqrt{nh^6} \left( \hat{\mathcal{X}}_{x_0}(\hat{\theta}_{x_0}) - \mathcal{X}_{x_0}(\theta_{x_0}) \right) \right\| < B_h(z) \right) = \exp\{-2 \exp\{-z\}\},$$

where $B_h(z) = \sqrt{2 \log h^{-1} + \frac{1}{\sqrt{2 \log h^{-1}}} \left[ z + c \right]}$ and $\mathcal{G}$ is some properly chosen region of starting points.
Uniform convergence

**Theorem**

*Under the above assumptions there exists a constant \( c > 0 \) and a function \( b(x) \), both depending on \( f \) and the kernel \( K \), such that for any fixed \( z \), we have*

\[
\lim_{n \to \infty} P \left( \sup_{x_0 \in G} \left\| b(x_0(\theta_{x_0})) \sqrt{n} h^6 \left( \hat{X}_{x_0}(\hat{\theta}_{x_0}) - X_{x_0}(\theta_{x_0}) \right) \right\| < B_h(z) \right) = \exp\{-2 \exp\{-z\}\},
\]

*where \( B_h(z) = \sqrt{2 \log h^{-1}} + \frac{1}{\sqrt{2 \log h^{-1}}} \left[ z + c \right] \) and \( G \) is some properly chosen region of starting points.*

First use ideas similar to Bickel and Rosenblatt (1973). Main ingredient to the proof is a generalization of a theorem by Mikhaleva and Piterbarg (1996).
Generalization of a theorem by Mikhaleva and Piterbarg

**Definition (Local equi-$D_t$-stationarity)**

Let $X_h(t), t \in \mathcal{G} \subset \mathbb{R}^2$ be a class of process indexed by $h \in \mathbb{H}$ with covariance function $r_h(t_1, t_2)$. The sequence $X_h(t)$ is *locally equi-$D_t^h$-stationary*, if for any $\epsilon > 0$ there exists a positive $\delta(\epsilon)$ independent of $h$ such that for any $s \in \mathcal{G}$ one can find a non-degenerated matrix $D_s^h$ such that

$$1 - (1 + \epsilon)||D_s^h(t_1 - t_2)||^2 \leq r_h(t_1, t_2) \leq 1 - (1 - \epsilon)||D_s^h(t_1 - t_2)||^2$$

provided $||t_1 - s|| < \delta(\epsilon)$ and $||t_2 - s|| < \delta(\epsilon)$ where $|| \cdot ||$ is Frobenius norm.
Overview
Integral curves
Level set estimation
Inference for modes / modal clustering

Generalization of a theorem by Mikhaleva and Piterbarg

**Theorem**

Let $\mathcal{M}_1 \subset \mathcal{H}$ be a smooth compact 1-dimensional manifold with bounded curvature, $\{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\}$ a class of centered, locally $D^h_t$-stationary Gaussian fields. Under below assumptions, there exists $M > 0$ such that with $x_h(z) = (2 \log \frac{1}{h})^{\frac{1}{2}} (1 + \frac{M+z}{2 \log \frac{1}{h}})$ we have

$$\lim_{h \to 0} P\{ \sup_{t \in \mathcal{M}_h} |X_h(t)| \leq x_h(z) \} = \exp\{-2 \exp\{-z\}\}$$

where $\mathcal{M}_h = \frac{\mathcal{M}_1}{h}$. 

Nonparametric Inference for Geometric Objects
Generalization of a theorem by Mikhaleva and Piterbarg

**ASSUMPTIONS:**
\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact 1-dimensional manifold with bounded curvature.
Generalization of a theorem by Mikhaleva and Piterbarg

**ASSUMPTIONS:**
\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact 1-dimensional manifold with bounded curvature. \( \{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\} \) a class of centered, locally \( D_t^h \)-stationary Gaussian fields with...
Generalization of a theorem by Mikhaleva and Piterbarg

-Assumptions:
\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact 1-dimensional manifold with bounded curvature. \( \{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\} \) a class of centered, locally \( D^h_t \)-stationary Gaussian fields with
- \( D^h_t \) positive definite and \( (t, h) \rightarrow D^h_t \), continuous;
Assumptions:
\[ \mathcal{M}_1 \subset \mathcal{H} \text{ smooth compact 1-dimensional manifold with bounded curvature.} \]
\[ \{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\} \text{ a class of centered, locally } D_t^h\text{-stationary Gaussian fields with} \]
- \( D_t^h \) positive definite and \((t, h) \rightarrow D_t^h\), continuous;
- \( \inf_{0 < h \leq 1, hs \in \mathcal{H}} \lambda_2(\{D_s^h\}D_s^h) \geq C; \)
Generalization of a theorem by Mikhaleva and Piterbarg

Assumptions:

$\mathcal{M}_1 \subset \mathcal{H}$ smooth compact 1-dimensional manifold with bounded curvature. \( \{X_h(t), \; t \in \mathbb{R}^2, \; 0 < h \leq 1 \} \) a class of centered, locally $D_t^h$-stationary Gaussian fields with

- $D_t^h$ positive definite and $(t, h) \rightarrow D_t^h$, continuous;
- $\inf_{0 < h \leq 1, hs \in \mathcal{H}} \lambda_2(D_s^h)'D_s^h \geq C$;
- $\lim_{h \rightarrow 0, ht = t^*} D_t^h = D_{t^*}^0$ uniformly in $t^* \in \mathcal{H}$;
Generalization of a theorem by Mikhaleva and Piterbarg

Assumptions:
\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact 1-dimensional manifold with bounded curvature. \( \{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\} \) a class of centered, locally \( D_t^h \)-stationary Gaussian fields with

- \( D_t^h \) positive definite and \((t, h) \rightarrow D_t^h\), continuous;
- \( \inf_{0 < h \leq 1, h \in \mathcal{H}} \lambda_2(\{D_s^h\}'D_s^h) \geq C \);
- \( \lim_{h \rightarrow 0, ht = t^*} D_t^h = D_t^0 \) uniformly in \( t^* \in \mathcal{H} \);
- \( t^* \rightarrow D_t^0, t^* \in \mathcal{H} \) is continuous.
Generalization of a theorem by Mikhaleva and Piterbarg

**Assumptions:**
\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact 1-dimensional manifold with bounded curvature. \( \{X_h(t), t \in \mathbb{R}^2, 0 < h \leq 1\} \) a class of centered, locally \( D_t^h \)-stationary Gaussian fields with

- \( D_t^h \) positive definite and \((t, h) \to D_t^h\), continuous;
- \( \inf_{0<h\leq1, h\in\mathcal{H}} \lambda_2(\{D_s^h\}'D_s^h) \geq C; \)
- \( \lim_{h \to 0, h t = t^*} D_t^h = D_t^{0^*} \) uniformly in \( t^* \in \mathcal{H}; \)
- \( t^* \to D_t^{0^*}, \ t^* \in \mathcal{H} \) is continuous.

With
\[
Q(\delta) := \sup_{0<h\leq1} \{|r_h(x + y, y)|, \|x\| > \delta\},
\]

where \( r_h(x, y) \) the covariance function of \( X_h(t) \), we have

\[
0 \leq Q(\delta) < 1
\]

\( \exists \tilde{\delta} > 0 : Q(\delta) = 0 \) for all \( \delta \geq \tilde{\delta}. \)
Heuristics of the proof.
A more general result

**Definition (Local equi-(\(\alpha, D_t\))-stationarity)**

Let \(X_h(t), t \in \mathcal{G} \subset \mathbb{R}^d\) be a class of process indexed by \(h \in \mathbb{H}\) with covariance function \(r_h(t_1, t_2)\). The sequence \(X_h(t)\) is *locally equi-(\(\alpha, D_t^h\))-stationary*, if for any \(\epsilon > 0\) there exists a positive \(\delta(\epsilon)\) independent of \(h\) such that for any \(s \in \mathcal{G}\) one can find a non-degenerated matrix \(D_s^h\) such that

\[
1 - (1 + \epsilon)\|D_s^h(t_1 - t_2)\|_\alpha \leq r_h(t_1, t_2) \leq 1 - (1 - \epsilon)\|D_s^h(t_1 - t_2)\|_\alpha
\]

provided \(\|t_1 - s\| < \delta(\epsilon)\) and \(\|t_2 - s\| < \delta(\epsilon)\) where \(\|\cdot\|\) is Frobenius norm.
Generalization of a theorem by Mikhaleva and Piterbarg

**Assumptions:**
\( M_1 \subset \mathcal{H} \) smooth compact \( r \)-dimensional manifold with positive condition number.
Generalization of a theorem by Mikhaleva and Piterbarg

**Assumptions:**

\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact \( r \)-dimensional manifold with positive condition number. \( \{X_h(t), t \in \mathbb{R}^d, 0 < h \leq 1\} \) sequence of centered, locally \((\alpha, D^h_t)\)-stationary Gaussian fields with
Generalization of a theorem by Mikhaleva and Piterbarg

**Assumptions:**

\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact \( r \)-dimensional manifold with **positive condition number.** \( \{X_h(t), t \in \mathbb{R}^d, 0 < h \leq 1\} \) sequence of centered, locally \((\alpha, D^h_t)\)-stationary Gaussian fields with

- \( D^h_t \) positive definite and \((t, h) \rightarrow D^h_t\), continuous in \( h \in (0, 1], t \in \mathbb{R}^2 \);
- \( \inf_{0 < h \leq 1, hs \in \mathcal{H}} \lambda_2(\{D^h_s\} \cap D_s^h) \geq C \),
- \( \lim_{h \rightarrow 0, ht = t^*} D^h_t = D^0_{t^*} \) uniformly in \( t^* \in \mathcal{H} \);
- \( t^* \rightarrow D^0_{t^*}, t^* \in \mathcal{H} \) is continuous.
Generalization of a theorem by Mikhaleva and Piterbarg

**Assumptions:**

\( \mathcal{M}_1 \subset \mathcal{H} \) smooth compact \( r \)-dimensional manifold with **positive condition number**. \( \{X_h(t), t \in \mathbb{R}^d, 0 < h \leq 1\} \) sequence of centered, locally \((\alpha, D^h_t)\)-stationary Gaussian fields with

- \( D^h_t \) positive definite and \((t, h) \rightarrow D^h_t\), continuous in \( h \in (0, 1], t \in \mathbb{R}^2 \);
- \( \inf_{0 < h \leq 1, h \in \mathcal{H}} \lambda_2(\{D^h_s\}'D^h_s) \geq C \),
- \( \lim_{h \rightarrow 0, h \rightarrow t^*} D^h_t = D^0_{t^*} \) uniformly in \( t^* \in \mathcal{H} \);
- \( t^* \rightarrow D^0_{t^*}, t^* \in \mathcal{H} \) is continuous.

With \( Q(\delta) \) as above

\[
Q(\delta) < 1 \quad \text{for all} \quad \delta > 0,
\]

\[
Q(\delta) \left| (\log \delta)^{2r/\alpha} \right| \leq (\log \delta)^{-\beta} \quad \text{for some} \quad \beta > 0.
\]
Generalization of a theorem by Mikhaleva and Piterbarg

**Theorem**

There exists $M > 0$ such that with

$$x_h(z) = \left(2r \log \frac{1}{h}\right)^{\frac{1}{2}} \left(1 + \frac{M + z + \left(\frac{r}{\alpha} - \frac{1}{2}\right) \log \log \frac{1}{h}}{2r \log \frac{1}{h}}\right)$$

we have

$$\lim_{h \to 0} P\{ \sup_{t \in \mathcal{M}_h} |X_h(t)| \leq x_h(z)\} = \exp\{-2 \exp\{-z\}\}$$

where $\mathcal{M}_h = \frac{M_1}{h}$.
References


References


References

**References**


References


References

References