Geometry of excursion sets: computing the surface area from discretized points

Céline Duval

j.w. H. Biermé, R. Cotsakis, E. Di Bernardino and A. Estrade

Université de Lille

Laboratoire Paul Painlevé

What is the question?

- Let $X : \mathbb{R}^d \mapsto \mathbb{R}$ be a **stationary isotropic random field**

For example ($d = 2$):

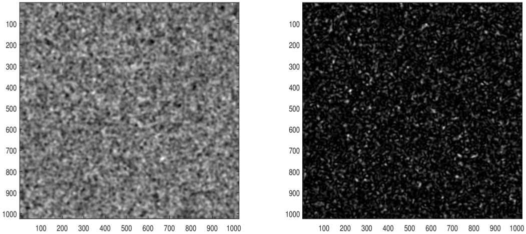


Figure: Gaussian field with covariance function $e^{-\kappa^2 \|x\|^2}$, $\kappa = 100/2^{10}$ (left), Chi square field with 2 degree of freedom (right).

What is the question?

- $X : \mathbb{R}^d \mapsto \mathbb{R}$ is a **stationary isotropic random field**
- X is observed on a **window** $T \subset \mathbb{R}^d$ trough its *excursion sets*

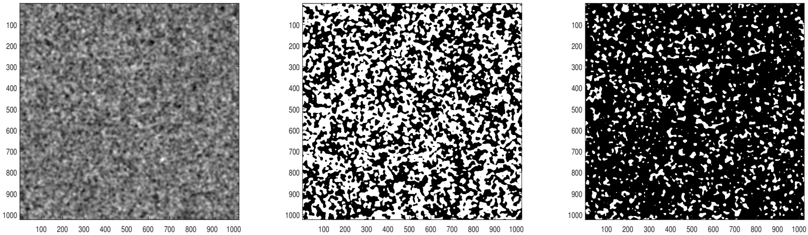


Figure: Gaussian field with covariance function $e^{-\kappa^2 \|x\|^2}$, $\kappa = 100/2^{10}$ (left) and two excursion sets for $u = 0$ (center) and $u = 1$ (right).

What is the question?

- $X : \mathbb{R}^d \mapsto \mathbb{R}$ is a **stationary isotropic random field**
- X is observed on a **window** T trough its *excursion sets at level* $u \in \mathbb{R}$

$$E_X(u) := X^{-1}\big([u,\infty)\big) = \left\{t \in \mathbb{R}^d,\ X(t) \geq u\right\}$$

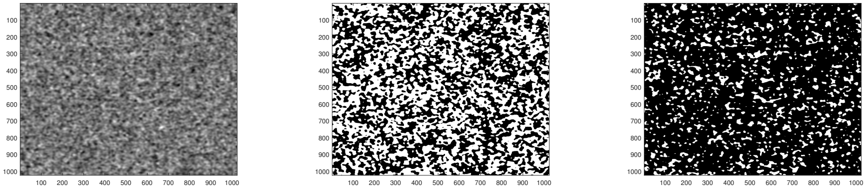
we observe: $\boxed{T \cap E_X(u_0)}$ for a fixed level u_0: *sparse information*.

Problems

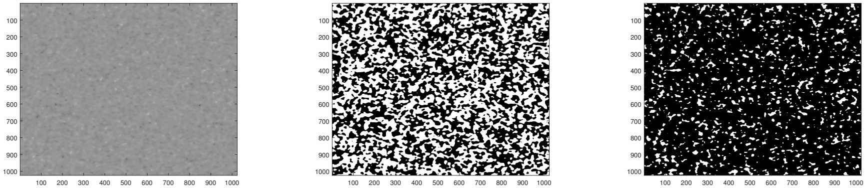
1. **Inference problem:** is it possible to recover parameters of X?
2. **Testing:** Is X Gaussian or not?

Tool: Geometry of the excursion sets $T \cap E_X(u)$

A not so trivial question...



Normalized Gaussian field with covariance function $e^{-\kappa^2\|x\|^2}$ (left) and two excursion sets for $u = 0$ (center) and $u = 1$ (right).



Normalized Student field with 4 degrees of freedom (left) and two excursion sets for $u = 0$ (center) and $u = 1$ (right).

Contents

1. Lipschitz-Killing curvatures for excursion sets
2. Pixelisation: Computing the surface area from discrete points

Lipschitz-Killing curvatures in dimension 2

If $d = 2$, for a "nice" set A one can define 3 LK curvatures:

- $\chi(A)$: Euler characteristic ($d = 2$): ($\sharp$(connected component) - $\sharp$(holes)) of A, related to the *connectivity*,
- $\sigma_1(A)$: Surface area (perimeter $d = 2$) of A, related to the *regularity*,
- $\sigma_2(A)$: Area of A, related to the *occupation density*.

Applications: Cosmology, 2D x-ray images (detection of osteoporosis, mammograms),...

Lipschitz-Killing curvatures

In dimension $d \geq 2$, there exist $d+1$ - LK curvatures:

$\sigma_d(A)$: Area of A: the Lebesgue measure of A,

$\sigma_{d-1}(A)$: Surface area of A, $d-1$-dimensional Hausdorff measure of ∂A.

For $0 \leq k \leq d$ the k-dimensional Hausdorff measure of $B \subset \mathbb{R}^d$ is

$$\sigma_k(B) := \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)} \lim_{\delta \to 0} \inf_{\substack{\operatorname{diam}(U_i) < \delta \\ \bigcup_{i=1}^{\infty} U_i \supseteq B}} \sum_{i \in \mathbb{N}} \left(\frac{\operatorname{diam}(U_i)}{2}\right)^k,$$

where the infimum is taken over all countable covers of B by arbitrary subsets U_i of $\mathbb{R}^d$.

Lipschitz-Killing curvatures for excursion sets

Let X be stationary, isotropic and a.s. twice differentiable such that the probability density of $\big(X(0), \nabla X(0)\big)$ is bounded uniformly on $\mathbb{R}^{d+1}$.

Let $u \in \mathbb{R}$, $T \subset \mathbb{R}^d$, define the excursion set within T above level u:

$$E_X^T(u) := \{t \in T : X(t) \geq u\} = T \cap E_X(u), \quad E_X(u) := X^{-1}([u, +\infty)).$$

and the level curves within T

$$L_X^T(u) := \{t \in T : X(t) = u\} = T \cap \partial E_X(u), \textit{ a.s.}$$

Lipschitz-Killing curvatures for excursion sets

- **Empirically accessible quantities**:

$$C_d^T(u) := \frac{1}{\sigma_d(T)} \sigma_d\left(E_X^T(u)\right) = \frac{1}{\sigma_d(T)} \int_T \mathbb{1}_{\{X(t) \geq u\}} \mathrm{d}t,$$

$$C_{d-1}^T(u) := \frac{1}{\sigma_d(T)} \sigma_{d-1}\left(L_X^T(u)\right) = \frac{1}{\sigma_d(T)} \int_{L_X^T(u)} \sigma_{d-1}(\mathrm{d}s).$$

Matlab functions: `bwarea`, `bwperim` (pixelization error)

- **LK densities**: involve parameters of the field

$$C_k^*(u) := \mathbb{E}[C_k^T(u)], \ \text{for } k = d, d-1, \forall T.$$

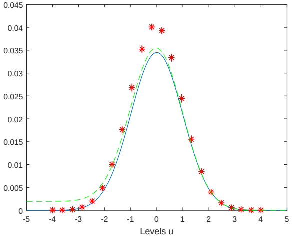
Computing $C_k^*(u)$ go to ?
Area ✓: $C_d^*(u) = \mathbb{E}[C_d^T(u)] = \mathbb{P}(X(0) \geq u)$.
Others? : Tube formula, characteristic function...

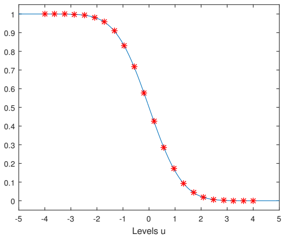
Statistical strategy

1. Observations: $T \cap E_X(u)$ for T a large hypercube in $\mathbb{R}^d$,
2. Compute: $C_k^T(u)$, $k = d-1, d$
3. Relate them to the parameters of the field: estimation / testing procedures

Example: For $d = 2$, X a Gaussian random field X centered with unit variance and second spectral moment $\lambda > 0$

$\frac{1}{2}$perimeter: $\frac{1}{4}\,\lambda^{1/2}\,e^{-u^2/2}$ area: $\int_u^\infty \frac{e^{-v^2/2}}{\sqrt{2\pi}}dv$





Estimated $C^T_{d,d-1}(u)$ and Theoretical $C^*_{d,d-1}(u)$.

Statistical strategy: technical limitations

1. Difficult to establish "general CLT results" for $C_k^T(u)$, as $T \nearrow \mathbb{R}^d$,
    - Known in particular cases (Gaussian, chi-square for $d = 1$),
    - Asymptotic variances not (always) explicit.

Statistical strategy: technical limitations

1. Difficult to establish "general CLT results" for $C_k^T(u)$, as $T \nearrow \mathbb{R}^d$,
2. Unreasonable assumption X centered with unit variance
   - The mean/variance provide information on the LK curvatures
   - From the excursion set: impossible to estimate mean/variance(X)
   - Image comparison: fields with distant mean/variances?

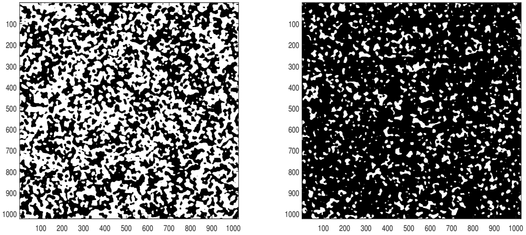


Figure: Two excursion sets for $u = 0$ of the same realization of a Gaussian field (left: initial field, right: shifted field).

Statistical strategy: technical limitations

1. Difficult to establish "general CLT results" for $C_k^T(u)$, as $T \nearrow \mathbb{R}^d$,
2. Unreasonable assumption X centered with unit variance
   Notion of effective level in the Gaussian case

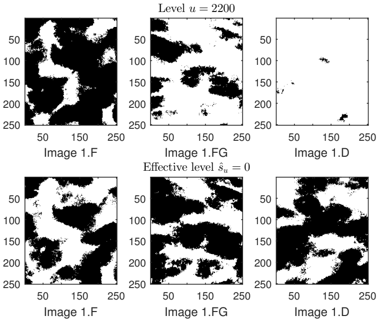
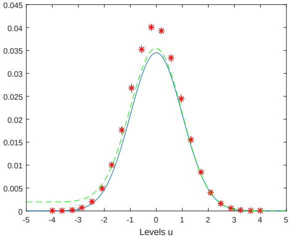


Figure: Synthetic digital mammograms. Excursion sets for a fixed level $u = 2200$ (first row) and for the three adaptive levels $\widetilde{u}$, such that the effective levels coincide (second row).

Statistical strategy: technical limitations

1. Difficult to establish “general CLT results” for $C_k^T(u)$, as $T \nearrow \mathbb{R}^d$,
2. Unreasonable assumption X centered with unit variance
3. Pixelisation: Area ✓, Surface area ✗

$$\frac{1}{2}\text{perimeter: } \frac{1}{4}\lambda^{1/2} e^{-u^2/2}$$



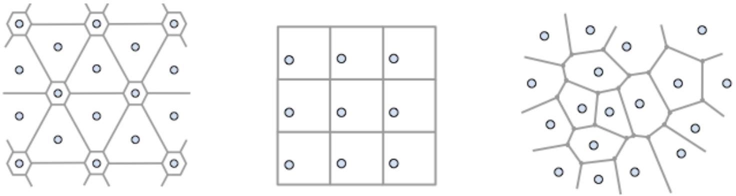
Estimated $C_{d,d-1}^T(u)$ and Theoretical $C_{d,d-1}^*(u)$.

Contents

1. Lipschitz-Killing curvatures for excursion sets

2. Pixelisation: Computing the surface area from discrete points

Polytopic tessellations based on point clouds



(a) Hexagones and truncated triangles lattice

(b) Square lattice

(c) Voronoi tessellation

Figure: Three examples of tilings with a particular choice of reference points.

Let $\mathscr{H}$ be a set of convex, closed polytopes that tessellates $\mathbb{R}^d$, let $\dot{\mathscr{H}}^T = \{(P, P^\bullet) : P \in \mathscr{H}, P \subset T\}$ be a *point-referenced honeycomb*.

- $\sigma_d(P_1 \cap P_2) = 0$, for any $P_1 \neq P_2 \in \mathscr{H}$ and for P_1, P_2 *adjacent*, then $\sigma_{d-1}(P_1 \cap P_2) > 0$.

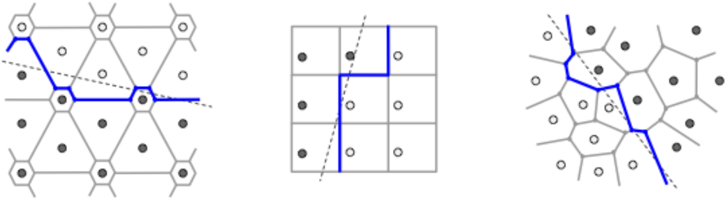
Estimated area and surface area on point clouds

Define

$$\widehat{C}_d^{(\dot{\mathscr{H}},T)}(u) := \frac{1}{\sigma_d(T)} \sum_{P \in \mathscr{H}^T} \sigma_d(P) \mathbb{1}_{X(P^\bullet) \geq u}$$

and

$$\widehat{C}_{d-1}^{(\dot{\mathscr{H}},T)}(u) := \frac{1}{\sigma_d(T)} \sum_{\substack{P_1, P_2 \in \mathscr{H}^T \\ P_1 \neq P_2}} \sigma_{d-1}(P_1 \cap P_2) \mathbb{1}_{X(P_1^\bullet) \leq u < X(P_2^\bullet)}.$$



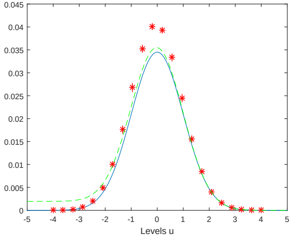
Estimated area and surface area on point clouds

- (Almost) unibiased area

$$\mathbb{E}[\widehat{C}_d^{(\dot{\mathscr{H}},T)}(u)] = \frac{\sigma_d\left(\bigcup_{P\in\mathscr{H}^T} P\right)}{\sigma_d(T)} C_d^*(u). \quad \checkmark$$

- For the surface area? It has been reported (Miller (1999), Biermé and Desolneux (2021)) for $d=2$, specific tilings and particular fields

$$\mathbb{E}\left[\widehat{C}_{d-1}^{(\dot{\mathscr{H}}_\delta,T)}(u)\right] \underset{\delta\to 0}{\not\to} C_{d-1}^*(u).$$



A result on crossings: $\mathbb{P}(X(P_1^\bullet) \leq u < X(P_2^\bullet))$

Theorem (Cotsakis, Di Bernardino, D.)

Fix $\mathbf{w} \in \partial B_1^d$ *and* $u \in \mathbb{R}$,

$$\lim_{q \to 0} \frac{1}{q} \mathbb{P}\left(X(\mathbf{0}) \leq u < X(q\mathbf{w})\right) = \frac{C_{d-1}^*(u)}{\beta_d},$$

$$\text{where} \qquad \beta_d = \frac{2\sqrt{\pi}\,\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})}.$$

Fix $\varepsilon \in (0,1)$, *under additional assumptions on* X, *there exists a constant* $K > 0$ *such that for all* $q > 0$,

$$0 \leq \frac{C_{d-1}^*(u)}{\beta_d} - \frac{1}{q} \mathbb{P}\left(X(\mathbf{0}) \leq u < X(q\mathbf{w})\right) \leq Kq^{1-\varepsilon}.$$

Comments

- Improve a related result for $d = 1$ Leadbetter et al. (1983)

- Numerically evaluation of $C^*_{d-1}(u)$: generate, for q small, M *i.i.d.* copies of $(Z_1, Z_2) \sim \big(X(\mathbf{0}), X(q\mathbf{e}_1)\big)$,

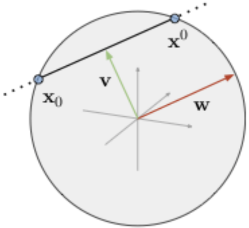
$$\left| C^*_{d-1}(u) - \frac{\beta_d}{q} \sum_{j=1}^{M} \mathbb{1}_{Z_1^j \leq u \leq Z_2^j} \right| \lesssim q^{1-\varepsilon} \vee M^{-1/2}.$$

Elements of proof: Crofton's formula (Schneider and Weil (2008))

Define the Affine Grassmanian $A(d,1)$: the set of lines in $\mathbb{R}^d$.

- **Parametrization of $A(d,1)$:** for $\mathbf{w} \in \partial B_1^d$ and $\mathbf{v} \in \mathrm{vect}(\mathbf{w}^\perp)$, denote

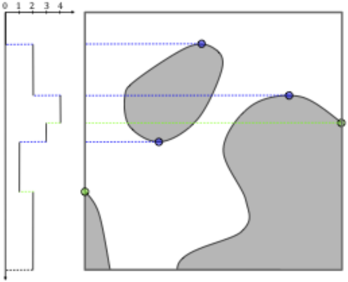
$$l_{\mathbf{w},\mathbf{v}} := \{\mathbf{v} + \lambda \mathbf{w} : \lambda \in \mathbb{R}\}.$$



$$\sigma_{d-1}(M) = \frac{\sqrt{\pi}\,\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \int_{\mathbb{R}^{d-1}} \int_{\partial B_1^d} \frac{\sigma_0\left(M \cap l_{\mathbf{w},\mathbf{v}_\mathbf{w}(\mathbf{x})}\right)}{\sigma_{d-1}(\partial B_1^d)} \,\mathrm{d}\mathbf{w}\,\mathrm{d}\mathbf{x}$$

where σ_0 is a counting measure of isolated points.

Elements of proof: Crofton's formula (Schneider and Weil (2008))



Let $L_X^B(u) = L_X(u) \cap B(0,1)$:

$$\mathbb{E}[\sigma_{d-1}(L_X^B(u))] = \frac{\sqrt{\pi}\,\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \int_{\mathbb{R}^{d-1}} \int_{\partial B_1^d} \frac{1}{\sigma_{d-1}(\partial B_1^d)} \mathbb{E}[\sigma_0(L_X^B(u) \cap l_{\mathbf{w},\mathbf{v}_\mathbf{w}(\mathbf{x})})]\, \mathrm{d}\mathbf{w}\, \mathrm{d}\mathbf{x}.$$

$\mathbb{E}[\sigma_{d-1}(L_X^B(u))] \to C_{d-1}^*(u)$

$\frac{\sqrt{\pi}\,\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \to \frac{1}{2} \times \beta_d$

$\mathbb{E}[\sigma_0(L_X^B(u) \cap l_{\mathbf{w},\mathbf{v}_\mathbf{w}(\mathbf{x})})] \to \frac{2}{q} \times$ crossing

Implication: Explicit (asymptotic) bias for the surface area

Theorem (Cotsakis, Di Bernardino, D.)

Let $D^{(\mathscr{H})} := \sup\{\mathrm{diam}(P \cap T) : P \in \mathscr{H}\}$ *and* $\delta\mathscr{H} := \{\delta P : P \in \mathscr{H}\}$ *for* $\delta \in \mathbb{R}^+$. *Suppose that* $\lim_{\delta\to 0} D^{(\delta\mathscr{H})} = 0$, *then, it holds*

$$\mathbb{E}\left[\widehat{C}_{d-1}^{(\delta\dot{\mathscr{H}},T)}(u)\right] \xrightarrow[\delta\to 0]{} \frac{2d}{\beta_d} C_{d-1}^{*}(u).$$

- $d = 2$ we get $\frac{2d}{\beta_d} = \frac{4}{\pi}$.
- We can derive: For all $d \geq 1$, square lattice and *nice* fields

$$\sqrt{\sigma_d(T_N)}\left(\mathbb{E}[\widehat{C}_{d-1}^{(\delta,T_N)}(u)] - \frac{2d}{\beta_d} C_{d-1}^{*}(u)\right) \to 0 \qquad (\star)$$

for $N\delta \to \infty$, $(N\delta)^{d/2}\delta^{1-\varepsilon} \to 0$, $\varepsilon \in (0,1)$.

Implication: CLT results

- Existing CLT: for the unavaible *continuoulsly* observed $C^{T}_{d,d-1}(u)$.
  $\rightarrow$ Cannot use them for $\widehat{C}^{(\delta\dot{\mathscr{H}},T)}_{d,d-1}(u)$.
- Square lattice: Adapt a CLT result for mixing random fields of Iribarren (1989) and ($\star$) for the observed pixelized $\widehat{C}^{(\delta,T)}_{d,d-1}(u)$.
  Constraints: $(N\delta)^{d/2}\delta^{1-\varepsilon} \rightarrow 0$, $\varepsilon \in (0,1)$, and X is strongly alpha mixing such that for some $\eta > 0$

$$\sum_{r=1}^{+\infty} r^{3d-1}\alpha(r)^{\frac{\eta}{2+\eta}} < +\infty.$$

To conclude

Estimating the surface area from a point cloud based on polytopic tesselations generates a bias, that depends only on the dimension : (*almost*) not on the nature of the field nor on the pixelization

To conclude

Estimating the surface area from a point cloud based on polytopic tesselations generates a bias, that depends only on the dimension : (*almost*) not on the nature of the field nor on the pixelization

Merci!

Non-exhaustive list of references

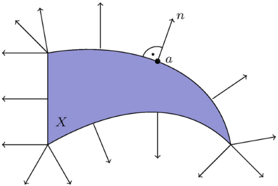
- ↪ Adler, Taylor. *Random fields and geometry*. Springer, New York, 2007.
- ↪ Biermé and Desolneux. The effect of discretization on the mean geometry of a 2D random field. *Annales Henri Lebesgue*, 4:1295–1345, (2021).
- Biermé, Di Bernardino, Duval, Estrade. Lipschitz-Killing curvatures of excursion sets for two dimensional random fields, *EJS*, 13, 536-581, (2019).
- ↪ Bulinski, Spodarev, and Timmermann. Central limit theorems for the excursion set volumes of weakly dependent random fields. *Bernoulli*, 18(1):100–118, (2012).
- Cotsakis, Di Bernardino, Duval. Surface area and volume of excursion sets observed on point cloud based polytopic tessellations, Preprint *ArXiv*.
- Di Bernardino, Duval. Statistics for Gaussian Random Fields with Unknown Location and Scale using Lipschitz-Killing Curvatures, *Scand. Journal of Stat.* 49, 143-184, (2022).
- ↪ Iribarren. Asymptotic behaviour of the integral of a function on the level set of a mixing random field. *Probability and Mathematical Statistics*, 10(1):45 – 56, (1989).
- ↪ Leadbetter, Lindgren, and Rootzén. *Extremes and related properties of random sequences and processes. Springer Series in Statistics.* Springer, Germany, (1983).
- ↪ Miller. Alternative Tilings for Improved Surface Area Estimates by Local Counting Algorithms. *Computer Vision and Image Understanding,* 74(3):193–211 (1999).
- ↪ Schneider and Weil. *Stochastic and integral geometry. Probability and its Applications*. Springer-Verlag, Berlin, (2008).

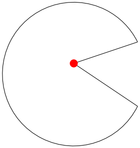
Lipschitz-Killing curvatures for excursion sets

Question: How to **properly** define these quantities for $T \cap E_X(u)$?

Tool: Curvature measures for Positive Reach sets (Federer 59)

Intuitively: *"A is a **positive reach set** if one can roll a ball of positive radius along the exterior boundary of A keeping in touch with A."*





Curvature **measures**

Let A be a Positive Reach set, define for any Borel set $U \subset \mathbb{R}^2$

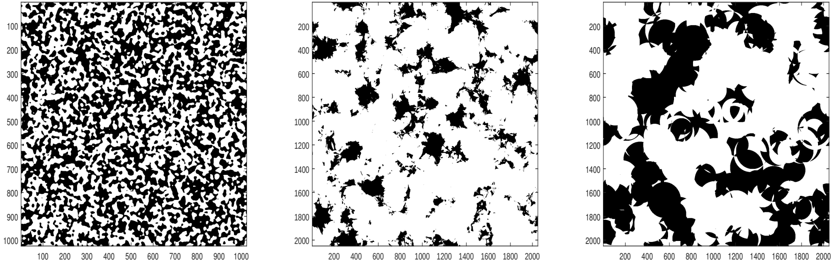
$$\underbrace{\Phi_0(A,U)}_{\text{Euler characteristic}} = \frac{\mathrm{TC}(\partial A, U)}{2\pi}, \quad \underbrace{\Phi_1(A,U)}_{\frac{1}{2}\text{Perimeter}} = \frac{|\partial A \cap U|_1}{2} \quad \textit{and} \quad \underbrace{\Phi_2(A,U)}_{\text{Area}} = |A \cap U|,$$

- $\mathrm{TC}(\partial A, U)$ is the integral over U of the curvature along the positively oriented curve ∂A
- $|.|_1$ the one-dimensional Hausdorff measure
- $|.|$ the two-dimensional Lebesgue measure.

Remarks:

- The measures $\Phi_i(A, \cdot)$ are additive: Union of Positive Reach go to
- We take $A = T \cap E_X(u)$, which is in the UPR class *a.s.* if, *e.g.*
  - X is of class C^2 *a.s.*
  - $E_X(u)$ is locally given by a finite union of disks.

Lipschitz-Killing curvatures for excursion sets return



✓ Student random field; $E_X(u) \in$ UPR *a.s.*

✓ Shot noise field, $B = 1$ *a.s.* $E_X(u) \notin$ PR *a.s.* but $E_X(u) \in$ UPR *a.s.*

✗ Shot noise field, $B \pm 1$ *a.s.* $E_X(u) \notin$ PR *a.s.* and $E_X(u) \notin$ UPR *a.s.* Biermé and Desolneux (2017)

Lipschitz-Killing curvatures for excursion sets back

How can we compute $C^*_{0,1,2}(X,u)$?

- *Gaussian type fields*: $X = F(\mathbf{G})$ where $\mathbb{V}(G'_i(0)) = \lambda I_2$, $\lambda > 0$,

$$\mathbb{P}\left(\mathbf{G}(0) \in Tube(F,\rho)\right) = C^*_2(X,u) + \rho \frac{2\sqrt{2}}{\sqrt{\lambda\pi}} C^*_1(X,u) + \rho^2 \frac{\pi}{\lambda} C^*_0(X,u) + O(\rho^3).$$

e.g. for a Student field:

$$F : (x,y) \in \mathbb{R} \times \mathbb{R}^k \mapsto \frac{x}{||y||/\sqrt{k}},$$