Different scaling regimes for geometric Poisson functionals

Raphaël Lachieze-Rey, Univ. Paris 5 René Descartes

Joint work with Giovanni Peccati, University of Luxembourg

Framework

- $\eta_\lambda$: Poisson measure with intensity $\mu_\lambda$.
- Chaos decomposition: $F_\lambda = F(\eta_\lambda) = \sum_{q=1}^{k} F_{q,\lambda}$, $L^2$ variable for each $\lambda$.
- Asymptotic regime of renormalized variables

$$\tilde{F}_\lambda = \frac{F_\lambda - \mathbb{E}F_\lambda}{\sqrt{\text{var}(F_\lambda)}}$$

$$\tilde{F}_{q,\lambda} = \frac{F_{q,\lambda} - \mathbb{E}F_{q,\lambda}}{\sqrt{\text{var}(F_{q,\lambda})}}$$

- $\mathcal{N}$: Standard law
- $\mathcal{P}(c)$: Poisson law with parameter $c$
- $d_W$: Wasserstein distance.
Finite decompositions and $U$-statistics

Under proper integrability assumptions ($\eta$: Poisson point process):

- $k$-th order stochastic integral with kernel $f$:

$$I_k(f) = \int_{\mathcal{X}^k} f(x_k) d(\eta - \mu)^\otimes k.$$ 

- $k$-th order $U$-statistic with kernel $h$:

$$U_k(h) = \sum_{x_k \subseteq \eta_\lambda} h(x_k) = \int_{\mathcal{X}^k} h(x_k) d\eta^k(x_k).$$

- Each $k$-tuple of points gives a contribution independent of the other points of $\eta_\lambda$. 
Geometric framework

- $\eta_\lambda$: Marked Poisson process.
- $\ell$: Lebesgue measure
- $X_\lambda = [-\lambda^{1/d}, \lambda^{1/d}]$
- $(M, \nu)$: Marks probability space
- $x = (t, m)$: marked point.
- $\mu_\lambda = 1_{X_\lambda} \ell \otimes \nu$ (Lebesgue measure).
- Kernel scale change $\Rightarrow$ Equivalent to $\mu_\lambda = 1_{X_1} \lambda \ell \otimes \nu$. 
Graph model with unbounded connections

- $H_\lambda \subseteq \mathbb{R}^d$ measurable.
- $x, x' \in \eta$ connected if $x - x' \in H_\lambda$.
- If $H_\lambda = \text{Unit ball} \Rightarrow \text{Unit disk graph (Boolean model)}$.
- $F_\lambda$: Number of connections.

**Interaction volume:**

$$v_\lambda := \ell(H_\lambda \cap X_\lambda).$$
Different regimes: $H_\lambda = \alpha_\lambda H_1$

$H_1$:

\[ a = 1, \alpha_\lambda = 1, \lambda = 25: \]

- CLT at speed $\lambda^{-1/2}$

\[ a = 1, \alpha_\lambda = \lambda^{-1/d}, \lambda = 50: \]

- CLT in $\log(\lambda)^{-1}$

\[ a = 2, \alpha_\lambda = \lambda^{-1/d}, \lambda = 50: \]

- no CLT.
(R0) $\lambda v_\lambda$ is bounded.
(R1) $v_\lambda \to 0$ and $\lambda v_\lambda \to \infty$.
(R2) $v_\lambda \to c > 0$
(R3) $v_\lambda \to \infty$.

Theorem

Assume $O$-regular variation: $0 < c_1 \leq \frac{v_\lambda}{v_{c_\lambda}} \leq c_2 < \infty$ for $c > 0$.
There are $C_i, C_i' > 0$ such that

(R3) $\text{var}(F_\lambda) \sim C_3 \lambda v_\lambda^2$

CLT: $d_W(\tilde{F}_\lambda, N) \leq C_3' \lambda^{-1/2}$

(R2) $\text{var}(\tilde{F}_\lambda, N^\epsilon) \sim C_2 \lambda$

CLT: $d_W(\tilde{F}_\lambda, N) \leq C_2' \lambda^{-1/2}$

(Sandard behaviour)

(R1) $\text{var}(F_\lambda(H)) \sim C_1 \lambda v_\lambda$

CLT: $d_W(\tilde{F}_\lambda, N^\epsilon) \leq C_1'(\lambda v_\lambda)^{-1/2}$

(R0) $\tilde{F}_\lambda$ converges to a Poisson law (or to 0).
Hierarchy of chaoses

As $\nu_\lambda$ grows, $F_{1,\lambda}$ becomes more predominant.

- Under (R3)

$$\frac{\text{var}(F_{1,\lambda})}{\text{var}(F_{2,\lambda})} \to \infty,$$

the first chaos dominates.

- Under (R2)

$$0 < c' \leq \frac{\text{var}(F_{1,\lambda})}{\text{var}(F_{2,\lambda})} \leq C' < \infty$$

- Under (R1),(R0)

$$\frac{\text{var}(F_{2,\lambda})}{\text{var}(F_{1,\lambda})} \to \infty$$

the second chaos dominates.

Remark: CLT $\Leftrightarrow \text{var}(F_\lambda) \to \infty.$
Summary:

- **(R3) Large interactions**: CLT, first chaos dominates.
- **(R2) Constant size interactions**: CLT, chaos co-dominates.
- **(R1) Small interactions**: Slow CLT, 2d chaos dominates.
- **(R0) Rare interactions**: Poisson limit, 2d chaos dominates.
Subgraph counting

- $G$: connected formal graph with cardinality $k \geq 1$.
- $\eta_\lambda$: Homogeneous Poisson process on $X_\lambda$.
- $G_\lambda$: Graph obtained by connecting points $(x, y) \in \eta_\lambda$ with distance $\|x - y\| \leq \alpha_\lambda$. $\nu_\lambda := \alpha_\lambda^d$.
- $F_\lambda(G)$: Number of occurrences of $G$ as a subgraph of $G_\lambda$. 

![Diagram of a graph with vertices A, B, C, D, E and a modified graph showing subgraph counting.]
\begin{itemize}
  \item \textbf{(R1)} $\nu_\lambda \to 0$ and $\lambda \nu_\lambda^{k-1} \to \infty$.
  \item \textbf{(R2)} $\nu_\lambda \to c > 0$
  \item \textbf{(R3)} $\nu_\lambda \to \infty$.
\end{itemize}

\textbf{Results:} Penrose, Peccati, LR.

\begin{itemize}
  \item \textbf{(R1)} $\text{var}(F_\lambda) \sim c_1 \lambda \nu_\lambda^{k-1}$ \textbf{CLT:} $d_W(\tilde{F}_\lambda, \mathcal{N}) \leq C_1 (\lambda \nu_\lambda^{k-1})^{-1/2}$
  \item \textbf{(R2)} $\text{var}(F_\lambda) \sim c_2 \lambda$ \textbf{CLT:} $d_W(\tilde{F}_\lambda, \mathcal{N}) \leq C_2 \lambda^{-1/2}$
  \item \textbf{(R3)} $\text{var}(F_\lambda) \sim c_3 \lambda \nu_\lambda^{2k-2}$ \textbf{CLT:} $d_W(\tilde{F}_\lambda, \mathcal{N}) \leq C_3 \lambda^{-1/2}$
\end{itemize}
Stationary rescaled marked kernels

- $k \geq 2$.

\[ F_\lambda = U_k(h_\lambda) = \sum_{y_1, \ldots, y_k \in \eta_\lambda} h_\lambda(y_1, \ldots, y_k), \quad y_i = (t_i, m_i) \in \mathbb{R}^d \times \mathcal{M}. \]

**Assumptions on $h_\lambda$:**

- $h_\lambda(\cdot) \sim h(\alpha_\lambda \cdot)$ in some sense.
- $\alpha_\lambda$: Scaling regime. ( $\nu_\lambda := \alpha_\lambda^d$: Interaction measure. )
- $h(y_1, \ldots, y_k)$ invariant under **spatial translations**.
- $h$ is **rapidly decreasing** away from the diagonal: there exists $\kappa(y) > 0$ bounded probability density such that for $p = 2, 4$,

\[
\int_{\mathcal{M} \times (\mathbb{R}^d \times \mathcal{M})^{k-1}} \frac{|h(0, m_0, y_1, \ldots, y_{k-1})|^p}{\kappa(y_1) \cdots \kappa(y_{k-1})^{p-1}} \mu \otimes \nu(dy_1, \ldots, dy_{k-1}) < \infty.
\]
Tool: Bounds on the contractions

Theorem

For $F = \sum_{q=1}^{k} l_q(f_q) \in L^2$ with $\text{var}(F) = 1$,

$$d_W(F, \mathcal{N}) \leq C(k)(\max_{q} \| f_q \star_l^r f_q' \|_2 + \max_{q} \| f_q \|_4)$$

where $1 \leq l \leq r \leq q \leq q'$ and $l \neq q'$

$$f_q \star_r^l f_q'(x_{r-l}, y_{q-r}, y'_{q'-r}) = \int f_q(x_l, x_{r-l}, y_{q-r}) f_q'(x_l, x_{r-l}, y'_{q'-r}) dx_l.$$ 

Theorem

If $h(x_k)$ and $g(x_q)$ are stationary and rapidly decreasing, for $r, l$ as above,

$$\| h \star_r^l g \|_{L^2(X \times (\mathbb{R}^d)^{k+q-r-l-1})}^2 \leq \ell(X)A(h)A(g)$$
Chaos behaviour

\[ F_\lambda = U_k(h_\lambda) = \sum_{q=1}^{k} F_{q,\lambda}. \]

**Theorem**

\( q \)-th chaos behaviour of \( U_k(h_\lambda) \):

\[ \text{var}(F_{q,\lambda}) \sim C_{\lambda} \nu_{\lambda}^{2k-q-1} \]

\[ d_W(\tilde{F}_{q,\lambda}, \mathcal{N}) \leq C' \sqrt{\frac{\nu_{\lambda}^{1-q}}{\lambda}} \left( 1 + 1_{\{q \neq 1\}} \right) \left\{ \begin{array}{ll} \nu_{\lambda}^q & \text{if } \nu_{\lambda} > 1 \\ \nu_{\lambda} & \text{otherwise} \end{array} \right\} \]

First term: kernel 4-th moments. second term: kernel contractions.

- First chaoses win if \( \nu_{\lambda} \to \infty \).
- Last chaoses win if \( \nu_{\lambda} \to 0 \).
- High order chaoses convergence is slower.
Applications

- $v_{\lambda} = \lambda$: All points interact $\Rightarrow$ Geometric U-statistics.
- $v_{\lambda} = 1$: Standard behaviour/Thermodynamic regime.
- $v_{\lambda}$ small: rarefaction of interactions: Slow CLT/no CLT.
Examples of functionals with a standard behaviour

**Boolean model:**

- $\mathcal{M}$: Compact sets (with Fell Borel $\sigma$-algebra).
- $\{M_k; k \geq 1\}$ IID Random compact sets.
- $\{x_k; k \geq 1\}$ Poisson point process with intensity $\lambda$.
- $\eta_\lambda = \{(x_k, M_k)\}$ marked Poisson measure.

\[
R_\lambda = \bigcup_{k: x_k \in X_\lambda} (M_k \oplus x_k)
\]

- $F_\lambda$: $U$-statistic with stationary kernel

\[
h((x, M); (x', M')) = \varphi(x - x')1\{(M \oplus x) \cap (M' \oplus x') \neq \emptyset\},
\]

\[
F_\lambda = \sum_{x_i \neq x_j \in \eta_\lambda} \varphi(x_i - x_j)1\{\text{The grains with centers } x_i \text{ and } x_j \text{ touch}\}.
\]
Magnitude assumption on $\varphi$

$$\varphi(x - y) \leq \|x - y\|^\beta, \; x, y \in \mathbb{R}^d.$$ 

**Theorem**

Assume that in the boolean model the typical grain has diameter $R$ such that for some $\varepsilon > 0$, $\beta > -d/2$, and for $r \geq 1$

$$\mathbb{P}(R \geq r) \leq Cr^{-(2(\beta+d)+1+\varepsilon)}$$

then for some $C, C' > 0$

$$\text{var}(F_\lambda) \sim C\lambda$$

$$d_W(\tilde{F}_\lambda, \mathcal{N}) \leq C'\lambda^{-1/2}.$$ 

The optimal condition bears actually upon the decay of

$$\chi(x) = \mathbb{P}(M_1 \cap (M_2 \oplus x) \neq \emptyset), \; x \in \mathbb{R}^d.$$
Number of intersections in a process of line segments.

- $M_k; k \geq 1$ Line segments with random IID lengths.
- $\varphi(x, y) = 1$.

Then $F_\lambda$ is the number of intersections of segments with centers in $X_\lambda$.

**Figure: Zbynek Pawlas**

**Theorem (Pawlas 2012)**

There is standard behaviour if for some $\varepsilon > 0$

$$\mathbb{P}(\text{length}(M_1) \geq r) \leq Cr^{-5-\varepsilon}, \ r \geq 1.$$
sub-hypergraph counting and telecommunications network

Decreusefond et al.

- \( \eta_\lambda = \{x_k; k \geq 1\} \) Poisson point process with intensity \( \lambda \ell \) on the torus.
- \( k \)-th order hypergraph \( \mathcal{H}_{k,\lambda} \): Data of \( k \)-tuples \( (x_1, \ldots, x_k) \in \eta_\lambda \) such that \( \|x_i - x_j\| \leq \varepsilon \)
- Exact mean formulas for \( \mathcal{H}_{k,\lambda} \) and related topological quantities.

Our approach:

- \( \{m_k; k \geq 1\} \) random radii around the points.
- Edge effects: \( \eta_\lambda = \{y_k = (x_k, m_k); k \geq 1\} \) on \( X_\lambda = [\lambda^{1/d}, \lambda^{1/d}] \).
- \( \tilde{\mathcal{H}}_{k,\lambda} \): \( k \)-tuples \( (y_1, \ldots, y_k) \) such that for all \( i \neq j \), \( \|x_i - x_j\| \leq m_i \).

Results (\( R \) = typical radius):

**Theorem**

If \( \mathbb{E} R^{4d+\varepsilon} < \infty \),

\[
\text{var}(F_{k,\lambda}) \sim C\lambda \\
d_W(\tilde{\mathcal{H}}_{k,\lambda}, \mathcal{N}) \leq C'\lambda^{-1/2}
\]
Geometric U-statistics: $\alpha_\lambda = \lambda^{1/d}$

Points interact regardless of the distance:

$$F_\lambda = \sum_{y_k \in \eta_\lambda} h(\lambda^{-1/d} y_1, \ldots, \lambda^{-1/d} y_k).$$

General form of a Geometric U-statistic:

- $(X, \mu)$ loc. compact measured space
- $\mu_\lambda = \lambda \mu$
- $h(y_1, \ldots, y_k):$ kernel on $X^k$.

$$F_\lambda = \sum_{y_k \in \eta_\lambda} h(x_1, \ldots, x_k).$$

**Examples** [Reitzner and Schulte]

- Number of $k$-tuples of points in convex position $\Rightarrow$ Approximation of Sylvester's constant in a convex body.
- Intersections of flats in a compact window.
Results

\[ F_\lambda = \sum_{q=1}^{k} \kappa_{k,q} F_{q,\lambda} \]

and \( F_{q,\lambda} = \lambda^{k-q} I_q(h_q) \)

Kernel projections:

\[ h_q(y_q) = \int_{\chi_q} h(y_q, y_{k-q}) d\mu(y_q) \]

Asymptotic behaviour:

**Theorem**

\[ \text{var}(F_{q,\lambda}) \sim C\lambda^{2k-q} \]

\[ F_{q,\lambda} \rightarrow G_q(h_q) \]

where \( G_q(h_q) \) is a Gaussian chaos of order \( q \).
Summary for geometric U-statistics

\[ q_0 = \min \{ q : \| h_q \|_2 \neq 0 \} \]

\( F_\lambda \) behaves like \( F_{q_0, \lambda} \).

- Reitzner and Schulte. \( q_0 = 1 \Rightarrow \) CLT at speed \( \lambda^{-1/2} \)
- LR and Peccati: \( q_0 \geq 2 \Rightarrow \) no CLT, convergence to \( G_{q_0} (h_{q_0}) \).
- Peccati and Thaele: \( q_0 = 2 \): Speed of convergence to \( G_2 \), a Gamma random variable.
Poisson regime: $\lambda \alpha \lambda \rightarrow c$

Peccati (2011) : sufficient conditions for the convergence to a Poisson law in terms of the contractions.
Mixed chaos behaviour

**Multi-dimensional CLTs**: Peccati, Zengh, Minh, Schulte, Thaele, Last, Penrose, Reitzner, LR ...

**Peccati, Bourguain 2012**: Portmanteau inequalities ⇒ Mixed limit theorems.

**Example.** Disk graph with influence volume $v_\lambda$ such that $\lambda v_\lambda^{3-1} \to 0$ and $\lambda v_\lambda^{2-1} \to \infty$. Consider

$$F_{2,\lambda} = \# \text{segments}$$

$$F_{3,\lambda} = \# \text{triangles}.$$  

Then

$$(\tilde{F}_{2,\lambda}, \tilde{F}_{3,\lambda}) \to (\mathcal{N}, \mathcal{P})$$

where $\mathcal{N}$ and $\mathcal{P}$ are independent.