

Geometry and topology of smooth Gaussian fields

Michael McAuley
Technological University Dublin

Probability seminar,
University College Dublin,
16th April 2025

Slides available at
<https://michael-mcauley.github.io>

Outline

1. Introduction

2. Geometric functionals

3. Topological functionals

Smooth Gaussian fields

Basic setting

Definition

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a random function. We say that f is a **Gaussian field** if for any $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{R}^d$, $(f(x_1), \dots, f(x_n))$ is a normal random vector.

The field is **smooth** if, with probability one $f \in C^2$.

The field is **stationary** if its distribution is invariant under translations.

- ▶ Given a stationary Gaussian field, we may normalise so that for all $x \in \mathbb{R}^d$, $f(x) \sim \mathcal{N}(0, 1)$.
- ▶ We can construct such a field as

$$f = \sum_{n=1}^{\infty} Z_n f_n$$

where $Z_n \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ and the f_n are deterministic C^2 functions (satisfying appropriate conditions).

- ▶ Analogously with Gaussian vectors, the distribution of f is specified by its covariance function $K : \mathbb{R}^d \rightarrow \mathbb{R}$ defined as

$$K(x - y) = \text{Cov}[f(x), f(y)] \quad \forall x, y \in \mathbb{R}^d.$$

Smooth Gaussian fields

Excursion sets

We will consider the geometry/topology of the **excursion sets**

$$\{f \geq \ell\} := \left\{x \in \mathbb{R}^d \mid f(x) \geq \ell\right\} \quad \text{for } \ell \in \mathbb{R}.$$

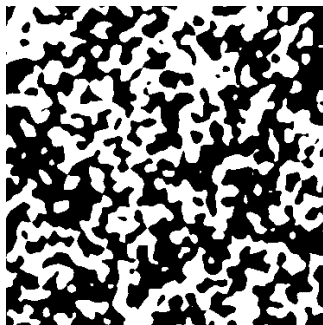
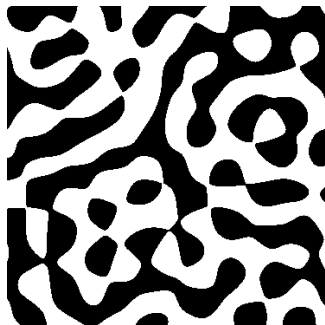


Figure: Excursion sets $\{f \geq 0\}$ in white for the fields on \mathbb{R}^2 with covariance functions $K(x) = J_0(|x|)$, the 0-th Bessel function, (left) and $K(x) = \exp(-|x|^2/2)$ (right).

Motivation

1) Studying classes of functions

A Gaussian field can be viewed as a measure on a particular class of functions. Statements about the field can be interpreted as statements about 'typical' functions in the class.

1. **Berry's conjecture:** on chaotic 2-dimensional manifolds, high-frequency eigenfunctions of the Laplacian can be approximated by the Gaussian field with $K(x) = J_0(|x|)$ [6].
2. **Hilbert's 16th problem** concerns the zero set of homogeneous polynomials. There is a canonical Gaussian measure on such polynomials which behaves locally like the stationary field with $K(x) = \exp(-|x|^2/2)$ [10].

Motivation

2) Percolation theory

- ▶ Percolation theory studies the long-range connectivity properties of random models.
- ▶ **Bernoulli percolation on the square lattice:** adjacent points of \mathbb{Z}^2 are joined by an edge independently with probability p .

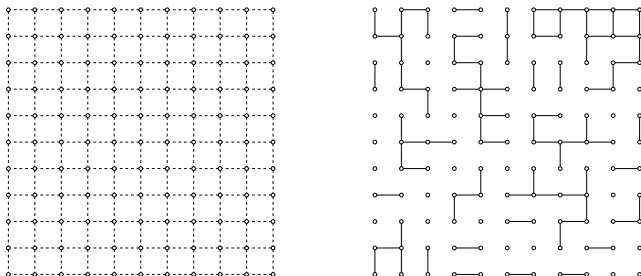


Figure: A section of the square lattice \mathbb{Z}^2 (left) and a realisation of the Bernoulli percolation model with $p = 0.4$ on this section (right).

Motivation

2) Percolation theory

- ▶ Progress has been made recently in studying percolation of Gaussian excursion sets [2].
- ▶ **Phase transition:** for a given field, there is a critical level ℓ_c such that
 - for $\ell > \ell_c$, $\{f \geq \ell\}$ contains only bounded components,
 - for $\ell < \ell_c$, $\{f \geq \ell\}$ contains a unique unbounded component.

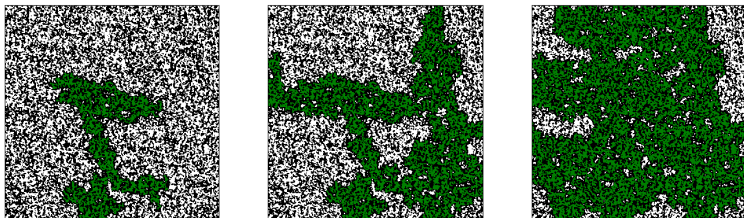


Figure: The excursion sets $\{f \geq \ell\}$ for $\ell = 0.05$ (left), $\ell = 0$ (middle) and $\ell = -0.05$ (right). Largest component highlighted in green.

Motivation

3) Statistical applications

- ▶ Gaussian fields arise in many areas of science:
 - Medical imaging [18],
 - Cosmology [16],
 - Topological data analysis [1].
- ▶ Geometric/topological properties of excursion sets can be used as test statistics. (See [17] for an overview.)

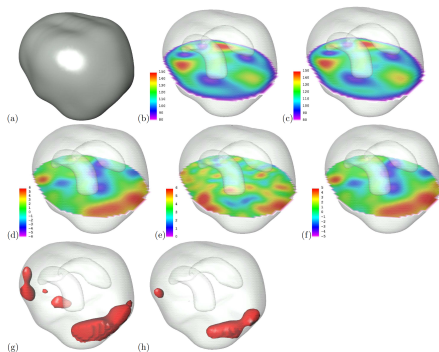


Figure: Measurements from a PET study of brain activity during a reading task. (Source: [17]).

Smooth Gaussian fields

Questions of interest

- ▶ What are the geometric and topological properties of smooth Gaussian excursion sets?
- ▶ We would like to analyse:

Geometric functionals

- volume
- boundary volume
- Euler characteristic

Topological functionals

- number of connected components
- Betti numbers

- ▶ What is the expectation, variance and distribution of such functionals on a bounded domain?
- ▶ How does this depend on the size of the domain? the level of the excursion set? the covariance of the field?

Outline

1. Introduction

2. Geometric functionals

3. Topological functionals

Geometric functionals

A rough definition

- ▶ A functional of a random field is described as **local** (or geometric) if it is an integral of a pointwise function of the field and its derivatives:

$$\int_D \varphi(f(x), \nabla f(x), \nabla^2 f(x), \dots) \nu(dx)$$

- ▶ We will consider functionals of the form

$$F_R = \int_{[-R,R]^d} \varphi(f(x) - \ell) dx$$

for some $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ (e.g. $\varphi(y) = \mathbb{1}_{y \geq 0}$).

- ▶ How does this behave as $R \rightarrow \infty$?
- ▶ First order behaviour is trivial: by Fubini's theorem,

$$\mathbb{E}[F_R] = (2R)^d \mu(\ell)$$

where $\mu(\ell) := \mathbb{E}[\varphi(f(0) - \ell)]$.

Second order properties

Hermite polynomials

The variance and limiting distribution of local functionals can be studied using Hermite polynomials.

- ▶ The Hermite polynomials $(H_n)_{n \geq 0}$ are defined inductively by setting

$$H_0(x) = 1 \quad \text{and} \quad H_{n+1}(x) = xH_n(x) - H'_n(x)$$

which yields

$$H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x.$$

- ▶ If X, Y are jointly normal with mean zero and variance one then

$$\mathbb{E}[H_n(X)H_m(Y)] = \begin{cases} n! \text{Cov}[X, Y]^n & \text{if } n = m \\ 0 & \text{if } n \neq m. \end{cases}$$

- ▶ If $\mathbb{E}[\varphi^2(Z)] < \infty$ for $Z \sim \mathcal{N}(0, 1)$ then

$$\varphi = \sum_{n=0}^{\infty} a_n H_n$$

where $\sum_n a_n^2 n! < \infty$.

Second order properties

Orthogonal decomposition

- Considering the expansion $\varphi(\cdot - \ell) = \sum_n a_n(\ell) H_n$ yields

$$F_R = \sum_{n=0}^{\infty} a_n(\ell) \int_{[-R,R]^d} H_n(f(x)) dx =: \sum_{n=0}^{\infty} Q_n.$$

- The variance of F_R can be computed by considering

$$\begin{aligned} \text{Cov}[Q_n, Q_m] &= a_n(\ell) a_m(\ell) \iint_{[-R,R]^{2d}} \text{Cov}[H_n(f(x)), H_m(f(y))] dx dy \\ &= \begin{cases} a_n(\ell)^2 n! \iint_{[-R,R]^{2d}} K(x-y)^n dx dy & \text{if } n = m \neq 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

- Hence $\text{Var}[F_R] = \sum_n \text{Var}[Q_n]$ which depends on the integrability of K .

Second order properties

Covariance function examples

Three general classes of covariance function are considered in the literature:

1. K is **integrable**

- Example: the Bargmann-Fock field has covariance

$$K(x) = \exp\left(-\frac{|x|^2}{2}\right).$$

2. K is **regularly varying** at infinity with index $\alpha \in (0, d)$

- Example: The Cauchy field has covariance

$$K(x) = (1 + |x|^2)^{-\alpha/2}.$$

3. K is **oscillating and slowly decaying**

- Example: The Random Plane Wave is the two-dimensional field with covariance

$$K(x) = J_0(|x|) \sim \sqrt{\frac{2}{\pi}} \cos(|x| - \pi/4) |x|^{-1/2} \quad \text{as } |x| \rightarrow \infty.$$

Second order properties

Case 1: Integrable covariance

Recall:

$$F_R = \sum_{n=0}^{\infty} Q_n, \quad \text{Var}[Q_n] = a_n(\ell)^2 n! \int_{[-R,R]^{2d}} K(x-y)^n dx dy$$

- If K is integrable then for $n \neq 0$

$$\text{Var}[Q_n] \sim \left(a_n(\ell)^2 n! \int_{\mathbb{R}^d} K(x)^n dx \right) (2R)^d$$

so each chaos has variance of order R^d (or 0).

- Since $\sum_n a_n(\ell)^2 n! < \infty$, for each ℓ

$$\text{Var}[F_R] \sim c_\ell R^d \text{ as } R \rightarrow \infty$$

where $c_\ell > 0$ assuming φ is not too degenerate.

Second order properties

Case 1: Integrable covariance

Using similar, but more involved, computations one can compute the higher order moments of F_R to prove:

Theorem (Breuer-Major theorem)

If f has rotation invariant distribution, then as $R \rightarrow \infty$

$$\frac{F_R - \mu(\ell)}{\text{Var}[F_R]} \xrightarrow{d} \mathcal{N}(0, 1)$$

Remark

More modern proofs of this result use the Malliavin-Stein method (in particular the fourth-moment theorem) which also yields a rate of convergence.

Second order properties

Case 2: Regularly varying covariance

- ▶ If $K(x) \sim c|x|^{-\alpha}$ then for $n \neq 0$

$$\begin{aligned}\text{Var}[Q_n] &= a_n(\ell)^2 n! \int_{[-R,R]^d} K(x-y)^n dx dy \\ &\sim a_n(\ell)^2 c_{K,n} \times \begin{cases} R^{2d-n\alpha} & \text{if } n\alpha < d, \\ R^d \log R & \text{if } n\alpha = d, \\ R^d & \text{if } n\alpha > d. \end{cases}\end{aligned}$$

- ▶ Hence a finite number of the Q_n terms have higher orders of variance.
- ▶ Since $\sum_n a_n(\ell)^2 n! < \infty$

$$\sum_{n > d/\alpha} \text{Var}[Q_n] \sim c_\ell R^d$$

and so $F_R = \sum_n Q_n$ will be asymptotically dominated by a single term if $a_n(\ell) \neq 0$ for some $n \leq d/\alpha$.

Second order properties

Case 2: Regularly varying covariance

Theorem (Dobrushin-Major theorem)

Let $n^*(\ell) = \inf\{n : a_n(\ell) \neq 0\}$. If f satisfies some technical conditions, then

$$\text{Var}[F_R] \sim c_{K,\varphi,\ell} \times \begin{cases} R^{2d-n^*\alpha} & \text{if } n^*\alpha < d, \\ R^d \log R & \text{if } n^*\alpha = d, \\ R^d & \text{if } n^*\alpha > d, \end{cases} \quad \text{as } R \rightarrow \infty.$$

Moreover if $n^* = 1$ or $n^*\alpha > d$ then

$$\frac{F_R - \mu(\ell)}{\text{Var}[F_R]} \xrightarrow{d} \mathcal{N}(0, 1).$$

For other values of n^* , the limiting distribution is a Hermite distribution.

Remark

- ▶ Typically $n^*(\ell) = 1$ for all but finitely many values of ℓ , which are described as **anomalous levels**.
- ▶ If φ is regular then $a_n(\ell) = (-1)^n \mu^{(n)}(\ell)/n!$ so that anomalous levels correspond to critical points of μ .

- ▶ The classical Breuer-Major theorem [7]. A modern proof using the Malliavin-Stein method [15].
- ▶ The Dobrushin-Major theorem [8] was first proven using multiple Wiener-Itô integrals.
- ▶ More recently, a general CLT has been proven for some fields with slowly decaying oscillating correlations [11].

Outline

1. Introduction

2. Geometric functionals

3. Topological functionals

Topological functionals

What is known?

- ▶ Much less is known about non-local/topological functionals of Gaussian fields.
- ▶ The previous approach fails due to the lack of an integral representation.
- ▶ There is no unifying theory, but many partial results using a variety of methods:
 - Law of large numbers [13] (ergodic argument)
 - Variance bounds [14, 5, 4] (coupling and interpolation methods)
 - Central limit theorem [3] (martingale limit theorem)

Topological functionals

A new approach

- ▶ In joint work with Stephen Muirhead [12], we adapt the Hermite expansion approach to non-local functionals.
- ▶ Let $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ be the Gaussian free field (in $d \geq 3$), so that

$$K(x - y) \sim c_d |x - y|^{-(d-2)}.$$

- ▶ The *cluster count* $N_R(f)$ is the number of clusters (i.e. connected components) of the graph $\{f \geq \ell\} \cap [-R, R]^d$.

Wiener chaos expansion

Abstract statement

Let H be a set of centred jointly Gaussian variables. Let \mathcal{P}_n be the space of all polynomials of degree n in H .

The n -th Wiener chaos of H is $H^{:n:} := \overline{\mathcal{P}_n} \cap \overline{\mathcal{P}_{n-1}}^\perp$.

Theorem (Wiener, Itô)

Let the random variable X be square integrable and $\sigma(H)$ -measurable, then

$$X \stackrel{L^2}{=} \sum_{n=0}^{\infty} Q_n[X]$$

where Q_n denotes projection onto $H^{:n:}$.

Remark

- ▶ While the result is very general, in practice the chaos projections can be difficult to characterise (especially if H is large).
- ▶ When H has a single element, this is just the Hermite expansion. Local functionals can be reduced to this case.

Proposition

Let $D \subset \mathbb{Z}^d$ be finite and $\Phi : \mathbb{R}^D \rightarrow \mathbb{R}$ be smooth and bounded. Then

$$Q_n[\Phi(f)] = \frac{1}{n!} \sum_{x_1, \dots, x_n \in D} \mathbb{E}[\partial_{x_1} \dots \partial_{x_n} \Phi(f)] Q_n[f(x_1) \dots f(x_n)].$$

- ▶ The proof uses Gaussian integration by parts and is quite elementary.
- ▶ The term $Q_n[f(x_1) \dots f(x_n)]$ is called a *Wick polynomial* and can be evaluated explicitly.

Wiener chaos expansion

Cluster count

Proposition

For the cluster count, the expected derivative at $\underline{x} = (x_1, \dots, x_n)$ can be replaced by

$$P_R(\underline{x}) := \mathbb{E}[d_{x_1} \dots d_{x_n} N_R(f) | f(\underline{x}) = \ell] \varphi_{f(\underline{x})}(\ell),$$

where d_{x_i} denotes the discrete derivative

$$d_{x_i} N_R(f) = N_R(\{f \geq \ell\} \cup \{x_i\}) - N_R(\{f \geq \ell\} \setminus \{x_i\}).$$

and $\varphi_{f(\underline{x})}$ is the density of $f(\underline{x})$.

Remark

For a local functional, $\partial_{x_1} \dots \partial_{x_n} \Phi(f) = 0$ unless $x_1 = \dots = x_n$ and $Q_n[f(x)^n] = H_n(f(x))$ so we reduce to the previous analysis.

Wiener chaos expansion

'Semi-locality' via percolation results

- ▶ If P_R decays rapidly away from the diagonal, then we can analyse the variance and limiting distribution on each chaos as in the local case. We can view this as '*semi-locality*' of the cluster count.

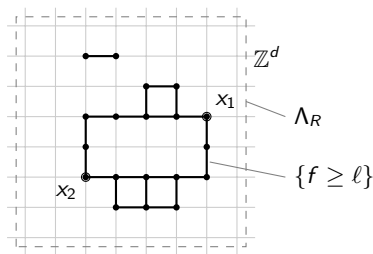


Figure: For this configuration $d_{x_1} d_{x_2} N_R(f) = 1$.

- ▶ In general, if $d_{x_1} \dots d_{x_n} N_R(f) \neq 0$ then x_1, \dots, x_n must be joined by bounded clusters of $\{f \geq \ell\}$.

Wiener chaos expansion

'Semi-locality' via percolation results

Theorem (Truncated arm decay [9])

Let $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ be the Gaussian free field for $d \geq 3$. There exists $\ell_c \in \mathbb{R}$ such that for every $\ell \neq \ell_c$, the probability that 0 is contained in a bounded cluster of $\{f \geq \ell\}$ of diameter at least n is at most e^{-cn^ρ} for some $c, \rho > 0$.

Corollary

For $\ell \neq \ell_c$ there exists $c, C, \rho > 0$ such that

$$P_R(\underline{x}) \leq Ce^{-c \text{diam}(\underline{x})^\rho}$$

where $\text{diam}(\underline{x})$ denotes the diameter of \underline{x} .

Wiener chaos expansion

Limit theorems for the cluster count

Let $\mu(\ell) = \lim_{R \rightarrow \infty} \mathbb{E}[N_R(f)] / (2R)^d$ be the *mean clusters-per-vertex*.

Theorem

Let $f : \mathbb{Z}^3 \rightarrow \mathbb{R}$ be the Gaussian free field and $\ell \neq \ell_c$.

$$\text{Var}[N_R(f)] \sim c_\ell \times \begin{cases} R^5 & \text{if } \mu'(\ell) \neq 0 \\ R^4 & \text{if } \mu'(\ell) = 0, \mu''(\ell) \neq 0 \\ R^3 \log R & \text{if } \mu'(\ell) = \mu''(\ell) = 0, \mu'''(\ell) \neq 0 \\ R^3 & \text{otherwise.} \end{cases}$$

In case 2, the (normalised) limiting distribution is a Hermite distribution, in all other cases it is Gaussian.

- ▶ Analogous results hold for $d \geq 4$ and other fields but are omitted here for brevity.
- ▶ Similar to results in local case, but the requirement that $\ell \neq \ell_c$ is new.

Open questions:

- ▶ Can this approach be extended to smooth fields?
- ▶ Does this approach enable the Malliavin-Stein method for non-local functionals?
- ▶ What happens at the critical level?

Thank you for listening!

Bibliography I

- [1] R. J. Adler et al. "Persistent homology for random fields and complexes". In: *Borrowing strength: theory powering applications—a Festschrift for Lawrence D. Brown*. Inst. Math. Stat. (IMS) Collect. Inst. Math. Statist., Beachwood, OH, 2010.
- [2] D. Beliaev. "Smooth Gaussian fields and percolation". In: *Probability Surveys* (2023). URL: <https://doi.org/10.1214/23-PS24>.
- [3] D. Beliaev, M. McAuley, and S. Muirhead. "A central limit theorem for the number of excursion set components of Gaussian fields". In: *Ann. Probab.* (2024). URL: <https://doi.org/10.1214/23-AOP1672>.
- [4] D. Beliaev, M. McAuley, and S. Muirhead. "A covariance formula for the number of excursion set components of Gaussian fields and applications". In: *Ann. Inst. Henri Poincaré Probab. Stat.* (2025). URL: <https://doi.org/10.1214/23-aihp1430>.
- [5] D. Beliaev, M. McAuley, and S. Muirhead. "Fluctuations of the number of excursion sets of planar Gaussian fields". In: *Probab. Math. Phys.* (2022). URL: <https://doi.org/10.2140/pmp.2022.3.105>.
- [6] M. V. Berry. "Regular and irregular semiclassical wavefunctions". In: *Journal of Physics A: Mathematical and General* (1977). URL: <https://dx.doi.org/10.1088/0305-4470/10/12/016>.

Bibliography II

- [7] P. Breuer and P. Major. "Central limit theorems for nonlinear functionals of Gaussian fields". In: *J. Multivariate Anal.* (1983). URL: [https://doi.org/10.1016/0047-259X\(83\)90019-2](https://doi.org/10.1016/0047-259X(83)90019-2).
- [8] R. L. Dobrushin and P. Major. "Non-central limit theorems for nonlinear functionals of Gaussian fields". In: *Z. Wahrsch. Verw. Gebiete* (1979). URL: <https://doi.org/10.1007/BF00535673>.
- [9] H. Duminil-Copin et al. "Equality of critical parameters for percolation of Gaussian free field level sets". In: *Duke Math. J.* (2023). URL: <https://doi.org/10.1215/00127094-2022-0017>.
- [10] A. Lerario and E. Lundberg. "Statistics on Hilbert's 16th Problem". In: *International Mathematics Research Notices* (2014). URL: <https://doi.org/10.1093/imrn/rnu069>.
- [11] L. Maini and I. Nourdin. "Spectral central limit theorem for additive functionals of isotropic and stationary Gaussian fields". In: *Ann. Probab.* (2024). URL: <https://doi.org/10.1214/23-aop1669>.
- [12] M. McAuley and S. Muirhead. *Limit theorems for the number of sign and level-set clusters of the Gaussian free field*. 2025. URL: <https://arxiv.org/abs/2501.14707>.

Bibliography III

- [13] F. Nazarov and M. Sodin. "Asymptotic Laws for the Spatial Distribution and the Number of Connected Components of Zero Sets of Gaussian Random Functions". In: *Journal of Mathematical Physics, Analysis, Geometry* (2016). URL: <https://jmag.ilt.kharkiv.ua/index.php/jmag/article/view/jm12-0205e>.
- [14] F. Nazarov and M. Sodin. "Fluctuations in the number of nodal domains". In: *J. Math. Phys.* (2020). URL: <https://doi.org/10.1063/5.0018588>.
- [15] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus*. Cambridge Tracts in Mathematics.
- [16] P. Pranav et al. "Topology and geometry of Gaussian random fields I: on Betti numbers, Euler characteristic, and Minkowski functionals". In: *Monthly Notices of the Royal Astronomical Society* (2019). URL: <https://doi.org/10.1093/mnras/stz541>.
- [17] K. J. Worsley. "The Geometry of Random Images". In: *CHANCE* (1996). URL: <https://www.math.mcgill.ca/keith/chance/chance3.pdf>.

Bibliography IV

- [18] K. J. Worsley et al. "A unified statistical approach for determining significant signals in images of cerebral activation". In: *Human brain mapping* (1996). URL: [https://doi.org/10.1002/\(SICI\)1097-0193\(1996\)4:1%3C58::AID-HBM4%3E3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-0193(1996)4:1%3C58::AID-HBM4%3E3.0.CO;2-0).