university of groningen / faculty of science and engineering

# High-dimensional sparse random geometric graphs

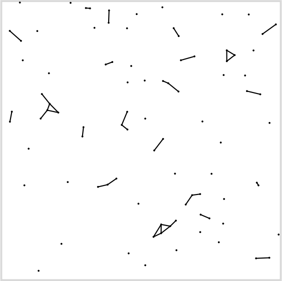
**Gilles Bonnet**

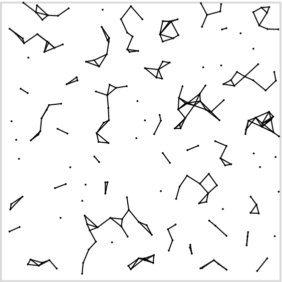
joint work with Daniel Willhalm (Groningen),
Christian Hirsch (Aarhus)
Daniel Rosen (Bochum)

Limit theory of sparse random geometric graphs in high dimensions,
Stochastic Processes and their applications, vol 163, 203-236 (2023)

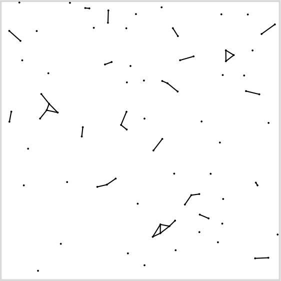
29/1/24, Dyogene seminar, Inria Paris

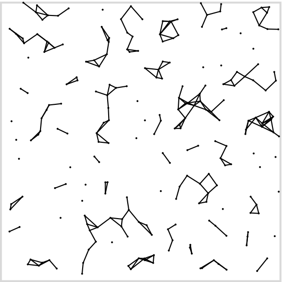
- $W_d = d$-dimensional cubical observation window with opposite facets identified = ***torus***
- $\mathcal{P}_d :=$ Poisson process in $W_d$ with ***intensity*** $\lambda_d^d > 0$
- $\mathsf{GG}_d(t) = l_\infty$-Gilbert graph with ***connection radius*** $t^{1/d}$, $0 \leqslant t \leqslant 1$





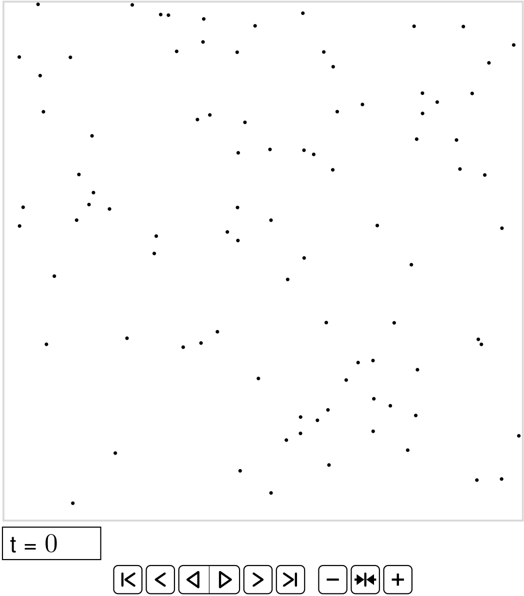
- ▷ $W_d = d$-dimensional cubical observation window with opposite facets identified = ***torus***
- ▷ $\mathcal{P}_d :=$ Poisson process in $W_d$ with ***intensity*** $\lambda_d^d > 0$
- ▷ $\mathsf{GG}_d(t) = l_\infty$-Gilbert graph with ***connection radius*** $t^{1/d}$, $0 \leqslant t \leqslant 1$





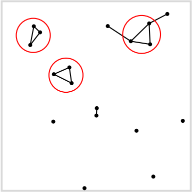
***Goal.*** Describe topological and geometric functionals as $\lambda_d \xrightarrow{d\to\infty} 0$ ***and*** $|W_d| \xrightarrow{d\to\infty} \infty$

# Animation of $(\mathrm{GG}_d(t))_{t \leqslant 1}$

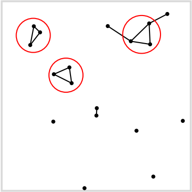


## *Subgraph count*

- $G_0$ = Fixed graph
- $\#\{G' \subseteq \mathsf{GG}_d(t) \colon G' \cong G_0\}$
- $G_0 = \Delta \rightsquigarrow$ Triangle count



## Subgraph count



- $G_0 =$ Fixed graph
- $\#\{G' \subseteq \mathsf{GG}_d(t) \colon G' \cong G_0\}$
- $G_0 = \Delta \rightsquigarrow$ Triangle count

## Betti numbers

- Geometric graph gives rise to abstract simplicial complexes, e.g. the Rips complex.
- Functional as $i$th Betti number of the clique complex on the graph $G$
- $\beta_0 =$ number of connected components
- $\beta_1 =$ number of loops
- $\beta_2 =$ number of cavities

# Additive functionals and moment asymptotics

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,t} := \mathfrak{a}(\mathsf{GG}_d(t))$, where

- $\mathfrak{a}(\cdot)$ is a ***nonnegative*** functional on abstract graphs,
- $\mathfrak{a}(\cdot)$ is ***additive***: $\mathfrak{a}(G \sqcup G') = \mathfrak{a}(G) + \mathfrak{a}(G')$,
- $\mathfrak{a}(\cdot)$ ***grows at most exponentially fast***: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- $\lambda_d \to 0$ (sparse), that $|W_d|^{\frac{1}{d}} \to \infty$ (large window).

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,t} := \mathfrak{a}(\mathsf{GG}_d(t))$, where

- $\mathfrak{a}(\cdot)$ is a ***nonnegative*** functional on abstract graphs,
- $\mathfrak{a}(\cdot)$ is ***additive***: $\mathfrak{a}(G \sqcup G') = \mathfrak{a}(G) + \mathfrak{a}(G')$,
- $\mathfrak{a}(\cdot)$ ***grows at most exponentially fast***: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- $\lambda_d \to 0$ (sparse), that $|W_d|^{\frac{1}{d}} \to \infty$ (large window).

***Theorem (Moment asymptotics)***

- $\mathbb{E}[A_{d,t}] \sim \rho_d\, t^{k_0} \frac{\sum_{G \in \mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)}{(k_0+1)!}$, *for* $0 \leqslant t \leqslant 1$, *as* $d \to \infty$
- $\mathsf{Cov}[A_{d,t_1}, A_{d,t_2}] \sim \rho_d\, t_1^{k_0} \frac{\sum_{G \in \mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)^2}{((k_0+1)!)^2}$, *for* $0 \leqslant t_1 \leqslant t_2 \leqslant 1$, *as* $d \to \infty$

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,t} := \mathfrak{a}(\mathsf{GG}_d(t))$, where

- $\mathfrak{a}(\cdot)$ is a ***nonnegative*** functional on abstract graphs,
- $\mathfrak{a}(\cdot)$ is ***additive***: $\mathfrak{a}(G \sqcup G') = \mathfrak{a}(G) + \mathfrak{a}(G')$,
- $\mathfrak{a}(\cdot)$ ***grows at most exponentially fast***: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- $\lambda_d \to 0$ (sparse), that $|W_d|^{\frac{1}{d}} \to \infty$ (large window).

***Theorem (Moment asymptotics)***

- $\mathbb{E}[A_{d,t}] \sim \rho_d\, t^{k_0} \dfrac{\sum_{G \in \mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)}{(k_0+1)!}$, *for* $0 \leqslant t \leqslant 1$, *as* $d \to \infty$
- $\mathsf{Cov}[A_{d,t_1}, A_{d,t_2}] \sim \rho_d\, t_1^{k_0} \dfrac{\sum_{G \in \mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)^2}{((k_0+1)!)^2}$, *for* $0 \leqslant t_1 \leqslant t_2 \leqslant 1$, *as* $d \to \infty$

- $\mathcal{A}_k :=$ Connected graphs on $\{0, \ldots, k\}$ with $\mathfrak{a}(G) > 0$
- $k_0 := \min\{k \geqslant 0 \colon \mathcal{A}_k \neq \varnothing\}$

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,t} := \mathfrak{a}(\mathsf{GG}_d(t))$, where

- $\mathfrak{a}(\cdot)$ is a ***nonnegative*** functional on abstract graphs,
- $\mathfrak{a}(\cdot)$ is ***additive***: $\mathfrak{a}(G \sqcup G') = \mathfrak{a}(G) + \mathfrak{a}(G')$,
- $\mathfrak{a}(\cdot)$ ***grows at most exponentially fast***: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- $\lambda_d \to 0$ (sparse), that $|W_d|^{\frac{1}{d}} \to \infty$ (large window).

***Theorem (Moment asymptotics)***

- $\mathbb{E}[A_{d,t}] \sim \rho_d\, t^{k_0} \dfrac{\sum_{G\in\mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)}{(k_0+1)!}$, *for* $0 \leqslant t \leqslant 1$, *as* $d \to \infty$
- $\mathsf{Cov}[A_{d,t_1}, A_{d,t_2}] \sim \rho_d\, t_1^{k_0} \dfrac{\sum_{G\in\mathcal{A}^{\mathsf{m}}_{k_0}} \mathfrak{a}(G)^2}{((k_0+1)!)^2}$, *for* $0 \leqslant t_1 \leqslant t_2 \leqslant 1$, *as* $d \to \infty$

- $\mathcal{A}_k :=$ Connected graphs on $\{0, \dots, k\}$ with $\mathfrak{a}(G) > 0$
- $k_0 := \min\{k \geqslant 0 \colon \mathcal{A}_k \neq \varnothing\}$
- $\rho_d := |W_d| \lambda_d^{(k_0+1)d} \max_{G\in\mathcal{A}_{k_0}} v(G)^d$

where $v(G)^d$ is an “integral over indicator that $k_0 + 1$ points form graph at least $G$”.

- $(B_t)_{t \leqslant 1}$ = Standard Brownian motion
- Assume that $\mathfrak{a}$ is ***additive*** and ***nonnegative***, that $\lambda_d \to 0$, that $|W_d|^{\frac{1}{d}} \to \infty$, and $\mathfrak{a}(G) \in e^{O(|G|)}$.

- $(B_t)_{t\leqslant 1}$ = Standard Brownian motion
- Assume that $\mathfrak{a}$ is ***additive*** and ***nonnegative***, that $\lambda_d \to 0$, that $|W_d|^{\frac{1}{d}} \to \infty$, and $\mathfrak{a}(G) \in e^{O(|G|)}$.

**Theorem (multivariate and functional CLTs)**

*If $\mathbb{E}[A_{d,1}] \to \infty$, then for any $0 \leq t_1 \leq \cdots \leq t_m \leq 1$, as $d \to \infty$,*

$$\left( \frac{A_{d,t_1} - \mathbb{E}[A_{d,t_1}]}{\sqrt{\mathrm{Var}[A_{d,1}]}}, \ldots, \frac{A_{d,t_m} - \mathbb{E}[A_{d,t_m}]}{\sqrt{\mathrm{Var}[A_{d,1}]}} \right) \Rightarrow (B_{t_1^{k_0}}, \ldots, B_{t_m^{k_0}}).$$

Remark: The functional CLT's assumption can be relaxed to $\mathfrak{a} = \mathfrak{a}_+ + \mathfrak{a}_-$ with both $\mathfrak{a}_+$ and $\mathfrak{a}_-$ ***increasing*** and nonnegative with $\mathfrak{a}_\pm(G) \in e^{O(|G|)}$.

- $(B_t)_{t\leqslant 1}$ = Standard Brownian motion
- Assume that $\mathfrak{a}$ is ***additive*** and ***nonnegative***, that $\lambda_d \to 0$, that $|W_d|^{\frac{1}{d}} \to \infty$, and $\mathfrak{a}(G) \in e^{O(|G|)}$.

***Theorem (multivariate and functional CLTs)***

*If* $\mathbb{E}[A_{d,1}] \to \infty$, *then for any* $0 \le t_1 \le \cdots \le t_m \le 1$, *as* $d \to \infty$,

$$\left( \frac{A_{d,t_1} - \mathbb{E}[A_{d,t_1}]}{\sqrt{\mathrm{Var}[A_{d,1}]}}, \ldots, \frac{A_{d,t_m} - \mathbb{E}[A_{d,t_m}]}{\sqrt{\mathrm{Var}[A_{d,1}]}} \right) \Rightarrow (B_{t_1^{k_0}}, \ldots, B_{t_m^{k_0}}).$$

*If additionally* $\mathfrak{a}$ *is* ***increasing***, *then*

$$\left( \frac{A_{d,t} - \mathbb{E}[A_{d,t}]}{\sqrt{\mathrm{Var}[A_{d,1}]}} \right)_{t\le 1} \overset{Skorokhod}{\Longrightarrow} (B_{t^{k_0}})_{t\le 1}.$$

Remark: The functional CLT's assumption can be relaxed to $\mathfrak{a} = \mathfrak{a}_+ + \mathfrak{a}_-$ with both $\mathfrak{a}_+$ and $\mathfrak{a}_-$ ***increasing*** and nonnegative with $\mathfrak{a}_\pm(G) \in e^{O(|G|)}$.

Assume that $\mathfrak{a}$ is ***additive*** and ***nonnegative***, that $\lambda_d \to 0$, that $|W_d|^{\frac{1}{d}} \to \infty$, and $\mathfrak{a}(G) \in e^{O(|G|)}$.

Assume that $\mathfrak{a}$ is ***additive*** and ***nonnegative***, that $\lambda_d \to 0$, that $|W_d|^{\frac{1}{d}} \to \infty$, and $\mathfrak{a}(G) \in e^{O(|G|)}$.

**_Theorem (Poisson approximation)_**

*If $\mathbb{E}[A_{d,1}]$ **converges in** $(0,\infty)$ as $d \to \infty$, then*

$$(A_{d,t})_{t\leqslant 1} \overset{Skorokhod}{\Longrightarrow} \Big( \sum_{G\in\mathcal{A}^{\mathrm{m}}_{k_0}} N_t^{(G)} \mathfrak{a}(G) \Big)_{t\leqslant 1},$$

*where $(N_t^{(G)})_{t\leq 1}$ are independent Poisson processes.*

Above: $\mathcal{A}^{\mathrm{m}}_{k_0} = \{G \colon v(G) = \max_{G'\in\mathcal{A}_{k_0}} v(G')\}$

These Poisson processes $(N_t^{(G)})_{t\leq 1}$ are described by:

- Expected value $K((k_0+1)!)^{-1} t^{k_0}$ at time $t$
- $K := \lim_{d\to\infty} \rho_d$

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,\boldsymbol{t}} := \mathfrak{a}(\mathsf{GG}_d(\boldsymbol{t}))$, where

- $\boldsymbol{t} := (t_1, \dots, t_m)$ for $0 \le t_1 \le \dots \le t_m \le 1$
- $\mathsf{GG}_d(\boldsymbol{t}) := \big(\mathsf{GG}_d(t_1), \dots, \mathsf{GG}_d(t_m)\big)$
- $\mathfrak{a}(\cdot) = m$-variate nonnegative functional on abstract graphs

university of groningen

***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,\boldsymbol{t}} := \mathfrak{a}(\mathsf{GG}_d(\boldsymbol{t}))$, where

- $\boldsymbol{t} := (t_1, \dots, t_m)$ for $0 \le t_1 \le \cdots \le t_m \le 1$
- $\mathsf{GG}_d(\boldsymbol{t}) := \big(\mathsf{GG}_d(t_1), \dots, \mathsf{GG}_d(t_m)\big)$
- $\mathfrak{a}(\cdot) = m$-variate nonnegative functional on abstract graphs

Assume ***multi-additivity.***

$$\mathfrak{a}(\boldsymbol{R} \sqcup \boldsymbol{R}') = \mathfrak{a}(\boldsymbol{R}) + \mathfrak{a}(\boldsymbol{R}')$$

- where $\boldsymbol{R} = (G_1, \dots, G_m)$ with $G_1 \subseteq \cdots \subseteq G_m$ and $\boldsymbol{R}' = (G'_1, \dots, G'_m)$ with $G'_1 \subseteq \cdots \subseteq G'_m$ and $G_m \cap G'_m = \varnothing$.

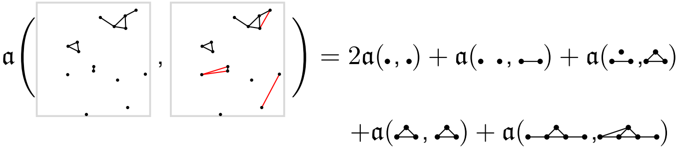
***Goal.*** Describe $d \uparrow \infty$-asymptotics of $A_{d,\boldsymbol{t}} := \mathfrak{a}(\mathsf{GG}_d(\boldsymbol{t}))$, where

- $\boldsymbol{t} := (t_1, \dots, t_m)$ for $0 \le t_1 \le \cdots \le t_m \le 1$
- $\mathsf{GG}_d(\boldsymbol{t}) := \big(\mathsf{GG}_d(t_1), \dots, \mathsf{GG}_d(t_m)\big)$
- $\mathfrak{a}(\cdot) = m$-variate nonnegative functional on abstract graphs

Assume ***multi-additivity.***

$$\mathfrak{a}(\boldsymbol{R} \sqcup \boldsymbol{R}') = \mathfrak{a}(\boldsymbol{R}) + \mathfrak{a}(\boldsymbol{R}')$$

- where $\boldsymbol{R} = (G_1, \dots, G_m)$ with $G_1 \subseteq \cdots \subseteq G_m$ and $\boldsymbol{R}' = (G'_1, \dots, G'_m)$ with $G'_1 \subseteq \cdots \subseteq G'_m$ and $G_m \cap G'_m = \varnothing$.



university of groningen

### *Linear combinations of univariate functionals*

$\mathfrak{a}(G_1, \dots, G_m) := \alpha_1 \mathfrak{a}'(G_1) + \cdots + \alpha_m \mathfrak{a}'(G_m)$, where $\alpha_1, \dots, \alpha_m \geqslant 0$ and where $\mathfrak{a}'$ is a nonnegative additive functional.

### *Linear combinations of univariate functionals*

$\mathfrak{a}(G_1, \dots, G_m) := \alpha_1 \mathfrak{a}'(G_1) + \cdots + \alpha_m \mathfrak{a}'(G_m)$, where $\alpha_1, \dots, \alpha_m \geqslant 0$ and where $\mathfrak{a}'$ is a nonnegative additive functional.

### *Dynamic subgraph count*

$(G_{0,1}, \dots, G_{0,m})$ fixed sequence of graphs and
$\mathfrak{a}(G_1, \dots, G_m) := \#\{G'_1 \subseteq \cdots \subseteq G'_m \colon G'_i \subseteq G_i$ and $G'_i \cong G_{0,i}$ for all $i \leqslant m\}$

### Linear combinations of univariate functionals

$\mathfrak{a}(G_1, \dots, G_m) := \alpha_1 \mathfrak{a}'(G_1) + \cdots + \alpha_m \mathfrak{a}'(G_m)$, where $\alpha_1, \dots, \alpha_m \geqslant 0$ and where $\mathfrak{a}'$ is a nonnegative additive functional.

### Dynamic subgraph count

$(G_{0,1}, \dots, G_{0,m})$ fixed sequence of graphs and

$\mathfrak{a}(G_1, \dots, G_m) := \#\{G'_1 \subseteq \cdots \subseteq G'_m \colon G'_i \subseteq G_i$ and $G'_i \cong G_{0,i}$ for all $i \leqslant m\}$

### Persistent Betti numbers

$\mathfrak{a}_i(G, G')$ as $i$th persistent Betti number associated with the clique complexes of the graphs $G \subseteq G'$.

university of groningen

### *Linear combinations of univariate functionals*

$\mathfrak{a}(G_1, \dots, G_m) := \alpha_1 \mathfrak{a}'(G_1) + \cdots + \alpha_m \mathfrak{a}'(G_m)$, where $\alpha_1, \dots, \alpha_m \geqslant 0$ and where $\mathfrak{a}'$ is a nonnegative additive functional.

### *Dynamic subgraph count*

$(G_{0,1}, \dots, G_{0,m})$ fixed sequence of graphs and
$\mathfrak{a}(G_1, \dots, G_m) := \#\{G'_1 \subseteq \cdots \subseteq G'_m : G'_i \subseteq G_i$ and $G'_i \cong G_{0,i}$ for all $i \leqslant m\}$

### *Persistent Betti numbers*

$\mathfrak{a}_i(G, G')$ as $i$th persistent Betti number associated with the clique complexes of the graphs $G \subseteq G'$.

**Theorem (CLT for multi-additive functionals)**

*Assume that $\mathfrak{a}$ is multi-additive, dominated and that $\mathfrak{a}(G, \cdots, G) \in e^{O(|G|)}$. Let $\boldsymbol{t} = (t_1, \dots, t_m)$ with $0 \leqslant t_1 \leqslant \cdots \leqslant t_m \leqslant 1$.*

1. *If $\rho_d^{1/d} \uparrow \infty$: $\mathbb{E}[A_{d,\boldsymbol{t}}] \asymp \rho_d$ and $\mathrm{Var}[A_{d,\boldsymbol{t}}] \asymp \rho_d$.*
2. *If $\rho_d^{1/d} \uparrow \infty$: $(\mathrm{Var}[A_{d,\boldsymbol{t}}])^{-1/2}(A_{d,\boldsymbol{t}} - \mathbb{E}[A_{d,\boldsymbol{t}}]) \Rightarrow \mathcal{N}(0,1)$*

# Proof of multivariate CLT

- ***Cramér-Wold theorem*** $\rightsquigarrow$ consider $\sum_{i\leqslant m} c_i A_{d,t_i}$ as $d \uparrow \infty$.

## Proof of multivariate CLT



- ***Cramér-Wold theorem*** $\rightsquigarrow$ consider $\sum_{i\leqslant m} c_i A_{d,t_i}$ as $d \uparrow \infty$.
- Using the moment (variance) asymptotics, it follows that we can ignore components with more than $k_0 + 1$ vertices.

## Proof of multivariate CLT

- ***Cramér-Wold theorem*** $\rightsquigarrow$ consider $\sum_{i\leqslant m} c_i A_{d,t_i}$ as $d \uparrow \infty$.
- Using the moment (variance) asymptotics, it follows that we can ignore components with more than $k_0 + 1$ vertices.
- After normalizing the centered process, we need to show that such an expression converges to a standard normal distribution. The variance asymptotics are handled by part 1 of the theorem.

## Proof of multivariate CLT

- ***Cramér-Wold theorem*** $\rightsquigarrow$ consider $\sum_{i\leqslant m} c_i A_{d,t_i}$ as $d \uparrow \infty$.
- Using the moment (variance) asymptotics, it follows that we can ignore components with more than $k_0 + 1$ vertices.
- After normalizing the centered process, we need to show that such an expression converges to a standard normal distribution. The variance asymptotics are handled by part 1 of the theorem.
- Divide $W_d$ in a grid consisting of boxes of side length about $2k_0$ and represent the functional of interest as a sum of random variables restricted to the boxes of the grid.

- ***Cramér-Wold theorem*** $\rightsquigarrow$ consider $\sum_{i \leqslant m} c_i A_{d,t_i}$ as $d \uparrow \infty$.
- Using the moment (variance) asymptotics, it follows that we can ignore components with more than $k_0 + 1$ vertices.
- After normalizing the centered process, we need to show that such an expression converges to a standard normal distribution. The variance asymptotics are handled by part 1 of the theorem.
- Divide $W_d$ in a grid consisting of boxes of side length about $2k_0$ and represent the functional of interest as a sum of random variables restricted to the boxes of the grid.
- Apply [Penrose, 2003, Theorem 2.4] ***(normal approximation for a sum of weakly dependent variables by Stein's method)***

By [Billingsley, 1999, Theorem 13.1] ***finite-dimensional convergence in distribution*** and ***tightness*** is sufficient for the ***convergence in distribution on a process level***.

By [Billingsley, 1999, Theorem 13.1] ***finite-dimensional convergence in distribution*** and ***tightness*** is sufficient for the ***convergence in distribution on a process level***.

To show tightness, we can show a ***condition for the 4th moment***. More, precisely that there exist $c, \varepsilon > 0$ such that

$$\rho_d^{-2}\mathbb{E}[\bar{A}_d(E)^4] \leqslant c|E|^{1+\varepsilon}$$

for all $d \geqslant 1$ and all intervals $E = [t_-, t_+] \subseteq [0, 1]$, where $\bar{A}_d(E)$ denotes the centered increment.

By [Billingsley, 1999, Theorem 13.1] ***finite-dimensional convergence in distribution*** and ***tightness*** is sufficient for the ***convergence in distribution on a process level***.

To show tightness, we can show a ***condition for the 4th moment***. More, precisely that there exist $c, \varepsilon > 0$ such that

$$\rho_d^{-2}\mathbb{E}[\bar{A}_d(E)^4] \leqslant c|E|^{1+\varepsilon}$$

for all $d \geqslant 1$ and all intervals $E = [t_-, t_+] \subseteq [0, 1]$, where $\bar{A}_d(E)$ denotes the centered increment.

***Strategy:***

- Reduce the moment condition to interval that are not too small.

By [Billingsley, 1999, Theorem 13.1] ***finite-dimensional convergence in distribution*** and ***tightness*** is sufficient for the ***convergence in distribution on a process level***.

To show tightness, we can show a ***condition for the 4th moment***. More, precisely that there exist $c, \varepsilon > 0$ such that

$$\rho_d^{-2}\mathbb{E}[\bar{A}_d(E)^4] \leqslant c|E|^{1+\varepsilon}$$

for all $d \geqslant 1$ and all intervals $E = [t_-, t_+] \subseteq [0,1]$, where $\bar{A}_d(E)$ denotes the centered increment.

***Strategy:***

- Reduce the moment condition to interval that are not too small.
- Use the ***cumulant identity***: Let $Z$ be a random variable with finite fourth moment. Then, $Z$ satisfies $\mathbb{E}[(Z - \mathbb{E}[Z])^4] = 3\mathrm{Var}(Z)^2 + c_4(Z)$. The fourth order cumulant is the $4$th derivative of the cumulant-generating function $s \mapsto \log \mathbb{E}[e^{sX}]$ evaluated at $0$.

By [Billingsley, 1999, Theorem 13.1] ***finite-dimensional convergence in distribution*** and ***tightness*** is sufficient for the ***convergence in distribution on a process level***.

To show tightness, we can show a ***condition for the 4th moment***. More, precisely that there exist $c, \varepsilon > 0$ such that

$$\rho_d^{-2}\mathbb{E}[\bar{A}_d(E)^4] \leqslant c|E|^{1+\varepsilon}$$

for all $d \geqslant 1$ and all intervals $E = [t_-, t_+] \subseteq [0,1]$, where $\bar{A}_d(E)$ denotes the centered increment.

***Strategy:***

- Reduce the moment condition to interval that are not too small.
- Use the ***cumulant identity***: Let $Z$ be a random variable with finite fourth moment. Then, $Z$ satisfies $\mathbb{E}[(Z - \mathbb{E}[Z])^4] = 3\mathsf{Var}(Z)^2 + c_4(Z)$. The fourth order cumulant is the $4$th derivative of the cumulant-generating function $s \mapsto \log \mathbb{E}[e^{sX}]$ evaluated at $0$.
- Derive bounds on the variance and the fourth-order cumulant $c_4$.

# Proof of Poisson Approximation

General approach similar to CLT: For high dimension $d$ only components in $\mathcal{A}'_{k_0}$ significantly contribute to the functional.

General approach similar to CLT: For high dimension $d$ only components in $\mathcal{A}'_{k_0}$ significantly contribute to the functional.

- Restrict $A_{d,t}$ to contributions from ***subgraphs in*** $\mathcal{A}^{\mathrm{m}}_{k_0}$, and that, once formed, ***do not change over time*** anymore. Call it $A'_{d,t,k_0}$. Then, $A'_{d,t,k_0}$ approximates $A_{d,t}$.

## Proof of Poisson Approximation

General approach similar to CLT: For high dimension $d$ only components in $\mathcal{A}'_{k_0}$ significantly contribute to the functional.

- Restrict $A_{d,t}$ to contributions from ***subgraphs in*** $\mathcal{A}^{\mathrm{m}}_{k_0}$, and that, once formed, ***do not change over time*** anymore. Call it $A'_{d,t,k_0}$. Then, $A'_{d,t,k_0}$ approximates $A_{d,t}$.
- With a general ***Poisson approximation result*** [Decreusefond et al., 2016, Theorem 3.1], we show that the sum in $A'_{d,t,k_0}$ over $(k_0+1)$-tuples of points can be approximated by a sum over a Poisson process. $\Rightarrow$ Approximate $A'_{d,t,k_0}$ by a Poisson functional.

General approach similar to CLT: For high dimension $d$ only components in $\mathcal{A}'_{k_0}$ significantly contribute to the functional.

- Restrict $A_{d,t}$ to contributions from ***subgraphs in $\mathcal{A}^{\mathrm{m}}_{k_0}$***, and that, once formed, ***do not change over time*** anymore. Call it $A'_{d,t,k_0}$. Then, $A'_{d,t,k_0}$ approximates $A_{d,t}$.
- With a general ***Poisson approximation result*** [Decreusefond et al., 2016, Theorem 3.1], we show that the sum in $A'_{d,t,k_0}$ over $(k_0 + 1)$-tuples of points can be approximated by a sum over a Poisson process. $\Rightarrow$ Approximate $A'_{d,t,k_0}$ by a Poisson functional.
- Categorizing the Poisson process into the graphs in $\mathcal{A}'_{k_0}$ that are formed yields ***independent*** Poisson Processes.

# THANK YOU!


Billingsley, P. (1999).
*Convergence of Probability Measures*.
John Wiley & Sons, New York, second edition.

Decreusefond, L., Schulte, M., and Thäle, C. (2016).
Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry.
*Ann. Probab.*, 44(3):2147–2197.

Penrose, M. (2003).
*Random geometric graphs*, volume 5 of *Oxford Studies in Probability*.
Oxford University Press, Oxford.