



university of
 groningen

faculty of science
 and engineering

High-dimensional sparse random geometric graphs

Gilles Bonnet

joint work with [Daniel Willhalm](#) (Groningen),

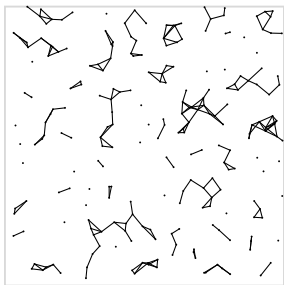
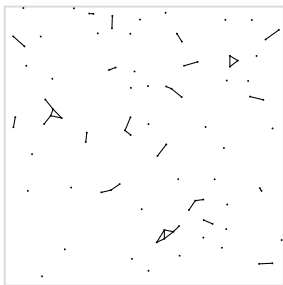
[Christian Hirsch](#) (Aarhus)

[Daniel Rosen](#) (Dortmund)

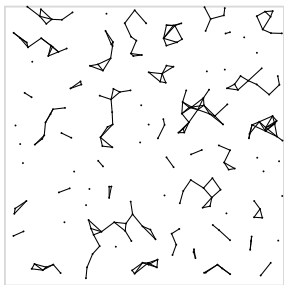
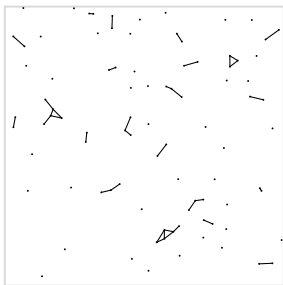
Limit theory of sparse random geometric graphs in high dimensions,
Stochastic Processes and their applications, vol 163, 203-236 (2023)

4th Italian Meeting on Prob. & Math. Stat., Rome

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



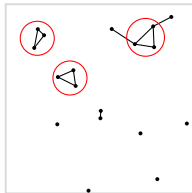
- ▷ $W_d = d$ -dimensional cubical observation window with opposite facets identified
= **torus**
- ▷ $\mathcal{P}_d :=$ Poisson process in W_d with **intensity** $\lambda_d^d > 0$
- ▷ $\text{GG}_d(t) = l_\infty$ -Gilbert graph with **connection radius** $t^{1/d}$, $0 \leq t \leq 1$



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

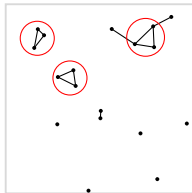
Subgraph count

- ▷ $G_0 =$ Fixed graph
- ▷ $\#\{G' \subseteq GG_d(t) : G' \cong G_0\}$
- ▷ $G_0 = \Delta \rightsquigarrow$ Triangle count



Subgraph count

- ▷ $G_0 =$ Fixed graph
- ▷ $\#\{G' \subseteq GG_d(t) : G' \cong G_0\}$
- ▷ $G_0 = \Delta \rightsquigarrow$ 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

1 Additive functionals

2 Multi-additive functionals

3 Proofs' strategies

4 The end

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,t} := \mathfrak{a}(\mathbb{G}\mathbb{G}_d(t))$, 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**: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- ▷ $\lambda_d \rightarrow 0$ (sparse), that $|W_d|^{\frac{1}{d}} \rightarrow \infty$ (large window).

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,t} := \mathfrak{a}(\text{GG}_d(t))$, 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**: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- ▷ $\lambda_d \rightarrow 0$ (sparse), that $|W_d|^{\frac{1}{d}} \rightarrow \infty$ (large window).

Theorem (Moment asymptotics)

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

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,t} := \mathfrak{a}(\mathbb{G}\mathbb{G}_d(t))$, 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**: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- ▷ $\lambda_d \rightarrow 0$ (sparse), that $|W_d|^{\frac{1}{d}} \rightarrow \infty$ (large window).

Theorem (Moment asymptotics)

$$\triangleright \mathbb{E}[A_{d,t}] \sim \rho_d t^{k_0} \frac{\sum_{G \in \mathcal{A}_{k_0}^m} \mathfrak{a}(G)}{(k_0+1)!}, \quad \text{for } 0 \leq t \leq 1, \text{ as } d \rightarrow \infty$$

$$\triangleright \text{Cov}[A_{d,t_1}, A_{d,t_2}] \sim \rho_d t_1^{k_0} \frac{\sum_{G \in \mathcal{A}_{k_0}^m} \mathfrak{a}(G)^2}{((k_0+1)!)^2}, \quad \text{for } 0 \leq t_1 \leq t_2 \leq 1, \text{ as } d \rightarrow \infty$$

- ▷ $\mathcal{A}_k :=$ Connected graphs on $\{0, \dots, k\}$ with $\mathfrak{a}(G) > 0$
- ▷ $k_0 := \min\{k \geq 0: \mathcal{A}_k \neq \emptyset\}$

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,t} := \mathfrak{a}(\text{GG}_d(t))$, 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**: $\mathfrak{a}(G) \in e^{O(|G|)}$,
- ▷ $\lambda_d \rightarrow 0$ (sparse), that $|W_d|^{\frac{1}{d}} \rightarrow \infty$ (large window).

Theorem (Moment asymptotics)

$$\text{▷ } \mathbb{E}[A_{d,t}] \sim \rho_d t^{k_0} \frac{\sum_{G \in \mathcal{A}_{k_0}^m} \mathfrak{a}(G)}{(k_0+1)!}, \quad \text{for } 0 \leq t \leq 1, \text{ as } d \rightarrow \infty$$

$$\text{▷ } \text{Cov}[A_{d,t_1}, A_{d,t_2}] \sim \rho_d t_1^{k_0} \frac{\sum_{G \in \mathcal{A}_{k_0}^m} \mathfrak{a}(G)^2}{((k_0+1)!)^2}, \quad \text{for } 0 \leq t_1 \leq t_2 \leq 1, \text{ as } d \rightarrow \infty$$

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

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

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

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

Theorem (multivariate and functional CLTs)

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

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

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

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

Theorem (multivariate and functional CLTs)

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

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

If additionally \mathbf{a} is **increasing**, then

$$\left(\frac{A_{d,t} - \mathbb{E}[A_{d,t}]}{\sqrt{\text{Var}[A_{d,1}]}} \right)_{t \leq 1} \xrightarrow{\text{Skorokhod}} (B_{t^{k_0}})_{t \leq 1}.$$

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

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

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

Theorem (Poisson approximation)

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

$$(A_{d,t})_{t \leq 1} \xrightarrow{\text{Skorokhod}} \left(\sum_{G \in \mathcal{A}_{k_0}^m} N_t^{(G)} \mathbf{a}(G) \right)_{t \leq 1},$$

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

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

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
- ▷ $K := \lim_{d \rightarrow \infty} \rho_d$

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,\mathbf{t}} := \alpha(\mathbb{G}\mathbb{G}_d(\mathbf{t}))$, where

- ▷ $\mathbf{t} := (t_1, \dots, t_m)$ for $0 \leq t_1 \leq \dots \leq t_m \leq 1$
- ▷ $\mathbb{G}\mathbb{G}_d(\mathbf{t}) := (\mathbb{G}\mathbb{G}_d(t_1), \dots, \mathbb{G}\mathbb{G}_d(t_m))$
- ▷ $\alpha(\cdot) = m$ -variate nonnegative functional on abstract graphs

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

- ▷ $\mathbf{t} := (t_1, \dots, t_m)$ for $0 \leq t_1 \leq \dots \leq t_m \leq 1$
- ▷ $\mathbb{G}\mathbb{G}_d(\mathbf{t}) := (\mathbb{G}\mathbb{G}_d(t_1), \dots, \mathbb{G}\mathbb{G}_d(t_m))$
- ▷ $\mathfrak{a}(\cdot) = m$ -variate nonnegative functional on abstract graphs

Assume **multi-additivity**.

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

- ▷ where $\mathbf{R} = (G_1, \dots, G_m)$ with $G_1 \subseteq \dots \subseteq G_m$ and $\mathbf{R}' = (G'_1, \dots, G'_m)$ with $G'_1 \subseteq \dots \subseteq G'_m$ and $G_m \cap G'_m = \emptyset$.

Goal. Describe $d \uparrow \infty$ -asymptotics of $A_{d,t} := \alpha(\mathbb{G}\mathbb{G}_d(\mathbf{t}))$, where

- ▷ $\mathbf{t} := (t_1, \dots, t_m)$ for $0 \leq t_1 \leq \dots \leq t_m \leq 1$
- ▷ $\mathbb{G}\mathbb{G}_d(\mathbf{t}) := (\mathbb{G}\mathbb{G}_d(t_1), \dots, \mathbb{G}\mathbb{G}_d(t_m))$
- ▷ $\alpha(\cdot) = m$ -variate nonnegative functional on abstract graphs

Assume **multi-additivity**.

$$\alpha(\mathbf{R} \sqcup \mathbf{R}') = \alpha(\mathbf{R}) + \alpha(\mathbf{R}')$$

- ▷ where $\mathbf{R} = (G_1, \dots, G_m)$ with $G_1 \subseteq \dots \subseteq G_m$ and $\mathbf{R}' = (G'_1, \dots, G'_m)$ with $G'_1 \subseteq \dots \subseteq G'_m$ and $G_m \cap G'_m = \emptyset$.

$$\alpha \left(\begin{array}{|c|} \hline \text{Diagram 1} \\ \hline \end{array}, \begin{array}{|c|} \hline \text{Diagram 2} \\ \hline \end{array} \right) = 2\alpha(\cdot, \cdot) + \alpha(\cdot, \cdot, \cdot) + \alpha(\cdot, \cdot, \cdot) \\ + \alpha(\cdot, \cdot, \cdot) + \alpha(\cdot, \cdot, \cdot)$$

The diagram shows two square boxes representing abstract graphs. The left box contains a small triangle and a larger triangle with a red edge. The right box contains the same small triangle, the larger triangle with a red edge, and two additional red edges connecting vertices. The equation shows that the functional α applied to the union of these two graphs is equal to the sum of α applied to each graph individually, which is then expanded into a sum of terms representing different graph configurations.

Linear combinations of univariate functionals

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

Linear combinations of univariate functionals

$\alpha(G_1, \dots, G_m) := \alpha_1 \alpha'(G_1) + \dots + \alpha_m \alpha'(G_m)$, where $\alpha_1, \dots, \alpha_m \geq 0$ and where α' is a nonnegative additive functional.

Dynamic subgraph count

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

$\alpha(G_1, \dots, G_m) := \#\{G'_1 \subseteq \dots \subseteq G'_m : G'_i \subseteq G_i \text{ and } G'_i \cong G_{0,i} \text{ for all } i \leq m\}$

Linear combinations of univariate functionals

$\alpha(G_1, \dots, G_m) := \alpha_1 \alpha'(G_1) + \dots + \alpha_m \alpha'(G_m)$, where $\alpha_1, \dots, \alpha_m \geq 0$ and where α' is a nonnegative additive functional.

Dynamic subgraph count

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

$\alpha(G_1, \dots, G_m) := \#\{G'_1 \subseteq \dots \subseteq G'_m : G'_i \subseteq G_i \text{ and } G'_i \cong G_{0,i} \text{ for all } i \leq m\}$

Persistent Betti numbers

$\alpha_i(G, G')$ as i th persistent Betti number associated with the clique complexes of the graphs $G \subseteq G'$.

Linear combinations of univariate functionals

$\alpha(G_1, \dots, G_m) := \alpha_1 \alpha'(G_1) + \dots + \alpha_m \alpha'(G_m)$, where $\alpha_1, \dots, \alpha_m \geq 0$ and where α' is a nonnegative additive functional.

Dynamic subgraph count

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

Persistent Betti numbers

$\alpha_i(G, G')$ 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 α is multi-additive, dominated and that $\alpha(G, \dots, G) \in e^{O(|G|)}$. Let $t = (t_1, \dots, t_m)$ with $0 \leq t_1 \leq \dots \leq t_m \leq 1$.

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

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

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

- ▷ **Cramér-Wold theorem** \rightsquigarrow consider $\sum_{i \leq m} c_i A_{d, t_i}$ 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.

- ▷ **Cramér-Wold theorem** \rightsquigarrow consider $\sum_{i \leq m} c_i A_{d,t_i}$ 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 $\sum_{i \leq m} c_i A_{d,t_i}$ 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

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

for all $d \geq 1$ 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

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

for all $d \geq 1$ 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

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

for all $d \geq 1$ 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 $\mathbb{E}[(Z - \mathbb{E}[Z])^4] = 3\text{Var}(Z)^2 + c_4(Z)$. The fourth order cumulant is the 4th 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

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

for all $d \geq 1$ 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 $\mathbb{E}[(Z - \mathbb{E}[Z])^4] = 3\text{Var}(Z)^2 + c_4(Z)$. The fourth order cumulant is the 4th 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 .

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* $\mathcal{A}_{k_0}^m$, and that, once formed, *do not change over time* anymore. Call it A'_{d,t,k_0} . Then, A'_{d,t,k_0} approximates $A_{d,t}$.

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 $\mathcal{A}_{k_0}^m$** , 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}'_{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.



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.