Karlsruhe Institute of Technology

Cluster counting in the random connection model

Franz Nestmann,
Karlsruhe Institute of Technology

joint work with Günter Last (Karlsruhe) and Matthias Schulte (Bern)

19th Workshop on Stochastic Geometry, Stereology and Image Analysis (SGSIA)

Stationary Poisson processes

Setting

Let η be a **stationary Poisson process** *on $\mathbb{R}^d$ with* **intensity** $\beta > 0$.

- η has intensity measure $\lambda := \beta \cdot \lambda_d$.
- The Poisson process η can be represented as $\eta = \sum_{n=1}^{\infty} \delta_{X_n}$, where the X_n, $n \in \mathbb{N}$ are random elements in $\mathbb{R}^d$.

The classical RCM

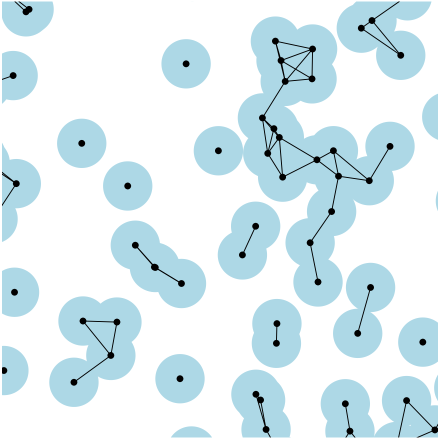
Setting

Let $\varphi : \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$ *be a measurable and symmetric* **connection function**.
Given η, *connect any two points* $x, y \in \eta$, $x \neq y$, *with probability*

$$\varphi(x, y) = \mathbb{P}(x \leftrightarrow y)$$

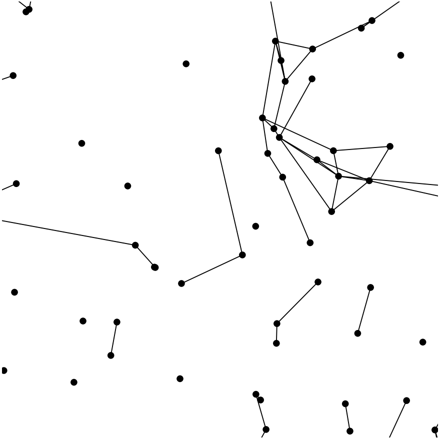
independently of all other pairs.
This gives the **random connection model** $\Gamma_\varphi(\eta) = (\eta, \chi)$, *where* χ *is the point process of the edges.*

Examples



- $\varphi(x, y) = \mathbf{1}\{\|x - y\| \le 2r\}, \; x, y \in \mathbb{R}^d$ with $r > 0$ *(Gilbert graph)*

Examples



- $\varphi(x,y) = \exp(-a\|x-y\|),\ x,y \in \mathbb{R}^d$ with $a > 0$

The marked RCM

Setting

Let η be an **independent $\mathbb{Q}$-marking** *of a stationary Poisson process η' on $\mathbb{R}^d$ with intensity $\beta > 0$, where $\mathbb{Q}$ is a distribution on $[0, \infty)$. The (marked) Poisson process η can be represented as*

$$\eta = \sum_{n=1}^{\infty} \delta_{(X_n, W_n)},$$

where the X_n, $n \in \mathbb{N}$ are random elements in $\mathbb{R}^d$ and $(W_n)_{n\in\mathbb{N}}$ is an iid-sequence of random variables on $[0, \infty)$ with distribution $\mathbb{Q}$, independent of $(X_n)_{n\in\mathbb{N}}$.

- η is a Poisson process on $\mathbb{R}^d \times [0, \infty)$ with **intensity measure** $\lambda := \beta\lambda_d \otimes \mathbb{Q}$.

The marked RCM

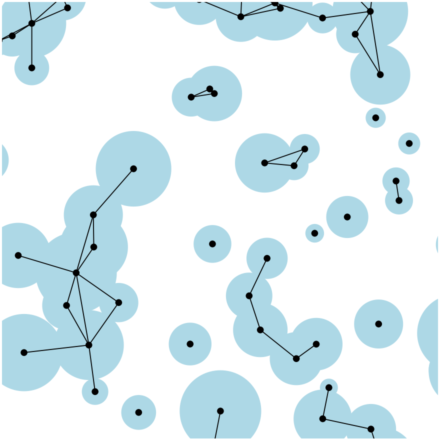
Setting

Let $\varphi : (\mathbb{R}^d \times [0,\infty))^2 \to [0,1]$ *be a measurable and symmetric* **connection function**.
Given η*, connect any two points* $(x,w),(y,w') \in \eta$*,* $(x,w) \neq (y,w')$*, with probability*

$$\varphi((x,w),(y,w')) = \mathbb{P}((x,w) \leftrightarrow (y,w'))$$

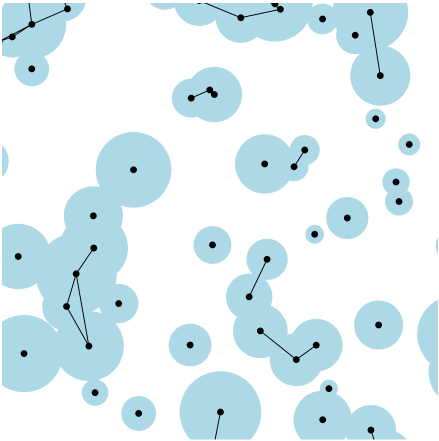
independently of all other pairs.
This gives the **marked random connection model** $\Gamma_\varphi(\eta) = (\eta,\chi)$*, where* χ *is the point process of the edges.*

Examples



- $\varphi((x,w),(y,w')) = \mathbf{1}\{\|x-y\| \leq w+w'\}$,
  $(x,w),(y,w') \in \mathbb{R}^d \times [0,1]$ *(Gilbert graph with random radii)*

Examples



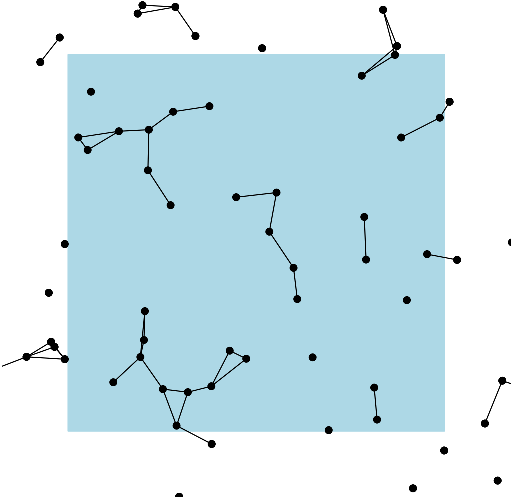
- $\varphi((x,w),(y,w')) = \mathbf{1}\{\|x-y\| \le w+w'\} \cdot \psi(x,y)$, $(x,w),(y,w') \in \mathbb{R}^d \times [0,1]$ with $\psi : \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$

Connected components and isomorphic graphs

- For $k \in \mathbb{N}$ let G be a connected graph with k vertices.
- For a compact and convex set $A \subset \mathbb{R}^d$ with $\lambda_d(A) > 0$ define

$$\eta_{\varphi,G}(A) := \#\{\text{clusters of } \Gamma_\varphi(\eta) \text{ isomorphic to } G \\ \text{with lexicographic minimum in } A\}.$$

Another picture



Expectation

Proposition (Last, N., Schulte 2017+)

Let $A \subset \mathbb{R}^d$ be compact and convex with $\lambda_d(A) > 0$. Then,

$$\mathbb{E}\eta_{\varphi,G}(A) = \int \mathbf{1}\{x_1 \in A \times [0,\infty)\} p_{\varphi,G}(x_1,\dots,x_k) \times \exp\left(\beta \int \left(\prod_{i=1}^{k}(1-\varphi(x_i,y)) - 1\right)\lambda(dy)\right) \lambda^k(d(x_1,\dots,x_k)),$$

where

$$p_{\varphi,G}(x_1,\dots,x_k) := \mathbf{1}\{x_1 < \cdots < x_k\}\mathbb{P}(\Gamma_\varphi(\delta_{x_1} + \cdots + \delta_{x_k}) \simeq G),$$

$x_1,\dots,x_k \in \mathbb{R}^d \times [0,\infty)$.

- The proof uses the multivariate Mecke equation and the formula for the generating functional of a Poisson process.

Assumptions

- Let G be a connected graph with k vertices, that occurs in $\Gamma_\varphi(\eta)$ with positive probability.
- Let $g : [0,\infty)^3 \to [0,\infty)$ be a measurable function, which is symmetric and increasing in the first two arguments. For $(x,w),(y,w') \in \mathbb{R}^d \times [0,\infty)$ assume that,

$$\varphi((x,w),(y,w')) = \varphi(\|x-y\|, w, w') = \mathbb{P}(\|x-y\| \leq g(w,w',S)),$$

where S is a random variable on $[0,\infty)$, independent of everything.

  - $g(w,w',s) = w + w'$, $w,w',s \in [0,\infty)$ yields the Gilbert graph with random radii.
  - $g(w,w',s) = s$, $w,w',s \in [0,\infty)$ yields the classical RCM.

A lower bound for the variance

- Assume that

$$\int \mathbb{E}\varphi(\|x\|, W, W)\, dx < \infty,$$

where W is a random variable on $[0, \infty)$ with distribution $\mathbb{Q}$.

Theorem (Last, N., Schulte 2017+)

There is a constant $c > 0$ such that

$$Var\,(\eta_{\varphi,G}(A)) \geq c \cdot \lambda_d(A),$$

for all compact and convex sets $A \subset \mathbb{R}^d$.

Probability distances

Definition (Kolmogorov distance)

For two random variables X and Y in $\mathbb{R}$ let

$$d_K(X,Y) := \sup_{u \in \mathbb{R}} \left| \mathbb{P}(X \leq u) - \mathbb{P}(Y \leq u) \right| .$$

Definition (Wasserstein distance)

For two random variables X and Y in $\mathbb{R}$ let

$$d_1(X,Y) := \sup_{h \in Lip(1)} \left| \mathbb{E}h(X) - \mathbb{E}h(Y) \right| ,$$

where $Lip(1)$ is the set of all functions $h : \mathbb{R} \to \mathbb{R}$ with a Lipschitz constant less than or equal to one.

Quantitative CLT for the classical RCM

- Classical RCM: η is a stationary Poisson process on $\mathbb{R}^d$.
- Let N be a standard Gaussian random variable.
- Let $r(A)$ denote the inradius of a compact and convex set $A \subset \mathbb{R}^d$.

Theorem (Last, N., Schulte 2017+)

Assume that

$$\int_{\mathbb{R}^d} \varphi(\|x\|)^{1/3}\, dx < \infty.$$

Then, there is a constant $c > 0$ such that

$$d_K\left(\frac{\eta_{\varphi,G}(A) - \mathbb{E}\eta_{\varphi,G}(A)}{\sqrt{Var\left(\eta_{\varphi,G}(A)\right)}}, N\right) \leq \frac{c}{\sqrt{\lambda_d(A)}},$$

for all compact and convex sets $A \subset \mathbb{R}^d$ with $r(A) \geq 1$.

- The assertion also holds for the Wasserstein distance.

Quantitative CLT for the marked RCM

- Marked RCM: η is an independent $\mathbb{Q}$-marking of a stationary Poisson process η' on $\mathbb{R}^d$.
- Remember: $\varphi(\|x-y\|, w, w') = \mathbb{P}(\|x-y\| \le g(w, w', S))$.

Theorem (Last, N., Schulte 2017+)

Assume that

$$\mathbb{E}g(W, W, S)^{11d+1} < \infty.$$

Then, there is a constant $c > 0$ such that

$$d_K\left(\frac{\eta_{\varphi,G}(A) - \mathbb{E}\eta_{\varphi,G}(A)}{\sqrt{Var\left(\eta_{\varphi,G}(A)\right)}}, N\right) \le \frac{c}{\sqrt{\lambda_d(A)}},$$

for all compact and convex sets $A \subset \mathbb{R}^d$ with $r(A) \ge 1$.

- The assertion also holds for the Wasserstein distance.

CLT for the RCM

Remark

Penrose '03 proved the CLT for the Gilbert graph with deterministic radii while van de Brug and Meester '04 proved the CLT for the classical RCM in the case of a connection function with compact support.

Pairwise marking of Poisson processes

- Let η be a Poisson process on a measurable space $(\mathbf{X}, \mathcal{X})$ with σ-finite intensity measure λ. η can be represented as

$$\eta = \sum_{n=1}^{\kappa} \delta_{X_n},$$

where the X_n, $n \in \mathbb{N}$ are random elements in $\mathbf{X}$ and κ is a random element in $\mathbb{N} \cup \{0, \infty\}$.

- Let $(\mathbf{M}, \mathcal{M})$ be a further measurable space and let $(Z_{m,n})_{m,n\in\mathbb{N}}$ be iid-sequence of random elements in $\mathbf{M}$ with common distribution $\mathbb{M}$, independent of η. Then

$$\xi := \sum_{m,n=1}^{\kappa} \mathbf{1}\{X_m < X_n\} \delta_{(\{X_m, X_n\}, Z_{m,n})},$$

where $<$ is a partial order on $\mathbf{X}$.

The classical RCM again

- To define the RCM we use $\mathbf{M} = [0,1]$ and $\mathbb{M} = \lambda_1(\,\cdot\,\cap[0,1])$. Then

$$\chi := \sum_{m,n=1}^{\kappa} \mathbf{1}\{X_m < X_n\}\mathbf{1}\{Z_{m,n} \leq \varphi(X_m, X_n)\}\delta_{\{X_m, X_n\}}$$

is the point process of the edges.

Adding deterministic points

- Let L_ξ be the space of all $\sigma(\xi)$-measurable random variables of $\mathbb{R}$.
- For each $F \in L_\xi$ there is a measurable **representative** f such that $F = f(\xi)$.
- Extend the sequence $(Z_{m,n})_{m,n\in\mathbb{N}}$ to $(Z_{m,n})_{m,n\in\mathbb{Z}}$.
- For $x_1, x_2 \in \mathbf{X}$ define

$$\xi_{x_1} := \sum_{m,n\in\{-1,1,\dots,\kappa\}} \mathbf{1}\{X_m < X_n\}\delta_{(\{X_m,X_n\},Z_{m,n})},$$

$$\xi_{x_1,x_2} := \sum_{m,n\in\{-2,-1,1,\dots,\kappa\}} \mathbf{1}\{X_m < X_n\}\delta_{(\{X_m,X_n\},Z_{m,n})},$$

where $X_{-i} := x_i$, $i \in \{1, 2\}$.

The difference operators

- For $x_1, x_2 \in \mathbf{X}$ and $F = f(\xi) \in L_\xi$ define

$$\Delta_{x_1} F := f(\xi_{x_1}) - f(\xi),$$
$$\Delta^2_{x_1,x_2} F := f(\xi_{x_1,x_2}) - f(\xi_{x_1}) - f(\xi_{x_2}) + f(\xi).$$

Second order Poincaré inequality

Theorem (Last, N., Schulte 2017+)

Let $F \in L_\xi$ be such that $\mathbb{E}F = 0$ and Var $F = 1$. Then, under further integrability assumptions on F,

$$d_1(F, N) \leq \gamma_1 + \gamma_2 + \gamma_3,$$

where

$$\gamma_1 := 2\Big[\int \big[\mathbb{E}(\Delta_{x_1}F)^2(\Delta_{x_2}F)^2\big]^{1/2} \times \big[\mathbb{E}(\Delta^2_{x_1,x_3}F)^2(\Delta^2_{x_2,x_3}F)^2\big]^{1/2}\lambda^3(d(x_1,x_2,x_3))\Big]^{1/2},$$

$$\gamma_2 := \Big[\int \mathbb{E}(\Delta^2_{x_1,x_3}F)^2(\Delta^2_{x_2,x_3}F)^2\lambda^3(d(x_1,x_2,x_3))\Big]^{1/2},$$

$$\gamma_3 := \int \mathbb{E}|\Delta_x F|^3\lambda(dx).$$

Thank you for your attention!

Literature

- van de Brug, T. and Meester, R. (2004). On central limit theorems in the random connection model. *Physica A* **332**, 263-278.
- Penrose, M. D. (2003). *Random Geometric Graphs.* Oxford University Press, Oxford.
- Last, G., Nestmann, F. and Schulte, M. (2017+). Second order properties and normal approximation of functions of pairwise markings of Poisson processes with an application to the random connection model. In preparation.
- Last, G., Peccati, G. and Schulte, M. (2016). Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Th. Rel. Fields* **165**, 667-723.
- Last, G. and Penrose, M. D. (2017+). *Lectures on the Poisson Process.* To be published by Cambridge University Press.
- Penrose, M. D. (2015). Inhomogeneous random graphs, isolated vertices, and Poisson approximation. arXiv:1507.07132