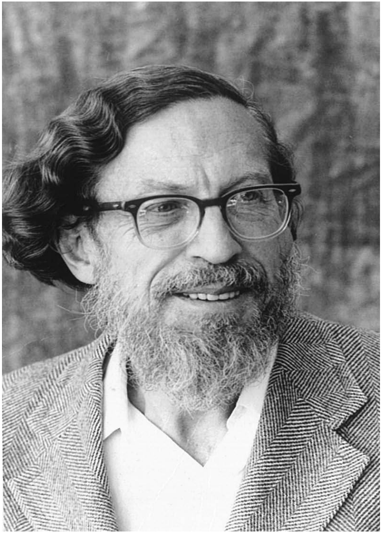
Stein's Method and Stochastic Geometry

Giovanni Peccati (Luxembourg University)

Firenze — 16 marzo 2018

INTRODUCTION



"Stein's method", as devised by **Charles Stein** at the end of the 60s, is a collection of probabilistic techniques, for measuring the **distance** between probability distributions, by means of **characterising differential operators**.

Stein's motivation was to develop an effective alternative to **Fourier methods**, for dealing with functionals of **dependent** random variables.

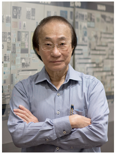
INTRODUCTION

- ⋆ Applications of Stein's method now span an enormous amount of domains, e.g.: **random matrices**, **statistics**, **biology**, **algebra**, **mathematical physics**, **finance**, **geometry**, ...

- ⋆ Main features: **quantitative**, and **"local to global"**.

- ⋆ In these lectures: my view of Stein's method, with focus on **Gaussian random fields** and **random geometric graphs**.

Some Names



L.H.Y. Chen

A.D. Barbour

E. Bolthausen

F. Götze

L. Goldstein

G. Reinert

I. Nourdin

S. Chatterjee

Conventions

- $\star$ From now on: everything is defined on a suitable triple

$$(\Omega, \mathscr{F}, \mathbb{P})$$

- $\star$ We write $N \sim \mathscr{N}(0,1)$ for a standard Gaussian random variable:

$$\mathbb{P}(N \in B) = \int_B e^{-y^2/2} \frac{dy}{\sqrt{2\pi}}.$$

- $\star$ Often: given a random element Y, we write $Y_1, Y_2, ...$ to indicate a sequence of independent copies of Y.

Theorem (CLT & Berry-Esseen bound)

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed r.v.'s, such that $\mathbb{E}[X_1] = 0$, and $\mathbf{Var}(X_1) = 1$. Write

$$S_n := X_1 + \cdots + X_n.$$

Then, as $n \to \infty$,

$$\Delta_n(z) := \mathbb{P}\left[\frac{1}{\sqrt{n}} S_n \leq z\right] - \int_{-\infty}^{z} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy \longrightarrow 0, \quad z \in \mathbb{R}.$$

Moreover,

$$\sup_z |\Delta_n(z)| \leq \frac{C\,\mathbb{E}|X_1|^3}{\sqrt{n}} \quad \Big(0.4 < C_{\text{optimal}} < 0.48\Big).$$

First Proof: Fourier (Lyapounov, Lévy)

- ⋆ Write the characteristic function $f_n(z)$ of $n^{-1/2}S_n$ as a n-product.

- ⋆ Prove that

$$f_n(z) \longrightarrow \exp\{-z^2/2\}, \quad \text{as } n \to \infty,$$

by a direct analytical argument.

Second Proof: Swapping (Lindeberg, Trotter)

- For a smooth φ, with red and blue independent, write

$$\begin{aligned}\Big|\,\mathbb{E}[\varphi(N)] - \mathbb{E}[\varphi(n^{-1/2}S_n)]\Big| \\ = \Big|\mathbb{E}[\varphi(n^{-1/2}(N_1 + \cdots + N_n))] \\ -\mathbb{E}[\varphi(n^{-1/2}(X_1 + \cdots + X_n))]\Big|.\end{aligned}$$

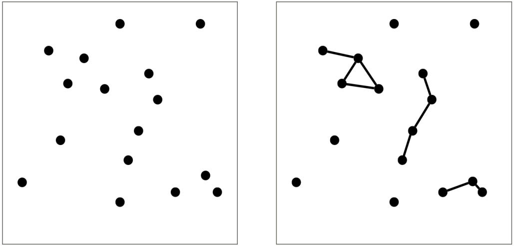
- Deduce that:

$$\begin{aligned}&\Big|\,\mathbb{E}[\varphi(N)] - \mathbb{E}[\varphi(n^{-1/2}S_n)]\Big| \\ &\leq \sum_{i=1}^{n} \Big|\mathbb{E}\varphi(n^{-1/2}(N_1 + \cdots + N_i + X_{i+1} + \cdots + X_n)) \\ &-\mathbb{E}\varphi(n^{-1/2}(N_1 + \cdots + N_{i-1} + X_i + \cdots + X_n))\Big|,\end{aligned}$$

and control each summand by a Taylor expansion.

Question

- ⋆ What happens if the summands $X_1, X_2, \ldots$ display some form of **dependence** and/or the considered random element is **not a linear mapping** ?
- ⋆ Typical example: length, number of edges / triangles / connected components / ... in a **random geometric graph**:



- ⋆ Even more extreme: random graphs arising in **combinatorial optimisation** (MST, TSP, MM, ...)

SETTING

- $\star$ In what follows, I will mainly focus on **one-dimensional normal approximations** in the 1-**Wasserstein distance**

$$\mathbf{W}_1(\bullet, \bullet).$$

- $\star$ Recall that

$$\begin{aligned} \mathbf{W}_1(X, Y) &:= \inf_{A \sim X\,;\, B \sim Y} \mathbb{E}\,|A - B| \\ &= \sup_{h \in \mathrm{Lip}(1)} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|, \end{aligned}$$

whenever $\mathbb{E}|X|, \mathbb{E}|Y| < \infty$.

INGREDIENTS

In order to implement Stein's method, one typically needs:

1. **A Lemma**
2. **A heuristic**
3. **An equation**
4. **Uniform bounds**

Stein's Lemma
Let Z be a real-valued random variable. Then, $Z \sim \mathcal{N}(0,1)$ if and only if

$$\mathbb{E}[f'(Z)] = \mathbb{E}[Z\,f(Z)],$$

for every smooth f.

[**Proof**: ($\Longrightarrow$) integration by parts. ($\Longleftarrow$) method of moments (or unicity of Fourier transform)]

Stein's Heuristic
Assume Z is a real random variable such that

$$\mathbb{E}[f'(Z)] \approx \mathbb{E}[Z f(Z)]$$

for a large class of smooth mappings f.

Then, the distribution of Z **has to be close to Gaussian** .

The Equation

- For $h \in \mathrm{Lip}(K)$ fixed and $N \sim \mathcal{N}(0,1)$, define the **Stein's equation**

$$f'(x) - xf(x) = h(x) - \mathbb{E}[h(N)], \; x \in \mathbb{R};$$

equivalent to

$$\frac{d}{dx} e^{-x^2/2} f(x) = e^{-x^2/2}(h(x) - \mathbb{E}[h(N)]).$$

- Every solution has the form

$$f(x) = ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} (h(y) - \mathbb{E}[h(N)]) e^{-y^2/2} \, dy, \; x \in \mathbb{R}.$$

- Set

$$f_h(x) := \int_{-\infty}^{x} (h(y) - \mathbb{E}[h(N)]) e^{-y^2/2} \, dy, \quad x \in \mathbb{R}.$$

By direct inspection, one proves

Stein's "Magic Factors" and Bounds

For every $h \in \mathrm{Lip}(K)$, $f_h \in \mathscr{C}^1$, and

$$\|f_h'\|_\infty \leq \sqrt{\frac{2}{\pi}}\,K.$$

As a consequence, for X integrable,

$$\begin{aligned}
\mathbf{W}_1(X, N) &= \sup_{h \in \mathrm{Lip}(1)} \Big| \mathbb{E}[h(X)] - \mathbb{E}[h(N)] \Big| \\
&= \sup_{h \in \mathrm{Lip}(1)} \Big| \mathbb{E}[f_h'(X) - X f_h(X)] \Big| \\
&\leq \sup_{f : |f'| \leq 1} \Big| \mathbb{E}[f'(X) - X f(X)] \Big|.
\end{aligned}$$

And Now ?

- ⋆ The name of the game is now to compare as sharply as possible

$$\mathbb{E}[f'(X)] \quad \text{and} \quad \mathbb{E}[Xf(X)],$$

for every smooth mapping f.

- ⋆ Several techniques: **exchangeable pairs**, **dependency graphs**, **zero-bias transforms**, **size-bias transforms**, ...

A Simple Example: Back to the CLT

- For a fixed n, write $Z := n^{-1/2}(X_1 + \cdots + X_n)$, and $Z^i = Z - n^{-1/2}X_i$.
- One has, by Taylor and independence,

$$\begin{aligned} \mathbb{E}[X_i f(Z)] &= \mathbb{E}[X_i(f(Z) - f(Z^i))] \approx \mathbb{E}[X_i(Z - Z^i) f'(Z)] \\ &= n^{-1/2}\mathbb{E}[X_i^2 f'(Z)]. \end{aligned}$$

- It follows that

$$\mathbb{E}[Zf(Z)] = n^{-1/2}\sum_i \mathbb{E}[X_i f(Z)] \approx \mathbb{E}[n^{-1}\sum_i X_i^2 \times f'(Z)].$$

- By the law of large numbers, $\mathbb{E}[Zf(Z)] \approx \mathbb{E}[f'(Z)]$ for n large, and using Stein's bounds one deduces the CLT.

Second Order Poincaré Estimates

- ⋆ Assume now $g = (g_1, ..., g_d) \sim \mathcal{N}_d(0, \mathrm{Id.})$, and define

$$F = \psi(g_1, ..., g_d),$$

for some smooth $\psi : \mathbb{R}^d \to \mathbb{R}$ s.t. $\mathbb{E}[F] = 0$ and $\mathbf{Var}(F) = 1$.

- ⋆ Remember the **Poincaré inequality** :

$$\mathbf{Var}(F) \leq \mathbb{E}[\| \nabla \psi(g) \|^2].$$

Second Order Poincaré Estimates

- $\star$ It turns out that F verifies an exact integration by parts formula:

$$\mathbb{E}[Ff(F)] = \mathbb{E}\Big[f'(F)\langle \nabla\psi(g), -\nabla L^{-1}\psi(g)\rangle\Big],$$

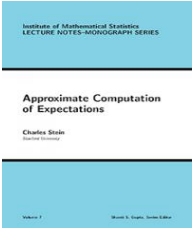
where L^{-1} is the **pseudo-inverse of the Ornstein-Uhlenbeck generator** $L = -\langle x, \nabla\rangle + \Delta$.

- $\star$ Plugging this into Stein's bound and applying once more Poincaré yields that, for $N \sim \mathcal{N}(0,1)$,

$$\begin{aligned} \mathbf{W}_1(F,N) &\leq \sqrt{\mathbf{Var}(\langle \nabla\psi(g), -\nabla L^{-1}\psi(g)\rangle)} \\ &\leq 2\mathbb{E}[\| \operatorname{Hess}\psi(g)\|_{op}^4]^{1/4} \times \mathbb{E}[\|\nabla\psi(g)\|^4]^{1/4}. \end{aligned}$$

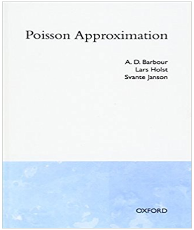
- $\star$ This is a **second order Poincaré inequality**, — see *Chatterjee* (2007), *Nourdin, Peccati and Reinert* (2010), and *Vidotto* (2017). Applications in random matrix theory & analysis of fractional fields.

Beyond Gaussian



Stein's approach extends to much more general densities — for instance to elements of the **Pearson family**. See Stein's 1986 monograph.

In the discrete setting, the equivalent of Stein's method is the **Chen-Stein method**. See the monograph by Barbour, Holst and Janson (1990).



TWO EXAMPLES

In what follows, I will illustrate two striking applications of Stein's method, that are relevant in a **geometric setting**:

(1) capturing the fluctuations of **chaotic random variables**, and

(2) quantifying **second order interactions**.

Both are connected to (generalized) **integration by parts formulae**.

Berry's Random Waves (1977)

- ⋆ Let $E > 0$. The **Berry's random wave model** on $\mathbb{R}^2$, with parameter E, written

$$B_E = \{B_E(x) : x \in \mathbb{R}^2\},$$

is defined as the unique (in law) centred, isotropic Gaussian field on $\mathbb{R}^2$ such that

$$\Delta B + E \cdot B = 0, \text{ where } \Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}.$$

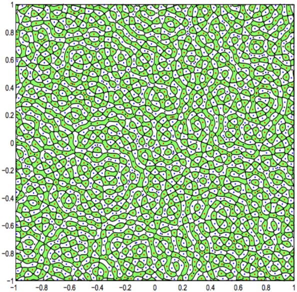
- ⋆ Equivalently, $\mathbb{E}[B_E(x)B_E(y)] = J_0(\sqrt{E}\|x - y\|)$ (J_0 = Bessel function of the 1st kind).
- ⋆ Its high-energy local behaviour is conjectured to be a **"universal model"** for Laplace eigenfunctions on arbitrary manifolds (Berry, 1977).
- ⋆ It is the local scaling limit of **monochromatic random waves** on arbitrary manifolds (Canzani & Hanin, 2016).

Nodal Sets

One is interested in the length L_E of the **nodal set** (components are the **nodal lines**):

$$B_E^{-1}(\{0\}) \cap \mathcal{Q} := \{x \in \mathcal{Q} : B_E(x) = 0\},$$

where $\mathcal{Q}$ is e.g. a square of size 1, as $E \to \infty$.



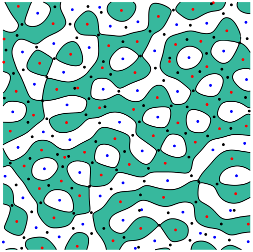
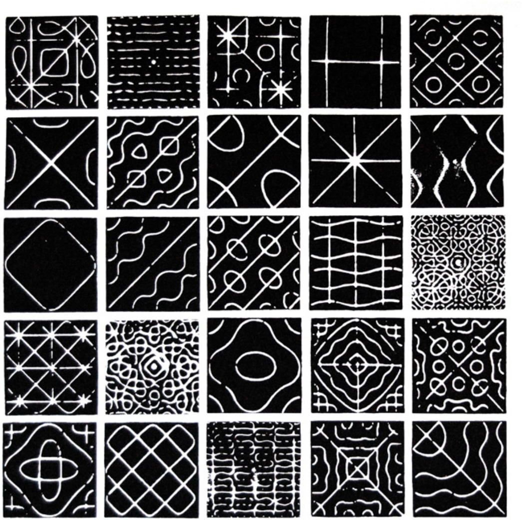


Image: D. Belyaev (2016)

Chladni Plates (1787)



Mean and Variance (Berry, 2002)

- *Berry* (J. Phys. A, 2002) : semi-rigorous computations give

$$\mathbb{E}[L_E] \sim \sqrt{E}, \quad \mathbf{Var}(L_E) \sim \log E,$$

  although the natural guess for the order of the variance is $\sim \sqrt{E}$. See Wigman (2010) for the spherical case.

- Such a variance reduction "... *results from a cancellation whose meaning is still obscure...* " (Berry (2002), p. 3032).

- **Question:** can one explain such a 'cancellation phenomenon', and characterise second-order fluctuations, involving the normalised length

$$\widetilde{L}_E := \frac{L_E - \mathbb{E}[L_E]}{\sqrt{\mathbf{Var}(L_E)}} \ ?$$

Explaining the Cancellation

- ⋆ Starting from seminal contributions by Marinucci and Wigman (2010, 2011): geometric functionals of random Laplace eigenfunctions on compact manifolds (e.g. tori and spheres) can be studied by means of **Wiener-Itô chaotic decompositions** – and in particular by detecting specific **domination effects**.
- ⋆ Such geometric functionals include: **lengths of level sets, excursion areas, Euler-Poincaré characteristics, # critical points, # nodal intersections**. See several works by Cammarota, Dalmao, Marinucci, Nourdin, Peccati, Rossi, Wigman, ... (2010–2018).
- ⋆ As first observed in *Marinucci, Peccati, Rossi and Wigman* (2016 — for arithmetic waves) domination of a single "chaotic projection" fully explains **cancellation phenomena** .

Vignette: Wiener Chaos

- Consider a generic Gaussian field $\mathbb{G} = \{G(u) : u \in \mathscr{U}\}$.
- For every $q = 0, 1, 2..., $ set

$$P_q := \overline{\mathbf{v.s.}}\left\{p\big(G(u_1), ..., G(u_r)\big) : d^\circ p \leq q\right\}.$$

  Then: $P_q \subset P_{q+1}$.
- Define the family of orthogonal spaces $\{C_q : q \geq 0\}$ as $C_0 = \mathbb{R}$ and $C_q := P_q \cap P_{q-1}^\perp$; one has

$$L^2(\sigma(\mathbb{G})) = \bigoplus_{q=0}^{\infty} C_q.$$

- C_q = qth **Wiener chaos** of $\mathbb{G}$.

Chaos and Integration by Parts

- ⋆ Elements of the Wiener chaos verify an **exact integration by parts formula**: for every $F \in C_q$, every $q \geq 2$ and every smooth f,

$$\mathbb{E}[Ff(F)] = \frac{1}{q}\mathbb{E}[\|DF\|^2 f'(F)],$$

  where D is a **generalized gradient** (Malliavin derivative).
- ⋆ This yields the striking inequality (*Nourdin and Peccati*, 2009):

$$\begin{aligned} |\mathbb{E}[f'(F)] - \mathbb{E}[Ff(F)]| &\leq \|f'\|_\infty \mathbf{Var}(q^{-1}\|DF\|^2)^{1/2} \\ &\leq \sqrt{\frac{q-1}{3q}}\sqrt{\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2}. \end{aligned}$$

A Rigid Asymptotic Structure

For fixed $q \geq 2$, let $\{F_k : k \geq 1\} \subset C_q$ (with unit variance).

- $\star$ *Nourdin and Poly (2013)*: If $F_k \Rightarrow Z$, then Z **has necessarily a density** (and the set of possible laws for Z does not depend on $\mathbb{G}$)
- $\star$ *Nualart and Peccati (2005)*: $F_k \Rightarrow Z \sim \mathscr{N}(0,1)$ if and only if $\mathbb{E}F_k^4 \to 3 (= \mathbb{E}Z^4)$, and

$$\mathbf{W}_1(F_k, Z) \leq \sqrt{\mathbb{E}[F_k^4] - 3} \quad (\textit{Nourdin and Peccati, 2009}).$$

- $\star$ *Peccati and Tudor (2005)*: **Componentwise convergence** to Gaussian implies **joint convergence**.
- $\star$ *Nourdin and Peccati (2009)*: $F_k \Rightarrow Z^2 - 1$ if and only if $\mathbb{E}F_k^4 - 12\mathbb{E}F_k^3 \to -36$.
- $\star$ *Nourdin, Nualart and Peccati (2015)*: given $\{H_k\} \subset C_p$, then F_k, H_k are **asymptotically independent** if and only if $\mathbf{Cov}(H_k^2, F_k^2) \to 0$.

Statement

Theorem (Nourdin, Peccati and Rossi, 2017)

1. **(Cancellation)** *For every fixed* $E > 0$*,*

$$\mathrm{proj}(L_E \,|\, C_{2q+1}) = 0, \quad q \geq 0,$$

   and $\mathrm{proj}(\widetilde{L}_E \,|\, C_2)$ *reduces to a "negligible boundary term", as* $E \to \infty$.

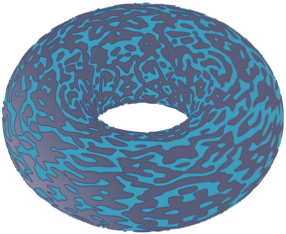
2. **(4^{th} chaos dominates)** *Let* $E \to \infty$*. Then,*

$$\widetilde{L}_E = \mathrm{proj}(\widetilde{L}_E \,|\, C_4) + o_{\mathbb{P}}(1).$$

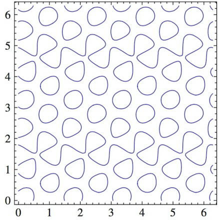
3. **(CLT)** *As* $E \to \infty$*,*

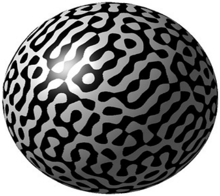
$$\widetilde{L}_E \Rightarrow Z \sim N(0, 1).$$

Other Manifolds?



What about the high-energy behaviour of random waves on $\mathbb{T}^2$ and $\mathbb{S}^2$?





Figures: A. Barnett, G. Poly and Z. Rudnick

Some Recent Findings

- ⋆ Similarly to planar waves, the projection of the (renormalized) nodal length on the second chaos **disappears exactly**, and global fluctuations are dominated (in L^2) by the projection on the **4th Wiener chaos**.

- ⋆ The nodal length of random spherical harmonics verifies a **Gaussian CLT** (*Marinucci, Rossi, Wigman* (2017)).

Some Recent Findings

- The nodal length of arithmetic random waves verifies a **non-central χ^2 limit theorem** (*Marinucci, Rossi, Peccati, Wigman* (2016)), and similar results hold for nodal intersections (*Dalmao, Nourdin, Peccati, Rossi*, 2016), as well as for local nodal lengths above the Planck scale (*Benatar, Marinucci and Wigman*, 2017).

- **Monochromatic random waves** on general manifolds can be coupled with Berry's wave at small scales, so that the local length fluctuations are Gaussian (*Dierinckx, Nourdin, Peccati, Rossi*, 2018+).

Poisson Setting

- ⋆ For every $t > 0$, let

$$\mathbb{R}^d \supset A \mapsto \eta_t(A)$$

  be a **Poisson random measure** with intensity $t \times \mathrm{Leb}$.

- ⋆ Malliavin calculus and Wiener Chaos are available for η_t: as in the Gaussian framework, they combine admirably well with Stein's bounds.

- ⋆ Stochastic analysis on the Poisson space is tightly connected to **add-one cost operators**, defined for every $x \in \mathbb{R}^d$ and every $F = F(\eta)$ as

$$D_x F(\eta_t) := F(\eta_t + \delta_x) - F(\eta_t).$$

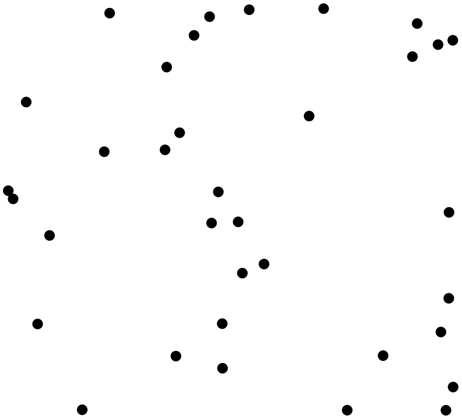
Beyond Chaos

- ⋆ The operators D_x are ersatz of gradients on the Poisson space. In particular, one has the **Poincaré inequality**: for every $F = F(\eta_t)$

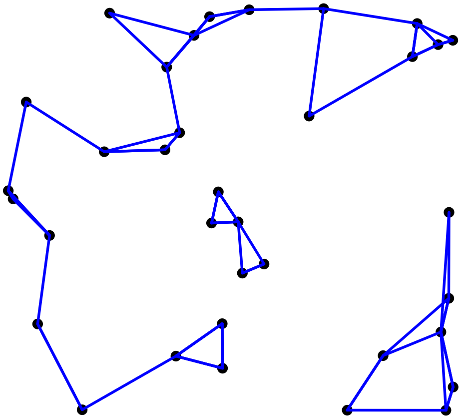
$$\mathbf{Var}(F) \leq t \times \mathbb{E} \int_{\mathbb{R}^d} (D_x F)^2 \, dx.$$

- ⋆ In this framework, several geometric quantities naturally emerge for which there is no dominating "chaotic projection": typically, characteristics of random graphs built from some 'intrinsic geometric rule' – like the **nearest neighbour graph**, or graphs emerging in **combinatorial optimization**.

Example: the Nearest Neighbour Graph



Example: the Nearest Neighbour Graph



Second Order Inequalities and Stabilization

- **Second order Poincaré inequalities** are available also in this framework (*Last, Peccati & Schulte* (2015)):

$$\mathbf{W}_1(F,N)^2 \lesssim \mathbb{E}\left[\int (D_xF)^4\,dx\right] + \mathbb{E}\left[\int (D_xF)^2\,dx\right] \times \mathbb{E}\left[\int\int (D^2_{x,y}F)^2\,dxdy\right],$$

  yielding that normality arises from "small local contributions", and "vanishing second order interactions".

- Such an estimate has recently been used to recover a generalised notion of **"stabilising geometric functionals"** (*Kesten and Lee*, 1996; *Penrose and Yukich*, 2001) – see *Lachièze-Rey, Schulte and Yukich* (2017).

- Applications to: Voronoi tessellations, radial graphs, volume approximations, ...

THANK YOU FOR YOUR ATTENTION!