Functional central limit theorem for Betti numbers of Gaussian excursion sets

Christian Hirsch

joint work with R. Lachièze-Rey

Geometric and Topological Properties of Random Algebraic Varieties, Cologne, 6/10/23

Setup

- $F := \{F(x)\}_{x \in \mathbb{R}^d} =$ centered Gaussian random field
- ***Stationarity*** $\rightsquigarrow F(\cdot + x) \overset{D}{=} F(\cdot)$, $x \in \mathbb{R}^d$
  $\rightsquigarrow (F(x_1), \ldots, F(x_n))$ Gaussian random vector, $x_1, \ldots, x_n \in \mathbb{R}^d$
- ***Covariance function***

$$C(x, y) := \mathrm{Cov}(F(x), F(y))$$

- $A(u; F) := \{x : F(x) \geqslant u\} =$ excursion set

Betti numbers

- $\beta_0, \beta_1, \beta_2 =$ number of connected components, loops, cavities
- $\beta_{q,n}(u) := \beta_q\big(A(u; F) \cap W_n\big) := \beta_q\big(A(u; F) \cap [0, n]^d\big)$

Dream goal. Functional central limit theorem (CLT)

$$|W_n|^{-1/2}(\beta_{q,n}(u) - \mathbb{E}[\beta_{q,n}(u)]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(u)).$$

as a stochastic process in u.

Motivation from natural sciences

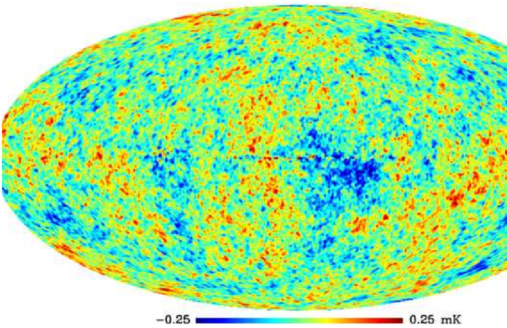
- Cosmic microwave background
- Materials science
- CLT used for hypothesis tests

(Estrade & León, '16)

- CLT for ***Euler characteristic***
- More local than Betti number
- Malliavin calculus

(Beliaev, McAuley, Muirhead, '23+)

- CLT for ***component count***
- Martingale CLT & stabilization
- Finite moments for critical points

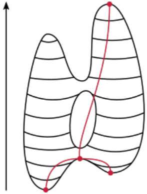


Morse theory

⇝ $\beta_{q,n}(u)$ only changes at values u of ***critical points***

⇝ ***Marked point process of critical points of the field***

$$Y(W \times I) := \{(x, F(x)) \in W \times I \colon \nabla F(x) = 0\}$$



https://bastian.rieck.me

Major problem. Notoriously difficult to get moments of critical points

- ▷ Progress by (Gass & Stecconi, '23) and (Ancona & Letendre)
- ▷ Functional CLT needs control of the moment for the number of critical points with ***values in small intervals***
- ▷ Restrict to ***subcritical level-set percolation***

Gaussian fields assumptions

- ***Regularity.*** F should be of class $\mathcal{C}^4$
- ***Spectral regularity.*** $\mathrm{supp}(\hat{C})$ contains an open set
- ***Correlation decay.*** C should decay at ∞ as $\|x\|^{-\beta}$ for some $\beta \gg 0$

$\mathcal{C}(Q; A) :=$ union of connected components of A hitting Q.

Percolation assumption

For Q compact, for some $u_c \in \mathbb{R}$ there is $\lambda > 0$ such that

$$\limsup_{m \uparrow \infty} \frac{\log \mathbb{P}(\mathrm{diam}(\mathcal{C}(Q; A(u_c))) \geqslant m)}{m^\lambda} < 0.$$

$\rightsquigarrow$ (Severo, '22)

Represent Betti numbers as

$$\beta(W, I) = \sum_{(x,v)\in Y(I)} \delta_x(W) + (\text{facet contributions})$$

where $\delta_x = \mathbf{1}_{\{x \text{ ref. point of } C\}}\beta_q(C, I)$.

Theorem (Fixed-level CLT)

Let $u > u_c$. Then, for $\beta_n(u) := \beta(W_n, A(u; F))$,

$$\widetilde{\beta_n}(u) := |W_n|^{-1/2}(\beta_n(u) - \mathbb{E}[\beta_n(u)]) \Rightarrow \mathcal{N}(0, \sigma(u)^2).$$

If $\liminf_{n\uparrow\infty} |W_n|^{-1}\mathbb{E}\big[\beta_n(u)\big] > 0$, then $\sigma(u)^2 > 0$.

Theorem (FCLT for regular Betti numbers)

On any compact contained in $[u_c, \infty)$, it holds that $\tilde{\beta}_n \overset{d}{\Rightarrow} G$, for some centered Gaussian process G on $[u_c, \infty)$.

A ***Gaussian white noise*** is a random signed measure $\mathcal{W}$ such that

- $\mathcal{W}(A) \sim \mathcal{N}(0, |A|)$ for $A \subseteq \mathbb{R}^d$ measurable
- $\mathcal{W}(A \cup B) = \mathcal{W}(A) + \mathcal{W}(B)$ for disjoints $A, B \subseteq \mathbb{R}^d$
- $\mathcal{W}(A)$ and $\mathcal{W}(B)$ are independent for disjoint $A, B \subseteq \mathbb{R}^d$.

In particular for f_1, f_2 square integrable

$$\mathrm{Cov}\Big(\int_{\mathbb{R}^d} f_1(x)\mathcal{W}(\mathrm{d}x), \int_{\mathbb{R}^d} f_2(x)\mathcal{W}(\mathrm{d}x)\Big) = \int_{\mathbb{R}^d} f_1(x) f_2(x)\mathrm{d}x.$$

Let $g : \mathbb{R} \to \mathbb{R}$ be symmetric, square integrable and define $F(x) := g \star \mathcal{W}$.
Then,

$$g \star g = C =: \hat{\rho} = \widehat{\sqrt{\rho}} \star \widehat{\sqrt{\rho}}.$$

Therefore,

$$g = \widehat{\sqrt{\rho}}$$

Idea. Martingale decomposition

For $B \subseteq \mathbb{R}^d$, define the ***independent resampling*** of $\mathcal{W}$ in B, i.e.

$$\mathcal{W}^{(B)}(A) = \mathcal{W}(A \setminus B) + \mathcal{W}'(A \cap B)$$

where $\mathcal{W}'$ is a white noise independent of $\mathcal{W}$.

- ▷ $\mathcal{G}_j :=$ σ-algebra generated by the restriction of $\mathcal{W}$ to boxes $Q_i := i + [0,1]^d$ with $i \leqslant_{\text{lex}} j$ lexicographically

Then, with $B^{(n)}_{\Delta,j}(u) := \beta_n(u; F) - \beta_n(u; F^{(Q_j)})$,

$$\beta_n(u) - \mathbb{E}[\beta_n(u)] = \sum_{j \in \mathbb{Z}^d} \mathbb{E}[B^{(n)}_{\Delta,j}(u) \mid \mathcal{G}_j] =: \sqrt{|W_n|} \sum_{j \in \mathbb{Z}^d} U_{j,n}$$

Conditions for martingale CLT (McLeish, '74); (BMM, '23+).

1. $\sup_{j \in \mathbb{Z}^d} |U_{j,n}| \to 0$ in probability
2. $\sup_{n \geqslant 1} \mathbb{E}\Big[\sup_{j \in \mathbb{Z}^d} U_{j,n}\Big] < \infty$;
3. $\sum_{j \in \mathbb{Z}^d} U_{j,n}^2 \xrightarrow{L^1} \sigma^2$ for some $\sigma^2 \geqslant 0$.
4. $\mathbb{E}\Big[\sum_{j \in \mathbb{Z}^d} |U_{j,n}|\Big] < \infty$ for all $n \geqslant 1$;

Issue. How to check conditions in practice?

- ▷ Simplifications (Penrose & Yukich, '01); (BMM, '23+)
- ⇝ Reduction to a ***stabilization*** and a ***moment*** condition.

Proposition (Stabilization condition; Lemma 3.7 of (BMM, '23+))

The sequence $\{B^{(n)}_{\Delta,o}\}_n$ *converges almost surely to some a.s. finite* $B^{(\infty)}_{\Delta,o}$*.*

Proof idea.

- ▷ Set $\gamma := \beta/d - 1/2$ and consider the decomposition

$$B^{(n)}_{\Delta,j}(u) =: \sum_{i \in \bar{W}_n} B_{\Delta,i,j},$$

Key task.

$$\mathbb{P}(B_{\Delta,i,o} \neq 0) = |i|^{-\gamma+o(1)}$$

Borel-Cantelli ⇝ a.s. only finitely many of the $B_{\Delta,i,o}$ are different from 0.

Moment condition

▷ $d_{j,n} := \mathrm{dist}(j, W_n) =$ distance of $j \in \mathbb{Z}^d$ from W_n

Proposition (Moment conditions; Lemma 3.6 of (BMM, '23+))

For any sufficiently small $\varepsilon > 0$, it holds with $q = q(\varepsilon) = 2 + \varepsilon$ that

1. $\sup_{n\geqslant 1} \sup_{j\in\mathbb{Z}^d} \mathbb{E}[|B^{(n)}_{\Delta,j}|^q] < \infty$;
2. $$\sup_{n,k\geqslant 1} \frac{\sum_{j\in\mathbb{Z}^d:\, d_{j,n}>k} \mathbb{E}[|B^{(n)}_{\Delta,j}|^q]}{|W_n|^3 k^{-\beta/3}(k^d + k|W_n|^{(d-1)/d})} < \infty;$$
3. $\sup_{n\geqslant 1} |W_n|^{-1} \sum_{j\in\mathbb{Z}^d} \mathbb{E}[|B^{(n)}_{\Delta,j}|^q] < \infty$.

Lemma (Moment bound on $B^{(n)}_{\Delta,j}$)

Let $3 \leqslant m < q_0(1-1/\gamma)$ *and* $\mathbb{E}[Y(Q_o)^{q_0}] < \infty$. *Then,*

$$\mathbb{E}[|B^{(n)}_{\Delta,j}|^m] \in O\big(|W_n|^m d_{j,n}^{-\gamma d(q_0-m)/q_0}(|W_n| \wedge d_{j,n}^d)\big).$$

Proof sketch. First,

$$\mathbb{E}[|B^{(n)}_{\Delta,j}|^m] \leqslant |\bar{W}_n|^m \mathbb{E}[\max_{i \in \bar{W}_n} |B^{(n)}_{\Delta,i,j}|^m] \leqslant |\bar{W}_n|^m \sum_{i \in \bar{W}_n} \mathbb{E}[B^m_{\Delta,i,j}].$$

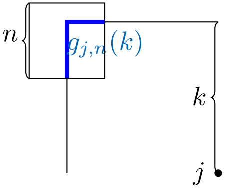
By Hölder with $q' = q_0/m$ and $p' = q_0/(q_0-m)$,

$$\mathbb{E}[B^m_{\Delta,i,j}] \leqslant \mathbb{E}[B^{q_0}_{\Delta,i,j}]^{1/q'} \mathbb{P}(B_{\Delta,i,j} \neq 0)^{1/p'} = |i-j|^{-\gamma d/p' + o(1)}$$

Moment bound (II)

Now,

$$\sum_{i \in \bar{W}_n} |i-j|^{-\gamma d/p'} \leqslant C \sum_{k \geqslant d_{j,n}} g_{j,n}(k) k^{-\gamma d/p'}$$
$$\in O\Big((|W_n| \wedge d_{j,n}^d) d_{j,n}^{-\gamma d/p'} \Big),$$



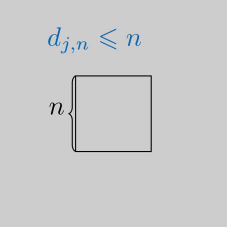
where
$g_{j,n}(k) := \Big| \big\{ i \in \bar{W}_n : |i-j| = k \big\} \Big| \in O\big(k^{d-1} \wedge n^{d-1} \big)$.

Proof of Proposition. Putting $\sigma := q/q_0$, we note that for $k \geqslant n$,

$$\sum_{j:\, d_{j,n} \geqslant k} \mathbb{E}[|B_{\Delta,j}^{(n)}|^q] \leqslant C \sum_{\ell \geqslant k} g_\ell^{(n)} |W_n|^{q+1} \ell^{-\gamma d(1-\sigma)},$$

where $g_\ell^{(n)} := |\{ i \in \mathbb{Z}^d : \mathrm{dist}(i, W_n) = \ell \}|$. Since $g_\ell^{(n)} \in O\big(\ell^{d-1}\big)$, this is of order $O\Big(|W_n|^{q+1} k^{d-\gamma d(1-\sigma)} \Big)$.

Moment bound (III)



$\#\{j\colon d_{j,n} \leqslant n\} \in O(|W_n|) \rightsquigarrow$ suffices to consider $d_{j,n} \geqslant n$.
Then, by the previous result with $k = n$,

$$|W_n|^{-1} \sum_{j\colon d_{j,n} \geqslant n} \mathbb{E}[|B_{\Delta,j}^{(n)}|^q] \in O\big(|W_n|^{q+1-\gamma(1-\sigma)}\big).$$

Part (2). For $k \geqslant n$,

$$|W_n|^{-1} \sum_{j\colon d_{j,n} \geqslant k} \mathbb{E}[|B_{\Delta,j}^{(n)}|^q] \in O\big(|W_n|^{q+1} k^{d-\gamma(q_0-\sigma))d}\big).$$

In particular,

$$\frac{|W_n|^{q+1} k^{(1-\gamma(1-\sigma))d}}{|W_n|^3 k^{d-\beta/3}} = |W_n|^{q-2} k^{(1/6-\gamma(2/3-\sigma))d} \in O(1).$$

Goal. Show $\mathbb{P}(B_{\Delta,i,o} \neq 0) = |i|^{-\gamma d + o(1)}$

- Set $\Delta_B(x) := F(x) - F^{(B)}(x) := F(x) - g \star \mathcal{W}^{(B)}(x)$
- ***Need:*** topology of $\{F \geqslant u\} \cap \mathcal{C}_i$ is the same as $\{F + \Delta \geqslant u\} \cap \mathcal{C}_i$ if no "topological event" occurs in the evolution $t \to F_t := F + t\Delta$.

Lemma (Deterministic topological lemma)

Assume that

1. *There is no critical point in $\partial \mathcal{C}_i$ (i.e. such that $\nabla F_t(x) = 0, F_t(x) = u$)*
2. *No $x \in Q_i$ shall lose (or gain) the "reference point" status*

Then,

$$B_{\Delta,i,j} = 0$$

Proof idea.

- Lemma 4.1 in (BMM, '23+)
- Morse-type theory; homotopy extension theorem

Question. How to bound probabilities of topological events?

- Why are they unlikely?

Idea. “More equations than unknowns”

- For $(x, \hat{y}) = (x, (y_2, \ldots, y_d)) \in \mathbb{R}^{2d-1}$, denote by $\tilde{y} = (x_1, y_2, \ldots, y_d)$.
- Let $A = (x, \hat{y}) \in Q \times \mathbb{R}^{d-1} : \tilde{y} \in Q$.
- Define the operator $D^2 F(x, \hat{y}) := \|x - \tilde{y}\|^{-1}(\nabla F(x) - \nabla F(\tilde{y}))$.
- ***Relevant event.*** $\exists t \in [0, 1]$:

$$\nabla F(x) + t\nabla\Delta(x) = \nabla F(\tilde{y}) + t\nabla\Delta(\tilde{y}) = 0.$$

or, equivalently

$$\nabla F(x) + t\nabla\Delta(x) = D^2 F(x, \hat{y}) + tD^2\Delta(x, \hat{y}) = 0.$$

In particular,

$$\|\nabla F(x)\| < \|\nabla\Delta(x)\|, \|D^2 F(x, \hat{y})\| < \|D^2\Delta(x, \hat{y})\|.$$

Lemma (Bulinskaya)

Let

- $A \subseteq \mathbb{R}^{m'}$ *compact with* $m' < m$.
- $(f_1(x), \ldots, f_m(x))$, $x \in A$*: non-degenerate centred Gaussian field.*

Then,

$$\limsup_{\tau \to 0} \frac{\log \mathbb{P}(\inf_{x \in A} \max_{i \leqslant m} |f_i(x)| \leqslant \tau)}{\log \tau} \leqslant -1.$$

Sketch.

- If for some $x \in A$, $\max_{i \leqslant m} |f_i(x)| < \tau$, then for $y \in B(x, \tau)$,

$$\max_{i \leqslant m} |f_i(y)| \leqslant \tau(1 + \|\nabla f\|_{B(x,\tau)}).$$

Let $\mu \in (0,1)$, and $X := \int_A \prod_{i \leqslant m} |f_i(x)|^{-\mu} dx$.
Work from now under the event $\Omega := \left\{ \inf_{x \in A} \max_{i \leqslant m} |f_i(x)| \leqslant \tau \right\}$

$$X \geqslant \prod_{i \leqslant m} |f_i(x)|^{-\mu} \geqslant \tau^{m'} \prod_{i \leqslant m} \frac{1}{\tau(1 + \|\nabla f\|_{B(x,\tau)})} =: \tau^{m' - m\mu} \frac{1}{B},$$

Hence, by Jensen's inequality,

$$\begin{aligned} \mathbb{P}(\Omega) \leqslant & \tau^{m\mu - m'} \mathbb{E}[BX] = \tau^{m\mu - m'} \mathbb{E}[B^q]^{\frac{1}{q}} \mathbb{E}\Big[\Big(\int_A \prod_{i \leqslant m} |f_i(x)|^{-\mu} dx\Big)^p\Big]^{1/p} \\ \leqslant & c\tau^{m\mu - m'} |A| \mathbb{E}\Big[\int_A \prod_{i \leqslant m} |f_i(x)|^{-\mu p} \frac{dx}{|A|}\Big]^{1/p}. \end{aligned}$$

- Choose μ such that $m\mu - m' = 1 - \varepsilon$ and set $p := (1 - \varepsilon)/\mu$.
- As the field is non-degenerate, it is uniformly comparable with an iid vector, and

$$\sup_{x \in A} \mathbb{E}\Big[\prod_i |f_i(x)|^{-\mu p}\Big] < \infty.$$

Corollary

Let φ, ψ two centred Gaussian fields $\mathbb{R}^{m'} \to \mathbb{R}^m$. Then

$$\mathbb{P}(\inf_x \|\varphi(x)\|/\|\psi(x)\| < 1) \leqslant \inf_\tau \left(c\tau^\alpha + \exp(-\tau \sup_{x\in A} \mathsf{Var}(\psi(x))^{-1})\right)$$

Apply this result to the field

$$\varphi(x, \hat{y}) = (\nabla F(x), D^2 F(x, \hat{y})) \in \mathbb{R}^{2d},$$
$$\psi(x, \hat{y}) = (\nabla \Delta(x), D^2 \Delta(x, \hat{y}))$$

on $A = \{(x, y) \colon x_1 = y_1\} \subseteq \mathbb{R}^{2d-1}$. Then,

$$\|\varphi\| \leqslant \|d^{\leqslant 2} F\|, \|\psi\| \leqslant \|d^{\leqslant 2}\Delta\|.$$

Proposition (Non-degeneracy of partial derivatives)

For all x, the derivatives $\partial^\alpha F(x)$ form a non-degenerate Gaussian vector.

Sketch. By continuity, if $\sum_i \lambda_i \partial^{\alpha_i} F(0) = 0$ a.s. Then, for some polynomial P,

$$0 = \mathsf{Var}(\sum_i \lambda_i \partial^{\alpha_i} F(0)) = \int P(\lambda_1, \ldots, \lambda_d) \rho(\lambda) d\lambda$$

- $\rightsquigarrow$ $\mathsf{supp}(\rho) \subseteq \{\lambda \colon P(\lambda) = 0\}$
- $\rightsquigarrow$ $\{\lambda \colon P(\lambda) = 0\}$ contains an open set
- $\rightsquigarrow$ P vanishes in a neighbourhood of some point

- Remains to bound the derivatives

Define

$$\Delta_B(x) := F(x) - F^{(B)}(x) := F(x) - g \star \mathcal{W}^{(B)}(x)$$

Proposition

Let $|\alpha| \leqslant 3$. *If for some* $\beta > d$, $|\partial_\alpha g(x)| \in O\big((1+|x|)^{-\beta}\big)$, *for* $A \subseteq \mathbb{R}^d$

$$\sup_A |\partial_\alpha \Delta_B(x)| \leqslant c(\log(\mathrm{diam}(A)) + U_A) d(A,B)^{-\gamma d}).$$

where U_A *has a Gaussian tail.*

First, by the Borell-TIS inequality and the entropy bound we have the following result

Proposition

Given any continuous centered Gaussian field G *on some domain* $A \subseteq \mathbb{R}^d$,

$$\|G\|_A := \sup_{x \in A} |G(x)| \leqslant (U_A + c(1 + \log(\operatorname{diam}(A) + 1))) \sup_A \sqrt{\mathsf{Var}(G(x))}$$

where $\mathbb{P}(U_A \geqslant s) \in O(\exp(-cs^2))$

Proof idea.

Let $\sigma = d(A, B)$. Then,

$$\mathsf{Var}(\Delta_B(x)) = 2\int_B |g(x-y)|^2 dy \leqslant \int_{B(0,\sigma)^c} (1 + \|x-y\|)^{-2\beta} dy \leqslant c\sigma^{-2\gamma d}.$$

Hence, the maximum over some A behaves as

$$\|\Delta_B\|_A \in O\big((\log(1 + \operatorname{diam}(A)) + U_A)(1+\rho)^{-\gamma d}\big)$$

where U_A has Gaussian tails.

Idea. By martingale CLT,

$$\sigma(u)^2 = \mathbb{E}\big[\mathbb{E}[B_{\Delta,o}^{(\infty)} \,|\, \mathcal{F}_0]\big],$$

where $\mathcal{F}_0 := \sigma\big(\mathcal{W} \cap Q_i \colon Q_i \prec Q_1\big)$

Key step. Expected Betti numbers decrease after perturbation

Lemma (Reduction of expected Betti number by perturbation)

There exists $m \geqslant 1$ *and an open set* $S \subseteq \mathbb{R}$ *such that*

$$\sup_{s \in S} \lim_{n \uparrow \infty} \big(\mathbb{E}[\beta_n(F + s(g \star \mathbb{1}_{mQ_0}), u)] - \mathbb{E}[\beta_n(F, u)]\big) < 0$$

For $w \in \mathcal{C}^4$, we set

$$D(w) := \lim_{n\uparrow\infty} \big(\beta_n(F+w;u) - \beta_n\big(F^{(Q_1)};u\big)\big).$$

In particular, $D(0) = B^{(\infty)}_{\Delta,o}$.

Positivity of the limiting variance (sketch).

- In the proof, we rely on a variant $B^{(\infty)}_{\Delta,0;m}$ of $B^{(\infty)}_{\Delta,o}$, where instead of a partition into side-length 1 boxes, we use boxes of side length $m \geqslant 1$.
- Decompose the white noise $\mathcal{W}$ on mQ_0 into $Z_0 \mathbb{1}_{mQ_0}(\cdot)$ and an orthogonal part, where Z_0 is a standard normal random variable.
- Consider $G(z) := \mathbb{E}[B^{(\infty)}_{\Delta,0;m} \,|\, Z_0 = z]$. Then, $\mathbb{E}[G(Z_0+s)] = \mathbb{E}[D(s(g \star \mathbb{1}_{mQ_0}))]$, so that $\mathbb{E}[G(Z_0)] = 0$.
- Reduction lemma $\rightsquigarrow \inf_{s\in\mathbb{R}} \mathbb{E}[G(Z_0+s)] < 0$.
- Jensen's inequality $\rightsquigarrow \sigma^2 \geqslant \mathsf{Var}(G(Z_0)) > 0$.

Key observation. $\lim_{u\uparrow\infty}\mu(u)=0$.

Reduction lemma $\rightsquigarrow \exists\zeta>0$ and $S\subseteq\mathbb{R}^d$ open such that $\sup_{s\in S}\mu(u-s)-\mu(u)<-7\zeta$

$$\rightsquigarrow \mathbb{E}\big[\beta\big(B(k),F+s\big)\big]-\mathbb{E}\big[\beta\big(B(k),F\big)\big]<-6\zeta k^d.$$

Four steps, valid for $n>m$, $k:=m-\sqrt{m}$, $s\in S$ and where we set $w:=w_{s,m}:=sg\star\mathbb{1}_{mQ_0}$:

1. $\mathbb{E}\big[\big|\beta\big(B(n),F\big)-\beta\big(B(k),F\big)-\beta\big(B(n)\setminus B(k),F\big)\big|\big]\leqslant\zeta m^d$
2. $\mathbb{E}\big[\big|\beta\big(B(n),F+w\big)-\beta\big(B(k),F+w\big)-\beta\big(B(n)\setminus B(k),F+w\big)\big|\big]\leqslant\zeta m^d$
3. $\mathbb{E}\big[\big|\beta\big(B(n)\setminus B(k),F+w\big)-\beta\big(B(n)\setminus B(k),F\big)\big|\big]\leqslant\zeta m^d$
4. $\mathbb{E}\big[\big|\beta\big(B(k),F+w\big)-\beta\big(B(k),F+s\big)\big|\big]\leqslant\zeta m^d$.

$$\rightsquigarrow \sup_{s\in S}\lim_{n\uparrow\infty}\big(\mathbb{E}[\beta_n(F+s(g\star\mathbb{1}_{mQ_0}))]-\mathbb{E}[\beta_n(F)]\big)\leqslant-\zeta m^d,$$

Expression I.

Betti numbers of components contained in $B(k)$ are taken into account in $\beta(B(n), F)$. Also the Betti numbers of components contained in $B(n) \setminus B(k)$ are accounted for in $\beta(B(n), F)$.

$\rightsquigarrow$ deviations come from components intersecting $\partial B(k)$.

- Stationarity $\rightsquigarrow$ suffices to show that $\mathbb{E}\big[\beta\big(\mathcal{C}(B(1))\big)\big] < \infty$.
- Let

$$\Omega_m := \{m \geqslant 1 \text{ is minimal such that all components hitting } B(1) \text{ are contained in } B(m)\}.$$

Then, by stationarity and Cauchy-Schwarz,

$$\begin{aligned}\mathbb{E}\big[\beta\big(\mathcal{C}(B(1))\big)\big] &\leqslant \sum_{m \geqslant 1} \mathbb{E}\big[Y(B(m))\mathbb{1}\{\Omega_m\}\big] \\ &\leqslant \sqrt{\mathbb{E}\big[Y(B(1))^2\big]} \sum_{m \geqslant 1} |B(m)|\sqrt{\mathbb{P}(\Omega_m)} < \infty.\end{aligned}$$

- First, decompose

$$\beta(B(k), F + w) =: \sum_{j \in B(k) \cap \mathbb{Z}^d} B_j(F + w).$$

- Next,

$$\sup_{x \in \mathbb{R}^d} |w(x)| \leqslant |s| \sup_{x \in \mathbb{R}^d} \int_{B(m)} |g(x - y)| \mathrm{d}y \leqslant |s| \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |g(x - y)| \mathrm{d}y < \infty,$$

  and similarly for the derivatives.

- Now, by Hölder, bound $\mathbb{E}[|B_j(F + w) - B_j(F)|]$ by

$$\mathbb{P}\big(B_j(F) \neq B_j(F + w)\big)^{1/p} \Big(\underbrace{\mathbb{E}[|B_j(F)|^q]}_{<\infty}{}^{1/q} + \underbrace{\mathbb{E}[|B_j(F + w)|^q]}_{<\infty}{}^{1/q}\Big).$$

$$\mathbb{E}\big[\big|\beta(B(n)\setminus B(k), F+w) - \beta(B(n)\setminus B(k), F)\big|\big]$$
$$\leqslant \sum_{j:\ k<|j|\leqslant m} \mathbb{E}[|B_j(F+w) - B_j(F)|] + \sum_{j:\ |j|\geqslant m+\sqrt{m}} \mathbb{E}[|B_j(F+w) - B_j(F)|]$$

Hence, need to show that

$$\max_{j:\ k<|j|\leqslant m} \mathbb{P}(B_j(F+w) \neq B_j(F)) \in o(1),$$
$$\sum_{j:\ |j|\geqslant m+\sqrt{m}} \mathbb{P}(B_j(F+w) \neq B_j(F))^{1/p} \in o(m^d).$$

Here, we consider the perturbations $F^{(t)} := F + tw$. Then, for $x \in B(m)^c$,

$$|w(x)| = \Big|\int_{B(m)} g(x-u)\mathrm{d}u\Big| \leqslant \int_{|u|>\mathrm{dist}(x,B(m))} |g(u)|\mathrm{d}u \in O(\mathrm{dist}(x, B(m))^{d-\beta}).$$

Thus, by Bulinskaya, $\mathbb{P}(B_j(F+w) \neq B_j(F)) = \mathrm{dist}(x, B(m))^{d-\beta+o(1)}$. Also,

$$\sum_{j:\ |j|\geqslant m+\sqrt{m}} \mathrm{dist}(j, B(m))^{-(\beta-d)/q_{\mathsf{M}}} \leqslant c_3 \sum_{i\geqslant m+\sqrt{m}} i^{d-1}(i-m)^{-(\beta-d)/q_{\mathsf{M}}} \in O(m^{d-1/2}),$$

We prove tightness by verifying the Chentsov condition,

$$\mathbb{E}\big[\tilde{\beta}_n(I)^4\big] \leqslant cn^{2d}|I|^{5/4},$$

where $\tilde{\beta}_n([u_-, u_+]) := \tilde{\beta}_n(u_+) - \tilde{\beta}_n(u_-)$

Cumulant expansion

$$\mathbb{E}\big[\tilde{\beta}_n(I)^4\big] = 3\mathsf{Var}\big(\tilde{\beta}_n(I)\big)^2 + c_4\big(\tilde{\beta}_n(I)\big).$$

Key step. Show that

$$\sup_{n\geqslant 1}\ \sup_{I \textit{ is } n\textit{-big}} |W_n|^{-1}|I|^{-5/8}\mathsf{Var}\big(\beta_n(I)\big) + |W_n|^{-7/6}c_4\big(\beta_n(I)\big) < \infty,$$

where I is n-***big*** if $|I| \geqslant |W_n|^{-2/3}$ and to show that

Proof of tightness. Note that

$$\mathbb{E}\big[\tilde{\beta}_n(I)^4\big] \leqslant 3c|W_n|^2|I|^{5/4} + c|W_n|^{7/6}.$$

Then, $|W_n|^{7/6} \leqslant |W_n|^2|I|^{5/4}$ because $|I|$ is n-big.

Lemma (Moment bounds)

Let $M \geqslant 1$*. Then,* $\mathbb{E}\big[Y(\mathcal{C}_0 \times \mathbb{R})^M\big] < \infty$ *and*

$$\sup_{I \subseteq I_\mathsf{b}} \sup_{i \in \mathbb{Z}^d} |I|^{-31/32} \mathbb{E}\big[Y(\mathcal{C}_i \times I)^M\big] < \infty \textit{ for every compact } I_\mathsf{b} \subseteq \mathbb{R}$$

Part (1). Let $K \geqslant 1$ be minimal with $\mathcal{C}_0 \subseteq B(0, K)$.

- Then, by Cauchy-Schwarz and stationarity,

$$\begin{aligned}\mathbb{E}\big[Y(\mathcal{C}_0 \times \mathbb{R})^M\big] &\leqslant \sum_{k \geqslant 1} \mathbb{E}\big[Y(B(0,k) \times \mathbb{R})^M \mathbb{1}\{K = k\}\big] \\ &\leqslant \sum_{k \geqslant 1} \sqrt{\mathbb{E}\big[Y(B(0,k) \times \mathbb{R})^{2M}\big]} \sqrt{\mathbb{P}(K = k)} \\ &\leqslant \sqrt{\mathbb{E}\big[Y(B(0,1) \times \mathbb{R})^{2M}\big]} \sum_{k \geqslant 1} (2k)^{M+1} \sqrt{\mathbb{P}(K = k)}.\end{aligned}$$

Part (2). Proceed similarly as in the proof of part (1). Moreover,

$$\mathbb{E}\big[Y(B(0,1) \times I)^{pM}\big] \leqslant \underbrace{\mathbb{E}[Y(B(0,1) \times \mathbb{R})^{pM^2}]^{1/M}}_{<\infty} \underbrace{\mathbb{E}[Y(B(0,1) \times I)]^{1/p}}_{O(|I|^{1/p})}.$$

Idea. Consider the approach from (Davydov, ’96)

- ▷ Write β_n as $\beta_n(u) = \beta_n^+(u) - \beta_n^-(u)$ with $\beta_n^+(u)$ and $\beta_n^-(u)$ decreasing in u.
- ⇝ may assume that β_n is decreasing (sum of tight processes is tight)

Need to check that $\mathbb{E}[\beta_n(I)] \in o\big(\sqrt{|W_n|}\big)$ for n-small I

$$\mathbb{E}\big[\beta_n(I)\big] \leqslant \sum_{j \in \bar{W}_n} \mathbb{E}\big[Y\big(Q_j \times I\big)\big] = |\bar{W}_n|\mathbb{E}\big[Y\big(Q_o \times I\big)\big] \in O\big(|W_n||I|^{7/8}\big).$$

Proof of cumulant bound. Martingale decomposition

- ▷ $\mathcal{G}_j :=$ σ-algebra generated by the restriction of $\mathcal{W}$ to boxes of the form Q_i with i preceding j in the lexicographic order $\leqslant_{\text{lex}}$.

Then, with $B^{(n)}_{\Delta,j}(I) := \beta_n(I;F) - \beta_n(I;F^{(Q_j)})$,

$$\beta_n(I) - \mathbb{E}[\beta_n(I)] = \sum_{j \in \mathbb{Z}^d} \mathbb{E}[B^{(n)}_{\Delta,j}(I) \,|\, \mathcal{G}_j]$$

First, since $\{B^{(n)}_{\Delta,i}\}_{i\in\mathbb{Z}^d}$ is a martingale-difference sequence,

$$\mathsf{Var}(\beta_n(I)) = \sum_{j\in\mathbb{Z}^d} \mathsf{Var}\big(\mathbb{E}[B^{(n)}_{\Delta,j}(I)\,|\,\mathcal{G}_j]\big) \leqslant \sum_{j\in\mathbb{Z}^d} \mathbb{E}\big[B^{(n)}_{\Delta,j}(I)^2\big]$$

Now, $\mathbb{E}\big[B_{\Delta,i,j}(I)^2\big] \leqslant \mathbb{E}\big[B_{\Delta,i,j}(I)^{2M}\big]^{1/M}\mathbb{P}(B_{\Delta,i,j}(I)\neq 0)^{1/p'}$ where

$$\mathbb{P}(B_{\Delta,i,j}(I)\neq 0) = |i-j|^{-\gamma d/5+o(1)}|I|^{4/5}$$

$$\rightsquigarrow \mathbb{E}\big[B^{(n)}_{\Delta,j}(I)^2\big]^{1/2} \leqslant \sum_{i\in\bar{W}_n} \mathbb{E}\big[B_{\Delta,i,j}(I)^2\big]^{1/2} \leqslant c\sum_{i\in\bar{W}_n} |i-j|^{-\gamma d/8}|I|^{3/8}.$$

- ▷ If $d_{j,n} := \mathsf{dist}(j,W_n) \leqslant n$, then $\mathbb{E}\big[B^{(n)}_{\Delta,j}(I)^2\big] \in O(|I|^{3/4})$.
- ▷ If $\mathsf{dist}(j,W_n) \geqslant n$, then

$$\mathbb{E}\big[B^{(n)}_{\Delta,j}(I)^2\big] \in O(d_{j,n}^{2(1-\gamma/8)d}|I|^{3/4}).$$

Finally,

$$\sum_{j:\ \mathsf{dist}(j,W_n)\geqslant n} d_{j,n}^{2(1-\gamma/8)d} \in O(|W_n|^{3-\gamma/4}).$$

First, by multilinearity of cumulants,

$$c_4\big(\tilde{\beta}_n(I)\big) = \sum_{i,j,k,\ell \in \mathbb{Z}^d} a_{i,j,k,\ell} c_4(B^{(n)}_{\Delta,i}, B^{(n)}_{\Delta,j}, B^{(n)}_{\Delta,k}, B^{(n)}_{\Delta,\ell}),$$

where each summand is bounded as

$$\big|c_4(B^{(n)}_{\Delta,i}, B^{(n)}_{\Delta,j}, B^{(n)}_{\Delta,k}, B^{(n)}_{\Delta,\ell})\big| \in O\Big(\prod_{m \in \{i,j,k,\ell\}} \mathbb{E}[(B^{(n)}_{\Delta,m})^4]^{1/4}\Big).$$

We now distinguish indices in $\mathcal{I} := \{i\colon d_{i,n} \leqslant n\}$ and those outside $\mathcal{I}$. First, $\mathbb{E}[|B^{(n)}_{\Delta,i}|^4] = |W_n|^5 d_{i,n}^{-\gamma d + o(1)}$. In particular,

$$\sum_{i\colon d_{i,n} \geqslant n} \mathbb{E}[|B^{(n)}_{\Delta,i}|^4]^{1/4} = |W_n|^{1+5/4-\gamma/4+o(1)},$$

so that

$$\sum_{\mathcal{I}^c} |c_4(B^{(n)}_{\Delta,i}, B^{(n)}_{\Delta,j}, B^{(n)}_{\Delta,k}, B^{(n)}_{\Delta,\ell})| = |W_n|^{4+5/4-\gamma/4+o(1)}.$$

Remains to deal with indices in $\mathcal{I}$. Here, partition the sum into $\Sigma_1 + \Sigma_2$.

- $\mathfrak{d}(\{i,j,k,\ell\}) := \max_{\{S,T\} \prec \{i,j,k,\ell\}} \mathrm{dist}\big(\{s\}_{s\in S}, \{t\}_{t \in T}\big)$
- Σ_1 contains indices with $\mathfrak{d}(\{i,j,k,\ell\}) \leqslant |W_n|^{o(1)}$
- Σ_2 contains rest
- $\Sigma_1 \rightsquigarrow O(|W_n|^{1+o(1)})$ ***summands;*** each in $O(1)$
- $\Sigma_2 \rightsquigarrow O(|W_n|^4)$ ***summands***.

Key step. Show that each of the Σ_2-factors decays at speed $|W_n|^{-\gamma/d+o(1)}$

Lemma (Spatial decorrelation)

It holds that

$$\left|c_4\big(B^{(n)}_{\Delta,i}, B^{(n)}_{\Delta,j}, B^{(n)}_{\Delta,k}, B^{(n)}_{\Delta,\ell}\big)\right| = \mathfrak{d}(\{i,j,k,\ell\})^{-\gamma+o(1)}.$$

Idea. Cluster-decomposition of the cumulant derived in (Penrose & Yukich, 2001).
Consider $k_0 := \mathfrak{d}(\boldsymbol{z}) := \mathrm{dist}\big(\{i,k\},\{j,\ell\}\big)$. Then, need to bound

$$\mathrm{Cov}\big(B^{(n)}_{\Delta,i}B^{(n)}_{\Delta,j}, B^{(n)}_{\Delta,k}B^{(n)}_{\Delta,\ell}\big) = \sum_{\boldsymbol{z}'} \mathrm{Cov}\big(B_{\Delta,i',i}B_{\Delta,k',k}, B_{\Delta,j',j}B_{\Delta,\ell',\ell}\big).$$

- Now, use a ***resampling representation***

For $\tilde B_{\Delta,i',i}$ resample in half-space of points closer to i than to j.

$$\rightsquigarrow \Big|\mathrm{Cov}\big(B_{\Delta,i',i}B_{\Delta,k',k}, B_{\Delta,j',j}B_{\Delta,\ell',\ell}\big)\Big|$$
$$= \Big|\mathbb{E}\Big[B_{\Delta,i',i}B_{\Delta,k',k}B_{\Delta,j',j}B_{\Delta,\ell',\ell} - \tilde B_{\Delta,i',i}\tilde B_{\Delta,k',k}\tilde B_{\Delta,j',j}\tilde B_{\Delta,\ell',\ell}\Big]\Big|,$$

Then, by Hölder,

$$\Big|\mathrm{Cov}\big(B_{\Delta,i',i}B_{\Delta,k',k}, B_{\Delta,j',j}B_{\Delta,\ell',\ell}\big)\Big|$$
$$\leqslant 2\mathbb{P}\big(B_{\Delta,i',i}B_{\Delta,k',k}B_{\Delta,j',j}B_{\Delta,\ell',\ell} \neq \tilde B_{\Delta,i',i}\tilde B_{\Delta,k',k}\tilde B_{\Delta,j',j}\tilde B_{\Delta,\ell',\ell}\big)^{1/p'}$$
$$\times\, \mathbb{E}\Big[\big|B_{\Delta,i',i}\big|^{4M}\Big]^{1/(4M)} = \mathfrak{d}(z)^{-\gamma+o(1)}.$$

Outlook

- CLTs for excursion set of Gaussian sets
  - Functional in the level ✓
- Level sets **!**
- Shot-noise fields **!**
- Critical/Super-critical regime **!**
- Persistent Betti numbers **!**

[1] C. A. N. Biscio, N. Chenavier, C. Hirsch, and A. M. Svane.
Testing goodness of fit for point processes via topological data analysis.
Electron. J. Stat., 14(1):1024–1074, 2020.

[2] C. A. N. Biscio and A. M. Svane.
A functional central limit theorem for the empirical Ripley's K-function.
Electron. J. Stat., 16(1):3060–3098, 2022.

[3] L. H. Y. Chen, A. Röllin, and A. Xia.
Palm theory, random measures and Stein couplings.
Ann. Appl. Probab., 31(6):2881–2923, 2021.

[4] A. D. Christoffersen, J. Møller, and H. S. Christensen.
Modelling columnarity of pyramidal cells in the human cerebral cortex.
Aust. N. Z. J. Stat., 63(1):33–54, 2021.

[5] T. Cong and A. Xia.
Convergence rate for geometric statistics of point processes with fast decay dependence.
arXiv preprint arXiv:2205.13211, 2022.

[6] C. Hirsch, J. Krebs, and C. Redenbach.
Persistent homology based goodness-of-fit tests for spatial tessellations.
arXiv preprint arXiv:2209.13151, 2022.

[7] J. Møller and R. P. Waagepetersen.
Statistical Inference and Simulation for Spatial Point Processes.
CRC, Boca Raton, 2004.