Lecture 3: Limit theory for statistics of geometric structures via stabilizing score functions

Joe Yukich

February 26, 2019

Probabilistic Analysis of Geometric Structures

· **Lecture 1**: Probabilistic analysis of Euclidean optimization problems

· **Lecture 2**: Central limit theorems for statistics of geometric structures

· **Lecture 3**: Limit theory for statistics of geometric structures via stabilizing score functions

· **Lecture 4**: Statistics of random polytopes

· **Lecture 5**: Rates of multivariate normal approximation for statistics of geometric structures

Lecture 3: Limit theory for statistics of geometric structures via stabilizing scores

- **I Examples and Goals**
- **II Stabilization**
- **III Binomial and Poisson Input**
  - WLLN
  - Gaussian fluctuations
  - Variance asymptotics
- **IV More General Input**

I Examples and Goals

Questions pertaining to geometric structures on random input $\mathcal{X} \subset \mathbb{R}^d$ often involve analyzing sums of spatially correlated terms

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}),$$

where the $\mathbb{R}$-valued score function ξ, defined on pairs $(x, \mathcal{X})$, represents the interaction of x with respect to $\mathcal{X}$.

Questions pertaining to geometric structures on random input $\mathcal{X} \subset \mathbb{R}^d$ often involve analyzing sums of spatially correlated terms

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}),$$

where the $\mathbb{R}$-valued score function ξ, defined on pairs $(x, \mathcal{X})$, represents the interaction of x with respect to $\mathcal{X}$.

The sums describe some global feature of the random structure in terms of local contributions $\xi(x, \mathcal{X}),\ x \in \mathcal{X}$.

I Examples and Goals

Questions pertaining to geometric structures on random input $\mathcal{X} \subset \mathbb{R}^d$ often involve analyzing sums of spatially correlated terms

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}),$$

where the $\mathbb{R}$-valued score function ξ, defined on pairs $(x, \mathcal{X})$, represents the interaction of x with respect to $\mathcal{X}$.

The sums describe some global feature of the random structure in terms of local contributions $\xi(x, \mathcal{X}),\ x \in \mathcal{X}$.

I Examples and Goals: Statistics of random graphs

Clique counts. $\mathcal{X} \subset \mathbb{R}^d$ finite, $\rho \in (0, \infty)$.

· Geometric graph: Join two points of $\mathcal{X}$ iff they are at distance at most ρ.

· Vietoris-Rips complex (with parameter ρ) is simplicial complex whose k-simplices correspond to unordered $(k+1)$-tuples of points in $\mathcal{X}$ all pairwise within ρ of each other.

Clique counts. $\mathcal{X} \subset \mathbb{R}^d$ finite, $\rho \in (0, \infty)$.

· Geometric graph: Join two points of $\mathcal{X}$ iff they are at distance at most ρ.

· Vietoris-Rips complex (with parameter ρ) is simplicial complex whose k-simplices correspond to unordered $(k+1)$-tuples of points in $\mathcal{X}$ all pairwise within ρ of each other.

· For $k \in \mathbb{N}$ and $x \in \mathcal{X}$, put $\sigma_k(x, \mathcal{X}) := \frac{\text{number of } k\text{-simplices containing } x}{k+1}$

Clique counts. $\mathcal{X} \subset \mathbb{R}^d$ finite, $\rho \in (0, \infty)$.

· Geometric graph: Join two points of $\mathcal{X}$ iff they are at distance at most ρ.

· Vietoris-Rips complex (with parameter ρ) is simplicial complex whose k-simplices correspond to unordered $(k+1)$-tuples of points in $\mathcal{X}$ all pairwise within ρ of each other.

· For $k \in \mathbb{N}$ and $x \in \mathcal{X}$, put $\sigma_k(x, \mathcal{X}) := \frac{\text{number of } k\text{-simplices containing } x}{k+1}$

· Total number of k-simplices in Vietoris-Rips complex: $\sum_{x \in \mathcal{X}} \sigma_k(x, \mathcal{X})$.

· Chatterjee, Decreusefond et al., Eichelsbacher, Lachièze-Rey + Peccati, Reitzner + Schulte, Thäle,...

Total edge length of graphs. $\mathcal{X} \subset \mathbb{R}^d$ finite. Given $x \in \mathcal{X}$, let x_{NN} be the nearest neighbor (NN) of x.

· Undirected nearest neighbor graph on $\mathcal{X}$: include an edge $\{x, y\}$ if $y = x_{NN}$ and/or $x = y_{NN}$.

· For $x \in \mathcal{X}$, put

$$\xi(x, \mathcal{X}) := \begin{cases} \frac{1}{2}||x - x_{NN}|| & \text{if} \quad x, x_{NN} \quad \text{are mutual NN} \\ ||x - x_{NN}|| & \text{otherwise.} \end{cases}$$

Total edge length of graphs. $\mathcal{X} \subset \mathbb{R}^d$ finite. Given $x \in \mathcal{X}$, let x_{NN} be the nearest neighbor (NN) of x.

· Undirected nearest neighbor graph on $\mathcal{X}$: include an edge $\{x, y\}$ if $y = x_{NN}$ and/or $x = y_{NN}$.

· For $x \in \mathcal{X}$, put

$$\xi(x, \mathcal{X}) := \begin{cases} \frac{1}{2}||x - x_{NN}|| & \text{if } \ x, x_{NN} \ \text{ are mutual NN} \\ ||x - x_{NN}|| & \text{otherwise.} \end{cases}$$

· Total edge length of NN graph on $\mathcal{X}$: $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$.

· Chatterjee; Last, Peccati, + Schulte; Steele; Penrose + Y

· $\mathcal{X} \subset \mathbb{R}^d$ a collection of 'germs'.

· $S_x, x \in \mathcal{X}$, a collection of 'grains' (closed bounded sets).

· Germ-grain model: $\bigcup_{x \in \mathcal{X}} (x \oplus S_x)$.

· Total surface area, volume, clump count,... may be expressed as $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ for appropriate ξ. For example, for $x \in \mathcal{X}$ we put

$$\xi_{\text{clump}}(x, \mathcal{X}) := (\text{size of clump of germ-grain model containing } x)^{-1}.$$

· $\mathcal{X} \subset \mathbb{R}^d$ a collection of 'germs'.

· $S_x, x \in \mathcal{X}$, a collection of 'grains' (closed bounded sets).

· Germ-grain model: $\bigcup_{x \in \mathcal{X}} (x \oplus S_x)$.

· Total surface area, volume, clump count,... may be expressed as $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ for appropriate ξ. For example, for $x \in \mathcal{X}$ we put

$$\xi_{\text{clump}}(x, \mathcal{X}) := (\text{size of clump of germ-grain model containing } x)^{-1}.$$

· Clump count in germ-grain model equals $\sum_{x \in \mathcal{X}} \xi_{\text{clump}}(x, \mathcal{X})$.

· Baddeley; Hall; Hug, Last + Schulte; Molchanov; Penrose + Y; Schneider + Weil; Stoyan; Thäle

I Examples and Goals: Random packing (Random sequential adsorption)

· $\mathcal{X} \subset \mathbb{R}^d$ finite. Assign $x \in \mathcal{X}$ a time mark $\tau_x \in [0, 1]$.

· Let $B_1, B_2, ...$ be a sequence of unit volume d-dimensional Euclidean balls with centers arriving sequentially at points $x_i \in \mathcal{X}$ and at arrival times τ_{x_i}.

· The first ball B_1 to arrive is packed. Recursively, for $i = 2, 3, ...$, the ith ball is packed if it does not overlap any ball in $B_1, B_2, ..., B_{i-1}$ which has already been packed.

I Examples and Goals: Random packing (Random sequential adsorption)

· $\mathcal{X} \subset \mathbb{R}^d$ finite. Assign $x \in \mathcal{X}$ a time mark $\tau_x \in [0,1]$.

· Let $B_1, B_2, ...$ be a sequence of unit volume d-dimensional Euclidean balls with centers arriving sequentially at points $x_i \in \mathcal{X}$ and at arrival times τ_{x_i}.

· The first ball B_1 to arrive is packed. Recursively, for $i = 2, 3, ...$, the ith ball is packed if it does not overlap any ball in $B_1, B_2, ..., B_{i-1}$ which has already been packed.

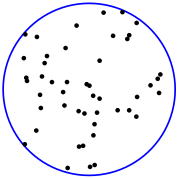
· For $x \in \mathcal{X}$ define packing functional

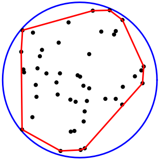
$$\xi_{\text{pack}}(x, \mathcal{X}) := \begin{cases} 1 & \text{if ball arriving at } x \text{ is packed} \\ 0 & \text{otherwise.} \end{cases}$$

Then total number of packed balls equals $\sum_{x \in \mathcal{X}} \xi_{\text{pack}}(x, \mathcal{X})$.

· Rényi, Coffman, Dvoretzky + Robbins; Flory, Itoh + Shepp; Torquato,...

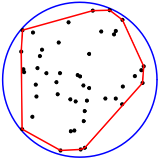
· $\mathcal{X} \subset \mathbb{R}^d$ finite. Let $\mathrm{co}(\mathcal{X})$ denote the convex hull of $\mathcal{X}$.





· For $x \in \mathcal{X}$, $k \in \{0, 1, ..., d-1\}$, we put

$$f_k(x, \mathcal{X}) := \frac{1}{k+1}(\text{number of k} - \text{dimensional faces containing } x).$$



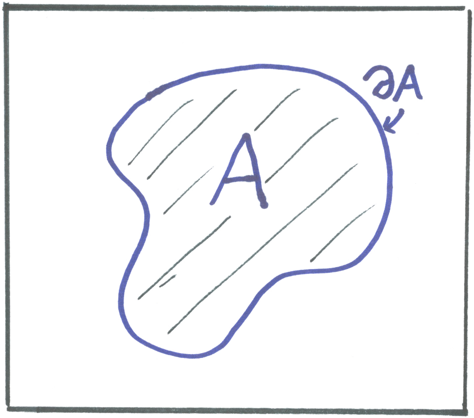
· For $x \in \mathcal{X}$, $k \in \{0, 1, ..., d-1\}$, we put

$$f_k(x, \mathcal{X}) := \frac{1}{k+1}(\text{number of k} - \text{dimensional faces containing } x).$$

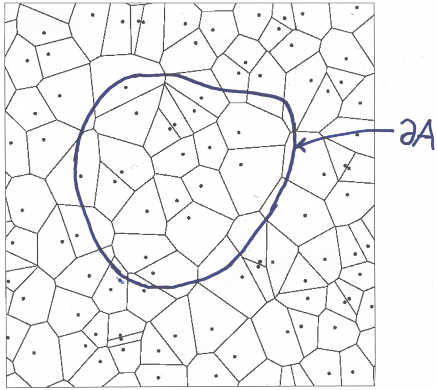
· Total number of k-dimensional faces of $\text{co}(\mathcal{X})$: $\sum_{x \in \mathcal{X}} f_k(x, \mathcal{X})$.

· Rényi+Sulanke; Bárány;Buchta;Calka,Groeneboom,Reitzner,Thäle,Vu,...

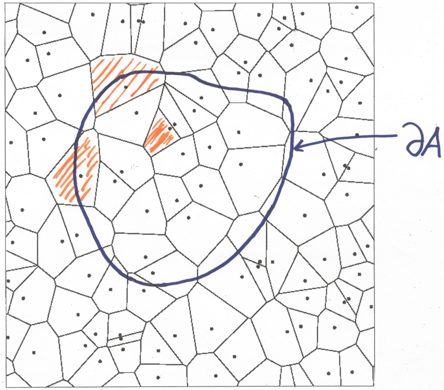
I Examples and Goals: Poisson - Voronoi tessellation

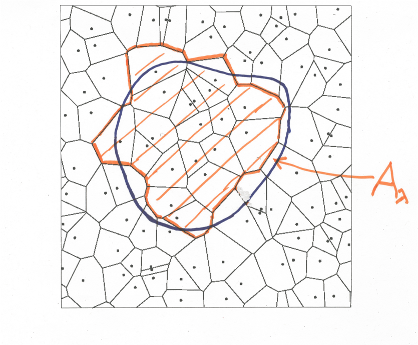


I Examples and Goals: Poisson - Voronoi tessellation



I Examples and Goals: Poisson - Voronoi tessellation





Def. A_λ is the Poisson - Voronoi approximation of A.

Question: What is $\mathrm{Vol}(A_\lambda)$?

Define volume score: $\xi_{Vol}(x, \mathcal{X}) = \mathrm{Vol}(C(x, \mathcal{X}))$ when $x \in A$, otherwise put the score to be zero. Sum of volume scores gives $\mathrm{Vol}(A_\lambda)$.

General questions.

· When $\mathcal{X} \subset \mathbb{R}^d$ is a random pt configuration, the sums $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ describe a global feature of some spatial random structure.

General questions.

· When $\mathcal{X} \subset \mathbb{R}^d$ is a random pt configuration, the sums $\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X})$ describe a global feature of some spatial random structure.

· What is the distribution of these sums for large pt configurations $\mathcal{X}$? LLN? CLT?

I Examples and Goals

$\mathcal{P}$: a stationary pt process on $\mathbb{R}^d$

Restrict to windows: $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$.

$\mathcal{P}$: a stationary pt process on $\mathbb{R}^d$

Restrict to windows: $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$.

Goal. Given a score function $\xi(\cdot, \cdot)$ defined on pairs $(x, \mathcal{X})$, given a pt process $\mathcal{P}$, we seek the limit theory (LLN, CLT, variance asymptotics) for the total score

$$\sum_{x \in \mathcal{P} \cap Q_\lambda} \xi(x, \mathcal{P} \cap Q_\lambda)$$

and total measure

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P} \cap Q_\lambda} \xi(x, \mathcal{P} \cap Q_\lambda) \delta_{\lambda^{-1/d} x}.$$

$\mathcal{P}$: a stationary pt process on $\mathbb{R}^d$

Restrict to windows: $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$.

Goal. Given a score function $\xi(\cdot, \cdot)$ defined on pairs $(x, \mathcal{X})$, given a pt process $\mathcal{P}$, we seek the limit theory (LLN, CLT, variance asymptotics) for the total score

$$\sum_{x \in \mathcal{P} \cap Q_\lambda} \xi(x, \mathcal{P} \cap Q_\lambda)$$

and total measure

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P} \cap Q_\lambda} \xi(x, \mathcal{P} \cap Q_\lambda) \delta_{\lambda^{-1/d} x}.$$

Tractable problems must be *local* in the sense that points far away from x should not play a role in the evaluation of the score $\xi(x, \mathcal{P} \cap Q_\lambda)$.

II Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.

Recall $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$

II Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.
Recall $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$

Key Definition. ξ is *stabilizing* wrt pt process $\mathcal{P}$ on $\mathbb{R}^d$ if for all $x \in \mathcal{P}$ there is $R := R^\xi(x, \mathcal{P}) < \infty$ a.s. (a 'radius of stabilization') such that

$$\xi(x, \mathcal{P} \cap B_R(x)) = \xi(x, (\mathcal{P} \cap B_R(x)) \cup (\mathcal{A} \cap B_R^c(x))).$$

for any locally finite $\mathcal{A} \subset \mathbb{R}^d$.

II Stabilization

We assume translation invariant scores: $\xi(x, \mathcal{X}) = \xi(\mathbf{0}, \mathcal{X} - x)$.

Recall $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$

Key Definition. ξ is *stabilizing* wrt pt process $\mathcal{P}$ on $\mathbb{R}^d$ if for all $x \in \mathcal{P}$ there is $R := R^\xi(x, \mathcal{P}) < \infty$ a.s. (a 'radius of stabilization') such that

$$\xi(x, \mathcal{P} \cap B_R(x)) = \xi(x, (\mathcal{P} \cap B_R(x)) \cup (\mathcal{A} \cap B_R^c(x))).$$

for any locally finite $\mathcal{A} \subset \mathbb{R}^d$. ξ is *exponentially stabilizing* wrt $\mathcal{P}$ if there is a constant $c \in (0, \infty)$ such that

$$\sup_{\lambda \geq 1} \sup_{x \in Q_\lambda} \mathbb{P}(R^\xi(x, \mathcal{P} \cap Q_\lambda) \geq r) \leq c \exp(-\frac{r}{c}), \quad r \in [1, \infty).$$

II Stabilization

· $\mathcal{P}$: a pt process on $\mathbb{R}^d$
· $\mathcal{P} \cap Q_\lambda := \mathcal{P} \cap [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$.

Definition (Moment condition). ξ satisfies the p moment condition wrt $\mathcal{P}$ if

$$\sup_{\lambda \geq 1} \sup_{x,y \in \mathbb{R}^d} \mathbb{E} \, |\xi(x, (\mathcal{P} \cap Q_\lambda) \cup \{y\})|^p < \infty.$$

II Stabilization

· Let $\mathcal{P}_1$ be a rate 1 Poisson pt process on $\mathbb{R}^d$
· $\mathcal{P}_1 \cap Q_\lambda := \mathcal{P}_1 \cap [\frac{-\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$. Put

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}.$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (1, \infty)$, then for all $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ we have

$$|\lambda^{-1} \mathbb{E} \langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P}_1 \cup \{\mathbf{0}\}) \int_{[-\frac{1}{2}, \frac{1}{2}]^d} f(x) dx| \leq \epsilon_\lambda.$$

II Stabilization

· Let $\mathcal{P}_1$ be a rate 1 Poisson pt process on $\mathbb{R}^d$
· $\mathcal{P}_1 \cap Q_\lambda := \mathcal{P}_1 \cap [\frac{-\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$. Put

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}.$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (1, \infty)$, then for all $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ we have

$$|\lambda^{-1} \mathbb{E} \langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P}_1 \cup \{\mathbf{0}\}) \int_{[-\frac{1}{2}, \frac{1}{2}]^d} f(x) dx| \leq \epsilon_\lambda.$$

Penrose and Y (2003): $\epsilon_\lambda = o(1)$.

II Stabilization

· Let $\mathcal{P}_1$ be a rate 1 Poisson pt process on $\mathbb{R}^d$
· $\mathcal{P}_1 \cap Q_\lambda := \mathcal{P}_1 \cap [\frac{-\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$. Put

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}.$$

Thm (WLLN): If ξ is stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (1, \infty)$, then for all $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ we have

$$|\lambda^{-1} \mathbb{E} \langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P}_1 \cup \{\mathbf{0}\}) \int_{[-\frac{1}{2}, \frac{1}{2}]^d} f(x) dx| \leq \epsilon_\lambda.$$

Penrose and Y (2003): $\epsilon_\lambda = o(1)$.

Lachièze-Rey, Schulte, + Y (2017): $\epsilon_\lambda = O(\lambda^{-1/d})$ if ξ is exponentially stabilizing wrt $\mathcal{P}_1$.

What about weak laws of large numbers on non-uniform input?

It is useful to consider ξ_λ defined as follows. For all $\lambda > 0$ define the λ *re-scaled version* of ξ by

$$\xi_\lambda(x, \mathcal{X}) := \xi(\lambda^{1/d} x, \lambda^{1/d} \mathcal{X}).$$

Re-scaling is natural when considering point sets $\mathcal{X}$ in compact sets K having cardinality roughly λ; dilation by $\lambda^{1/d}$ means that unit volume subsets of $\lambda^{1/d} K$ host on the average one point.

II Stabilization

$\mathcal{P}_{\lambda\kappa}$: PPP on $\mathbb{R}^d$ with intensity density $\lambda\kappa(x)dx$.

One may show that $\lambda^{1/d}(\mathcal{P}_{\lambda\kappa} - x_0) \xrightarrow{\mathcal{D}} \mathcal{P}_{\kappa(x_0)}$ as $\lambda \to \infty$, where convergence is in the sense of weak convergence of point processes.

II Stabilization

$\mathcal{P}_{\lambda\kappa}$: PPP on $\mathbb{R}^d$ with intensity density $\lambda\kappa(x)dx$.

One may show that $\lambda^{1/d}(\mathcal{P}_{\lambda\kappa} - x_0) \xrightarrow{\mathcal{D}} \mathcal{P}_{\kappa(x_0)}$ as $\lambda \to \infty$, where convergence is in the sense of weak convergence of point processes.

If $\xi(\cdot,\cdot)$ is a functional defined on $\mathbb{R}^d \times \mathbf{N}$, where we recall $\mathbf{N}$ is the space of locally finite point sets in $\mathbb{R}^d$, one might hope that ξ is *continuous* on the pairs $(\mathbf{0}, \lambda^{1/d}(\mathcal{P}_{\lambda\kappa} - x_0))$ in the sense that $\xi(\mathbf{0}, \lambda^{1/d}(\mathcal{P}_{\lambda\kappa} - x_0))$ converges in distribution to $\xi(\mathbf{0}, \mathcal{P}_{\kappa(x_0)})$ as $\lambda \to \infty$.

This turns out to be the case whenever ξ is stabilizing wrt to $\mathcal{P}_{\kappa(x_0)}$ and if x_0 is a Lebesgue point of κ.

II Stabilization

Almost every $x \in \mathbb{R}^d$ is a *Lebesgue point* of κ, that is to say for almost all $x \in \mathbb{R}^d$ we have

$$\lim_{\epsilon \to 0} \epsilon^{-d} \int_{B_\epsilon(x)} |\kappa(y) - \kappa(x)| \, dy = 0.$$

Lemma (convergence of re-scaled binomial pt process) Let $x \in \mathbb{R}^d$ be a Lebesgue point for κ. Then

$$n^{1/d}(-x + \mathcal{X}_n) \xrightarrow{\mathcal{D}} \mathcal{P}_{\kappa(x)}, \quad n \to \infty.$$

Lemma (convergence of re-scaled binomial pt process) Let $x \in \mathbb{R}^d$ be a Lebesgue point for κ. Then

$$n^{1/d}(-x + \mathcal{X}_n) \xrightarrow{\mathcal{D}} \mathcal{P}_{\kappa(x)}, \quad n \to \infty.$$

Key 'Continuity' Lemma. Let $x \in \mathbb{R}^d$ be a Lebesgue point for κ and assume that $R^\xi(x, \mathcal{P}_{\kappa(x)}) < \infty$ a.s. where $R^\xi(x, \mathcal{P}_{\kappa(x)})$ is the radius of stabilization for ξ at x wrt $\mathcal{P}_{\kappa(x)}$. Then

(a) $\xi_\lambda(x, \mathcal{P}_{\lambda\kappa}) \xrightarrow{\mathcal{D}} \xi(\mathbf{0}, \mathcal{P}_{\kappa(x)}), \quad \lambda \to \infty,$

(b) $\xi_n(x, \mathcal{X}_n) \xrightarrow{\mathcal{D}} \xi(\mathbf{0}, \mathcal{P}_{\kappa(x)}), \quad n \to \infty.$

So stabilization of a score function acts as a surrogate for continuity.

Theorem (Campbell Formula). Let $\mathcal{P}$ be a point process on $\mathbb{R}^d$ with intensity $\kappa(x)$ and let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then the random sum

$$\sum_{x \in \mathcal{P}} f(x)$$

has expected value

$$\mathbb{E} \sum_{x \in \mathcal{P}} f(x) = \int_{\mathbb{R}^d} f(x)\kappa(x)dx.$$

Theorem (Campbell Formula). Let $\mathcal{P}$ be a point process on $\mathbb{R}^d$ with intensity $\kappa(x)$ and let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then the random sum

$$\sum_{x \in \mathcal{P}} f(x)$$

has expected value

$$\mathbb{E} \sum_{x \in \mathcal{P}} f(x) = \int_{\mathbb{R}^d} f(x)\kappa(x)dx.$$

Theorem (Mecke Formula). The random sum

$$\sum_{x \in \mathcal{P}} f(x, \mathcal{P})$$

has expected value

$$\mathbb{E} \sum_{x \in \mathcal{P}} f(x, \mathcal{P}) = \int_{\mathbb{R}^d} \mathbb{E}\, f(x, \mathcal{P} \cup \{x\})\kappa(x)dx.$$

II Stabilization

Let $\mathbb{B}(K)$ denote the class of all bounded $f : K \to \mathbb{R}$ and for all measures μ on $\mathbb{R}^d$ let $\langle f, \mu \rangle := \int f d\mu$. Put $\bar{\mu} := \mu - \mathbb{E}\,\mu$.

For all $f \in \mathbb{B}(\mathbb{R}^d)$ we have by Mecke formula that

$$\mathbb{E}\,[\langle f, \mu_\lambda \rangle] = \lambda \int_{\mathbb{R}^d} f(x) \mathbb{E}\,[\xi_\lambda(x, \mathcal{P}_{\lambda\kappa})] \kappa(x)\, dx.$$

If the moment condition

$$\sup_{\lambda} \sup_{x,y \in \mathbb{R}^d} \mathbb{E}\,|\xi_\lambda(x, (\mathcal{P}_{\lambda\kappa}) \cup \{y\})|^p < \infty$$

holds for some $p > 1$, then uniform integrability and the 'continuity' Lemma show that for all Lebesgue points x of κ one has $\mathbb{E}\,\xi_\lambda(x, \mathcal{P}_{\lambda\kappa}) \to \mathbb{E}\,\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})$ as $\lambda \to \infty$.

· If the moment condition

$$\sup_{\lambda} \sup_{x,y \in \mathbb{R}^d} \mathbb{E}\, |\xi_\lambda(x, (\mathcal{P}_{\lambda\kappa}) \cup \{y\})|^p < \infty$$

holds for some $p > 1$, then uniform integrability and our 'continuity' Lemma show that for all Lebesgue points x of κ one has $\mathbb{E}\, \xi_\lambda(x, \mathcal{P}_{\lambda\kappa}) \to \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})$ as $\lambda \to \infty$.

· The set of points failing to be Lebesgue points has measure zero and so when the moment condition holds for some $p > 1$, the bounded convergence theorem gives

$$\lim_{\lambda \to \infty} \lambda^{-1} \mathbb{E}\, [\langle f, \mu_\lambda \rangle] = \int_{\mathbb{R}^d} f(x) \mathbb{E}\, [\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})] \kappa(x)\, dx.$$

II Stabilization

Convergence of means $\mathbb{E}[\langle f, \mu_\lambda \rangle]$ is now upgraded to convergence in L^q, $q = 1$ or 2.

We also prove LLN for

$$\mu_n^\xi := \sum_{i=1}^n \xi_n(X_i, \mathcal{X}_n)\delta_{X_i}$$

where $\mathcal{X}_n := \{X_i\}_{i=1}^n$, $X_i, i \geq 1$, i.i.d. with density κ.

III Binomial and Poisson Input: WLLN

Theorem (WLLN for sums of translation invariant stabilizing functionals) Let $q = 1$ or $q = 2$. Assume that ξ is translation invariant and stabilizing, so that $\xi(\mathbf{0}, \mathcal{P}_\tau) := \lim_{r \to \infty} \xi(\mathbf{0}, \mathcal{P}_\tau \cap B_r(\mathbf{0}))$ exists for all $\tau > 0$. If $\sup_n \mathbb{E}\, |\xi_n(X_1, \mathcal{X}_n)|^p < \infty$ for some $p \in (q, \infty)$, then for all $f \in \mathbb{B}(\mathbb{R}^d)$ we have

$$\lim_{n \to \infty} n^{-1} \langle f, \mu_n \rangle = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_n(X_i, \mathcal{X}_n) f(X_i)$$

$$= \int f(x) \mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})] \kappa(x) dx \text{ in } L^q.$$

Theorem (WLLN for sums of translation invariant stabilizing functionals) Let $q = 1$ or $q = 2$. Assume that ξ is translation invariant and stabilizing, so that $\xi(\mathbf{0}, \mathcal{P}_\tau) := \lim_{r \to \infty} \xi(\mathbf{0}, \mathcal{P}_\tau \cap B_r(\mathbf{0}))$ exists for all $\tau > 0$. If $\sup_n \mathbb{E}\, |\xi_n(X_1, \mathcal{X}_n)|^p < \infty$ for some $p \in (q, \infty)$, then for all $f \in \mathbb{B}(\mathbb{R}^d)$ we have

$$\lim_{n \to \infty} n^{-1} \langle f, \mu_n \rangle = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_n(X_i, \mathcal{X}_n) f(X_i)$$

$$= \int f(x) \mathbb{E}\, [\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})] \kappa(x) dx \text{ in } L^q.$$

If $\sup_\lambda \mathbb{E}\, |\xi_\lambda(\mathbf{0}, \mathcal{P}_{\lambda\kappa})|^p < \infty$ for some $p \in (q, \infty)$, then for all $f \in \mathbb{B}(\mathbb{R}^d)$ we have

$$\lim_{\lambda \to \infty} \lambda^{-1} \langle f, \mu_\lambda \rangle = \int f(x) \mathbb{E}\, [\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)})] \kappa(x) dx \text{ in } L^q.$$

III Binomial and Poisson Input: WLLN

Corollaries of WLLN. We can deduce a weak law of large numbers for the following statistics:

· clique counts in the random geometric graph on $\mathcal{P}_{\lambda\kappa}$
· total edge length of nearest neighbors graph on $\mathcal{P}_{\lambda\kappa}$
· clump count in the germ grain model on $\mathcal{P}_{\lambda\kappa}$
· number of balls accepted in RSA model on $\mathcal{P}_{\lambda\kappa}$

III Binomial and Poisson Input: Gaussian fluctuations

Recall $\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}$.

Thm (CLT): Assume ξ is exponentially stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (5, \infty)$. If $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ satisfies $\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle = \Omega(\lambda)$, then

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left[\frac{\langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\,\langle \mu_\lambda^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle}} \leq t \right] - \mathbb{P}[N(0,1) \leq t] \right| \leq \epsilon_\lambda.$$

III Binomial and Poisson Input: Gaussian fluctuations

Recall $\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}$.

Thm (CLT): Assume ξ is exponentially stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (5, \infty)$. If $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ satisfies $\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle = \Omega(\lambda)$, then

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left[\frac{\langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\,\langle \mu_\lambda^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle}} \leq t \right] - \mathbb{P}[N(0,1) \leq t] \right| \leq \epsilon_\lambda.$$

Penrose + Y (2005), Penrose (2007): $\epsilon_\lambda = O((\log \lambda)^{3d} \lambda^{-1/2})$.

III Binomial and Poisson Input: Gaussian fluctuations

Recall $\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_1 \cap Q_\lambda} \xi(x, \mathcal{P}_1 \cap Q_\lambda) \delta_{\lambda^{-1/d} x}$.

Thm (CLT): Assume ξ is exponentially stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (5, \infty)$. If $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ satisfies $\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle = \Omega(\lambda)$, then

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left[\frac{\langle \mu_\lambda^\xi, f \rangle - \mathbb{E}\langle \mu_\lambda^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_\lambda^\xi, f \rangle}} \leq t \right] - \mathbb{P}[N(0,1) \leq t] \right| \leq \epsilon_\lambda.$$

Penrose + Y (2005), Penrose (2007): $\epsilon_\lambda = O((\log \lambda)^{3d} \lambda^{-1/2})$.

Lachièze-Rey, Schulte, + Y (2017): $\epsilon_\lambda = O(\lambda^{-1/2})$ (Stein's method)

III Binomial and Poisson Input: Gaussian fluctuations

The gaussian fluctuation result may be extended to treat:

· binomial input
· stabilizing functionals on input on general metric spaces
· stabilizing functionals on input on manifolds

III Poisson Input: Variance asymptotics

Lemma (continuity lemma for pairs) Let x be a Lebesgue point for κ. If ξ is stabilizing w.r.t. $\mathcal{P}_{\kappa(x)}$, then for all $z \in \mathbb{R}^d$, we have as $\lambda \to \infty$

$$(\xi_\lambda(x, \mathcal{P}_{\lambda\kappa}), \xi_\lambda(x + \lambda^{-1/d} z, \mathcal{P}_{\lambda\kappa})) \xrightarrow{\mathcal{D}} (\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)}), \xi(z, \mathcal{P}_{\kappa(x)})).$$

We use this lemma to prove variance asymptotics. (Remember it for the next slide.)

By Mecke's Formula for the Poisson process $\mathcal{P}_{\lambda\kappa}$ we have

$$\begin{aligned}
&\lambda^{-1}\mathrm{Var}[\langle f, \mu_\lambda \rangle] \\
&= \lambda \int_K \int_K f(x) f(y) \{\mathbb{E}\,[\xi_\lambda(x, \mathcal{P}_{\lambda\kappa} \cup \{y\}) \xi_\lambda(y, \mathcal{P}_{\lambda\kappa} \cup \{x\})] \\
&\quad - \mathbb{E}\,[\xi_\lambda(x, \mathcal{P}_{\lambda\kappa})] \mathbb{E}\,[\xi_\lambda(y, \mathcal{P}_{\lambda\kappa})]\} \kappa(x) \kappa(y)\, dx\, dy \\
&\quad + \int_K f(x)^2 \mathbb{E}\,[\xi_\lambda^2(x, \mathcal{P}_{\lambda\kappa})] \kappa(x)\, dx.
\end{aligned}$$

Put $y = x + \lambda^{-1/d} z$ in the right-hand side of the above (so $\lambda dy = dz$). Then the two point correlation function $\{...\}$ becomes

$$\{...\} := \{\mathbb{E}\,[\xi_\lambda(x, \mathcal{P}_{\lambda\kappa} \cup \{x + \lambda^{-1/d} z\}) \xi_\lambda(x + \lambda^{-1/d} z, \mathcal{P}_{\lambda\kappa} \cup \{x\})]$$

$$-\mathbb{E}\,[\xi_\lambda(x, \mathcal{P}_{\lambda\kappa})] \mathbb{E}\,[\xi_\lambda(x + \lambda^{-1/d} z, \mathcal{P}_{\lambda\kappa})]\}.$$

Now use $\xi_\lambda(x, \mathcal{P}_{\lambda\kappa}) \xi_\lambda(x + \lambda^{-1/d} z, \mathcal{P}_{\lambda\kappa}) \xrightarrow{\mathcal{D}} \xi(\mathbf{0}, \mathcal{P}_{\kappa(x)}) \xi(z, \mathcal{P}_{\kappa(x)})$.

III Poisson Input: Variance asymptotics

· Assuming exponential stabilization, the integrand in the above is dominated by an integrable function of z over $\mathbb{R}^d$.
· For simplicity we assume that f is a.e. continuous.

· The double integral in the above thus converges to

$$\int_K \int_{\mathbb{R}^d} [\mathbb{E}\,[\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)} \cup \{z\})\xi(z, \mathcal{P}_{\kappa(x)} \cup \mathbf{0})] - (\mathbb{E}\,\xi(\mathbf{0}, \mathcal{P}_{\kappa(x)}))^2] f(x)^2 \cdot \kappa(x)^2 \, dz \, dx$$

by dominated convergence, the a.e. continuity of f, and the assumed moment bounds.

III Poisson Input: Variance asymptotics

Given homogenous rate 1 Poisson input $\mathcal{P}_1$ on $\mathbb{R}^d$, and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0}, \mathcal{P}_1) + \int_{\mathbb{R}^d} [\mathbb{E}\,\xi(\mathbf{0}, \mathcal{P}_1 \cup \{x\})\xi(x, \mathcal{P}_1 \cup \{\mathbf{0}\}) - \mathbb{E}\,\xi(\mathbf{0}, \mathcal{P}_1)\mathbb{E}\,\xi(x, \mathcal{P}_1)]dx.$$

III Poisson Input: Variance asymptotics

Given homogenous rate 1 Poisson input $\mathcal{P}_1$ on $\mathbb{R}^d$, and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0},\mathcal{P}_1) + \int_{\mathbb{R}^d} [\mathbb{E}\,\xi(\mathbf{0},\mathcal{P}_1 \cup \{x\})\xi(x,\mathcal{P}_1 \cup \{\mathbf{0}\}) - \mathbb{E}\,\xi(\mathbf{0},\mathcal{P}_1)\mathbb{E}\,\xi(x,\mathcal{P}_1)]dx.$$

Thm (variance asymptotics): If ξ is exponentially stabilizing wrt $\mathcal{P}_1$ and satisfies the p moment condition for some $p \in (2,\infty)$, then for all $f \in B([-\frac{1}{2},\frac{1}{2}]^d)$ we have

$$\lim_{\lambda\to\infty} \lambda^{-1}\mathrm{Var}\langle \mu_\lambda^\xi, f\rangle = \sigma^2(\xi)\int_{[-\frac{1}{2},\frac{1}{2}]^d} f^2(x)dx \in [0,\infty).$$

Baryshnikov + Y (2005); Penrose (2007)

IV General input

· **Question.** If the input pt process is neither Poisson nor binomial, when do we get results which are qualitatively similar?

· Soshnikov (2002): establishes asymptotic normality of the *count statistic*

$$\sum_{x \in \mathcal{P} \cap Q_n} \delta_{n^{-1/d} x}$$

where $\mathcal{P}$ is determinantal pt process, $\mathcal{P} \cap Q_n := \mathcal{P}_1 \cap [-\frac{1}{2} n^{1/d}, \frac{1}{2} n^{1/d}]^d$.

IV General input

· **Question.** If the input pt process is neither Poisson nor binomial, when do we get results which are qualitatively similar?

· Soshnikov (2002): establishes asymptotic normality of the *count statistic*

$$\sum_{x \in \mathcal{P} \cap Q_n} \delta_{n^{-1/d}x}$$

where $\mathcal{P}$ is determinantal pt process, $\mathcal{P} \cap Q_n := \mathcal{P}_1 \cap [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$.

· Nazarov and Sodin (2012): establish asymptotic normality of the *count statistic*

$$\sum_{x \in \mathcal{P} \cap Q_n} \delta_{n^{-1/d}x}$$

where $\mathcal{P}$ is zero set of Gaussian analytic function.

· We want to extend these results to more general statistics

$$\mu_n^{\xi} := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d}x}.$$

Def (correlation functions). Given a simple pt process $\mathcal{P}$ on $\mathbb{R}^d$, the k pt correlation function $\rho^{(k)} : (\mathbb{R}^d)^k \to [0, \infty)$ is defined via

$$\mathbb{E}\left[\Pi_{i=1}^k \mathrm{card}(\mathcal{P} \cap B_i)\right] = \int_{B_1} ... \int_{B_k} \rho^{(k)}(x_1, ..., x_k) dx_1 ... dx_k,$$

where $B_1, ..., B_k$ are disjoint subsets of $\mathbb{R}^d$.

Rks.

$\rho^{(k)}(x_1, ..., x_k) = \Pi_{i=1}^k \rho^{(1)}(x_i)$ characterizes the Poisson pt process

Def (correlation functions). Given a simple pt process $\mathcal{P}$ on $\mathbb{R}^d$, the k pt correlation function $\rho^{(k)} : (\mathbb{R}^d)^k \to [0, \infty)$ is defined via

$$\mathbb{E}\left[\Pi_{i=1}^k \mathrm{card}(\mathcal{P} \cap B_i)\right] = \int_{B_1} ... \int_{B_k} \rho^{(k)}(x_1, ..., x_k) dx_1 ... dx_k,$$

where $B_1, ..., B_k$ are disjoint subsets of $\mathbb{R}^d$.

Rks.

$\rho^{(k)}(x_1, ..., x_k) = \Pi_{i=1}^k \rho^{(1)}(x_i)$ characterizes the Poisson pt process

$\rho^{(k)}(x_1, ..., x_k) \geq \Pi_{i=1}^k \rho^{(1)}(x_i)$ implies $\mathcal{P}$ is attractive

$\rho^{(k)}(x_1, ..., x_k) \leq \Pi_{i=1}^k \rho^{(1)}(x_i)$ implies $\mathcal{P}$ is repulsive

IV General input

Key Definition (weak decay of correlations). A pt process $\mathcal{P}$ has *weak decay of correlations (w.d.c.)* if there is a fast decreasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $p, q \in \mathbb{N}$ there are constants $c_{p,q}$ and $C_{p,q}$ such that for all $x_1, ..., x_{p+q} \in \mathbb{R}^d$,

$$|\rho^{(p+q)}(x_1, ..., x_{p+q}) - \rho^{(p)}(x_1, ..., x_p)\rho^{(q)}(x_{p+1}, ..., x_{p+q})| \leq C_{p,q}\phi(-c_{p,q}s),$$

where $s := \inf_{i \in \{1,...,p\},\ j \in \{p+1,...,p+q\}} ||x_i - x_j||$.

(ϕ 'fast decreasing' means ϕ decaying faster than any power)

Key Definition (weak decay of correlations). A pt process $\mathcal{P}$ has *weak decay of correlations (w.d.c.)* if there is a fast decreasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $p, q \in \mathbb{N}$ there are constants $c_{p,q}$ and $C_{p,q}$ such that for all $x_1, ..., x_{p+q} \in \mathbb{R}^d$,

$$|\rho^{(p+q)}(x_1, ..., x_{p+q}) - \rho^{(p)}(x_1, ..., x_p)\rho^{(q)}(x_{p+1}, ..., x_{p+q})| \leq C_{p,q}\phi(-c_{p,q}s),$$

where $s := \inf_{i \in \{1,...,p\},\ j \in \{p+1,...,p+q\}} ||x_i - x_j||$.

(ϕ 'fast decreasing' means ϕ decaying faster than any power)

Note: 'weak decay of correlations' is called 'clustering' in physics literature.

IV General input

Ex. 1: Determinantal pt process. A pt process is determinantal (DPP) if its correlation functions satisfy

$$\rho^{(k)}(x_1, ..., x_k) = \det(K(x_i, x_j))_{1 \le i \le j \le k},$$

where $K(\cdot, \cdot)$ is Hermitian non-negative definite kernel of locally trace class integral operator from $L^2(\mathbb{R}^d)$ to itself.

DPP is repulsive

Fact If $|K(x, y)| \le \phi(||x - y||)$, with ϕ fast decreasing, then the DPP has weak decay of correlations.

Ex. 1: Determinantal pt process. A pt process is determinantal (DPP) if its correlation functions satisfy

$$\rho^{(k)}(x_1, ..., x_k) = \det(K(x_i, x_j))_{1 \le i \le j \le k},$$

where $K(\cdot, \cdot)$ is Hermitian non-negative definite kernel of locally trace class integral operator from $L^2(\mathbb{R}^d)$ to itself.

DPP is repulsive

Fact If $|K(x, y)| \le \phi(||x - y||)$, with ϕ fast decreasing, then the DPP has weak decay of correlations.

Ex. Infinite Ginibre ensemble on complex plane clusters with kernel

$$K(z_1, z_2) = \exp(i\mathrm{Im}(z_1 \bar{z}_2) - \frac{1}{2}||z_1 - z_2||^2).$$

IV General input

Ex. 2: Zero set of Gaussian entire function

· Let $X_j, j \geq 1$, be i.i.d. standard complex Gaussians. Consider the Gaussian entire function

$$F(z) := \sum_{j=1}^{\infty} \frac{X_j}{\sqrt{j!}} z^j.$$

· Zero set $Z_F := F^{-1}(\{0\})$ is trans. invariant (in the class of Gaussian power series, it is the only one which is trans. inv.).

· Z_F exhibits local repulsivity.

· Z_F has weak decay of correlations (Nazarov and Sodin (2012)).

IV General input

Other examples of pt processes with weak decay of correlations.

· Permanental pt processes with fast decreasing kernel,

· Certain Gibbs pt processes.

IV General input: WLLN

Let $\mathcal{P}$ be a pt process on $\mathbb{R}^d$ with weak decay of correlations (wdc). Recall $\mathcal{P} \cap Q_n := \mathcal{P} \cap [\frac{-n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d$ and

$$\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d} x}.$$

Thm (BYY '19): If ξ is stabilizing wrt $\mathcal{P}$ and satisfies the p moment condition for some $p \in (1, \infty)$, then for all $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$ we have

$$\lim_{n \to \infty} n^{-1} \mathbb{E} \langle \mu_n^\xi, f \rangle = \mathbb{E}\, \xi(\mathbf{0}, \mathcal{P} \cup \{\mathbf{0}\}) \int_{[-\frac{1}{2}, \frac{1}{2}]^d} f(x) dx \cdot \rho^{(1)}(\mathbf{0}).$$

IV General input: Gaussian fluctuations

Thm (BYY '19) $\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d}x}$. Assume

- $\mathcal{P}$ has wdc
- ξ has deterministic radius of stabilization wrt $\mathcal{P}$,
- ξ satisfies the p moment condition for some $p \in (2, \infty)$, and
- $\mathrm{Var}\langle \mu_n^\xi, f \rangle = \Omega(n^\alpha)$ for some $\alpha \in (0, 1)$, $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$.

IV General input: Gaussian fluctuations

Thm (BYY '19) $\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d}x}$. Assume

- $\mathcal{P}$ has wdc
- ξ has deterministic radius of stabilization wrt $\mathcal{P}$,
- ξ satisfies the p moment condition for some $p \in (2, \infty)$, and
- $\mathrm{Var}\langle \mu_n^\xi, f \rangle = \Omega(n^\alpha)$ for some $\alpha \in (0, 1)$, $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$. Then

$$\frac{\langle \mu_n^\xi, f \rangle - \mathbb{E}\,\langle \mu_n^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_n^\xi, f \rangle}} \xrightarrow{\mathcal{D}} N(0, 1).$$

IV General input: Gaussian fluctuations

Thm (BYY '19) $\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d} x}$. Assume

· $\mathcal{P}$ has wdc

· ξ has deterministic radius of stabilization wrt $\mathcal{P}$,

· ξ satisfies the p moment condition for some $p \in (2, \infty)$, and

· $\mathrm{Var}\langle \mu_n^\xi, f \rangle = \Omega(n^\alpha)$ for some $\alpha \in (0,1)$, $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$. Then

$$\frac{\langle \mu_n^\xi, f \rangle - \mathbb{E} \langle \mu_n^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_n^\xi, f \rangle}} \xrightarrow{\mathcal{D}} N(0,1).$$

Remarks. When $\mathcal{P}$ is determinantal with fast decreasing kernel, this extends Soshnikov (2002) and Shirai + Takahashi (2003) who restrict to the count statistics $\sum_{x \in \mathcal{P}_n} \delta_{n^{-1/d} x}$, i.e., they put $\xi \equiv 1$.

· If $\mathcal{P}$ is zero set of Gaussian entire function, this extends Nazarov and Sodin (2012), who also restrict to $\sum_{x \in \mathcal{P} \cap Q_n} \delta_{n^{-1/d} x}$.

IV General input: Gaussian fluctuations

Thm (BYY '19) $\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d}x}$. Assume

- $\mathcal{P}$ wdc and decay coeff. satisfy mild growth condition
- ξ exponentially stabilizing wrt $\mathcal{P}$,
- ξ satisfies the p moment condition for some $p \in (2, \infty)$, and
- $\mathrm{Var}\langle \mu_n^\xi, f \rangle = \Omega(n^\alpha)$ for some $\alpha \in (0,1)$, $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$. Then

$$\frac{\langle \mu_n^\xi, f \rangle - \mathbb{E}\, \langle \mu_n^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_n^\xi, f \rangle}} \xrightarrow{\mathcal{D}} N(0,1).$$

IV General input: Gaussian fluctuations

Thm (BYY '19) $\mu_n^\xi := \sum_{x \in \mathcal{P} \cap Q_n} \xi(x, \mathcal{P} \cap Q_n) \delta_{n^{-1/d}x}$. Assume

· $\mathcal{P}$ wdc and decay coeff. satisfy mild growth condition

· ξ exponentially stabilizing wrt $\mathcal{P}$,

· ξ satisfies the p moment condition for some $p \in (2, \infty)$, and

· $\mathrm{Var}\langle \mu_n^\xi, f \rangle = \Omega(n^\alpha)$ for some $\alpha \in (0,1)$, $f \in B([-\frac{1}{2}, \frac{1}{2}]^d)$. Then

$$\frac{\langle \mu_n^\xi, f \rangle - \mathbb{E} \langle \mu_n^\xi, f \rangle}{\sqrt{\mathrm{Var}\langle \mu_n^\xi, f \rangle}} \xrightarrow{\mathcal{D}} N(0,1).$$

Remark. If $\mathcal{P}$ is determinantal with fast decreasing kernel (e.g. Ginibre) then $\mathcal{P}$ satisfies stated condition

IV General input: Variance asymptotics

- Given wdc input $\mathcal{P}$ and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})$$
$$+ \int_{\mathbb{R}^d} [\mathbb{E}\,\xi(\mathbf{0},\mathcal{P}\cup x)\xi(x,\mathcal{P}\cup\mathbf{0})\rho^{(2)}(\mathbf{0},x)$$
$$- \mathbb{E}\,\xi(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})\mathbb{E}\,\xi(x,\mathcal{P})\rho^{(1)}(x)]dx.$$

IV General input: Variance asymptotics

· Given wdc input $\mathcal{P}$ and a score ξ, put

$$\sigma^2(\xi) := \mathbb{E}\,\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0}) + \int_{\mathbb{R}^d}[\mathbb{E}\,\xi(\mathbf{0},\mathcal{P}\cup x)\xi(x,\mathcal{P}\cup\mathbf{0})\rho^{(2)}(\mathbf{0},x) - \mathbb{E}\,\xi(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})\mathbb{E}\,\xi(x,\mathcal{P})\rho^{(1)}(x)]dx.$$

· **Thm (BYY '19)**: If ξ is exponentially stabilizing wrt $\mathcal{P}$, if ξ satisfies the p moment condition for some $p \in (2,\infty)$, then for all $f \in B([-\frac{1}{2},\frac{1}{2}]^d)$ we have

$$\lim_{n\to\infty} n^{-1}\mathrm{Var}\langle\mu_n^{\xi}, f\rangle = \sigma^2(\xi)\int_{[-\frac{1}{2},\frac{1}{2}]^d} f^2(x)dx \in [0,\infty).$$

· **Rk.** When $\mathcal{P}$ is determinantal with fast decreasing kernel this extends Soshnikov (2002), who assumes $\xi \equiv 1$.

Cumulants. For a random variable Y with all finite moments, expanding the logarithm of the Laplace transform in a formal power series gives

$$\log \mathbb{E}\,(e^{tY}) = \log(1 + \sum_{k=1}^{\infty} \frac{M_k t^k}{k!}) = \sum_{k=1}^{\infty} \frac{S_k t^k}{k!},$$

where $M^k = \mathbb{E}\,(Y^k)$ is the k th moment of Y and $S_k = S_k(Y)$ denotes the k th cumulant of Y.

Both of the above series can be considered as formal ones and no additional condition (on exponential moments of Y) are required for the cumulants to exist.

Cumulants.

$$\log \mathbb{E}\left(e^{tY}\right) = \log(1 + \sum_{k=1}^{\infty} \frac{M^k t^k}{k!}) = \sum_{k=1}^{\infty} \frac{S_k t^k}{k!},$$

We have

$$S_k = \sum_{\gamma \in \Pi[k]} (-1)^{|\gamma|-1}(|\gamma|-1)! \prod_{i=1}^{|\gamma|} M^{|\gamma(i)|},$$

where $\Pi[k]$ is the set of all unordered partitions of the set $\{1, ..., k\}$, and for a partition $\gamma = \{\gamma(1), \ldots, \gamma(l)\} \in \Pi[k]$, $|\gamma|$ denotes the number of its elements (in this case $|\gamma| = l$), while $|\gamma(i)|$ the number of elements of subset $\gamma(i)$. In view of the above, the existence of the kth cumulant S_k follows from the finiteness of the moment M^k.

First cumulant is the mean, second cumulant is the variance.

Proof idea for CLT.

· Let $X_n, n \geq 1$, be mean zero random variables, $\mathrm{Var} X_n = 1$.

· Put $c_n^k := c^k(X_n)$, $k \in \mathbb{N}$, to be kth order cumulants for X_n.

· Recall $c_n^1 = \mathbb{E}\, X_n = 0,\ c_n^2 = \mathrm{Var} X_n$.

· **Classic Theorem.** If $\lim_{n\to\infty} c_n^k = 0$ for all k large, then $X_n \xrightarrow{\mathcal{D}} N(0,1)$ as $n \to \infty$.

Proof idea for CLT.

· Let $X_n, n \geq 1$, be mean zero random variables, $\mathrm{Var} X_n = 1$.

· Put $c_n^k := c^k(X_n)$, $k \in \mathbb{N}$, to be kth order cumulants for X_n.

· Recall $c_n^1 = \mathbb{E}\, X_n = 0,\ c_n^2 = \mathrm{Var} X_n$.

· **Classic Theorem.** If $\lim_{n\to\infty} c_n^k = 0$ for all k large, then $X_n \xrightarrow{\mathcal{D}} N(0,1)$ as $n \to \infty$.

The next corollary gives a CLT when the cumulants have linear growth.

· **Corollary**. If $Y_n, n \geq 1$, are mean zero random variables with $c_n^k = O(n)$ for all k large, $\mathrm{Var} Y_n \geq n^\alpha$ for some $\alpha \in (0, \infty)$, then $Y_n/\sqrt{\mathrm{Var} Y_n} \xrightarrow{\mathcal{D}} N(0,1)$ as $n \to \infty$.

Proof idea for CLT

· To show $\langle \mu_n^\xi, f \rangle / \sqrt{\mathrm{Var}\langle \mu_n^\xi, f \rangle} \xrightarrow{\mathcal{D}} N(0,1)$, by the previous Corollary it suffices to show that kth order cumulant for $\langle \mu_n^\xi, f \rangle$ is $O(n)$.

· Given ξ, consider k mixed moment functions $m_{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}$ given by

$$m_{(k)}(x_1, ..., x_k; \mathcal{P}_n) := \mathbb{E}\, \Pi_{i=1}^k \xi(x_i, \mathcal{P}_n) \rho^{(k)}(x_1, ..., x_k).$$

Proof idea for CLT

· To show $\langle \mu_n^\xi, f\rangle / \sqrt{\mathrm{Var}\langle \mu_n^\xi, f\rangle} \xrightarrow{\mathcal{D}} N(0,1)$, by the previous Corollary it suffices to show that kth order cumulant for $\langle \mu_n^\xi, f\rangle$ is $O(n)$.

· Given ξ, consider k mixed moment functions $m_{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}$ given by

$$m_{(k)}(x_1, ..., x_k; \mathcal{P}_n) := \mathbb{E}\, \Pi_{i=1}^k \xi(x_i, \mathcal{P}_n) \rho^{(k)}(x_1, ..., x_k).$$

· Need to show that the mixed moments 'cluster', that is for all $k \in \mathbb{N}$ there are constants c_k and C_k s.t. for all $x_1, ..., x_{p+q} \in \mathbb{R}^d$,

$$|m_{(p+q)}(x_1, ..., x_{p+q}) - m_{(p)}(x_1, ..., x_p) m_{(q)}(x_{p+1}, ..., x_{p+q})| \leq C_{p+q} \varphi(-c_{p+q} s$$

where φ is fast decreasing and

$$s := \inf_{i \in \{1,...,p\},\ j \in \{p+1,...,p+q\}} ||x_i - x_j||.$$

· $\mathcal{P}$ has wdc and ξ exp. stabilizing $\Rightarrow$ mixed moments cluster. $\square$

IV General input: Applications

These general results immediately yield limit theory (WLLN, Gaussian fluctuations, variance asymptotics) for statistics of geometric structures on pt processes with wdc. This includes:

· Vietoris-Rips clique count on any pt process with wdc, including DPP with fast decreasing kernel, zero set of Gaussian entire funct.

IV General input: Applications

These general results immediately yield limit theory (WLLN, Gaussian fluctuations, variance asymptotics) for statistics of geometric structures on pt processes with wdc. This includes:

· Vietoris-Rips clique count on any pt process with wdc, including DPP with fast decreasing kernel, zero set of Gaussian entire funct.

· total volume and surface area of germ-grain model with germs given by points in pt process with wdc, i.i.d. grains with bounded diameter.

IV General input: Applications

These general results immediately yield limit theory (WLLN, Gaussian fluctuations, variance asymptotics) for statistics of geometric structures on pt processes with wdc. This includes:

· Vietoris-Rips clique count on any pt process with wdc, including DPP with fast decreasing kernel, zero set of Gaussian entire funct.

· total volume and surface area of germ-grain model with germs given by points in pt process with wdc, i.i.d. grains with bounded diameter.

· WLLN and variance asymptotics for total edge length in knn graph on DPP with fast decreasing kernel. Total edge length has Gaussian fluctuations (subject to lower bounds on variance).

IV General input: Applications

These general results immediately yield limit theory (WLLN, Gaussian fluctuations, variance asymptotics) for statistics of geometric structures on pt processes with wdc. This includes:

· Vietoris-Rips clique count on any pt process with wdc, including DPP with fast decreasing kernel, zero set of Gaussian entire funct.

· total volume and surface area of germ-grain model with germs given by points in pt process with wdc, i.i.d. grains with bounded diameter.

· WLLN and variance asymptotics for total edge length in knn graph on DPP with fast decreasing kernel. Total edge length has Gaussian fluctuations (subject to lower bounds on variance).

THANK YOU