rijksuniversiteit
groningen

# Rare events in random geometric graphs

**Christian Hirsch** (University of Groningen)

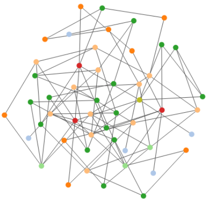
joint work with S. B. Moka, T. Taimre & D. P. Kroese

*Projet ANR GrHyDy - Kickoff meeting*, 20 Oct 2021

Spatial random networks

?

Rare events in random graphs



http://www.networkpages.nl

$G(n,p) =$ ER graph on $n$ nodes and prob $p$

$T(G(n,p)) = \#\Delta$ in $G(n,p)$

$$\mathbb{P}\Big(T(G(n,p)) \geqslant \binom{n}{3} t^3\Big) \rightsquigarrow ??$$

Bhamidi, Hanning, Lee & Nolen (2015) $\rightsquigarrow$ ***Importance sampling (IS)***

- ▷ Generate graph samples $G_1, \ldots, G_K$ under IS distribution $\mathbb{Q}$
- ▷ Compute $M_K := \frac{1}{K}\sum_{i \leqslant K} \mathbb{1}\{T(G_i) \geqslant \binom{n}{3} t^3\} \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}(G_i)$.

***How to choose $\mathbb{Q}$?***

- ▷ ***Edge tilt.*** $\mathbb{Q}^h(G) := c\exp(hE(G))$, $E(G) = \#$edges
- ▷ ***Triangle tilt.*** $\mathbb{Q}^{h,\beta,\alpha}(G) := c\exp\Big(hE(G) + \frac{\beta}{n}(n^3/6)^{1-\alpha} T(X)^\alpha\Big)$

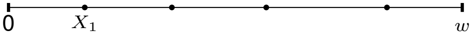
$\rightsquigarrow$ $\#$ ***Edges in random geometric graphs??***

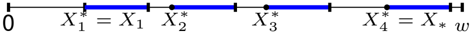
- $X = \{X_i\}_i =$ PPP on $[0, \infty)$ with intensity $\lambda$
- $G(X \cap W) =$ edges in RGG at level 1 in window $W = [0, w]$, $w \geqslant 1$

***Few edges.*** $\mathbb{P}(E_{\leqslant k}) := \mathbb{P}(|G(X \cap W)| \leqslant k) = ?$

$k = 0$.

- $X_1^* := X_1$; $X_{m+1}^* := \inf\{X_i \in X : X_i \geqslant X_m^* + 1\}$
- $X_* := X_{I_*}$; $I_* := \sup\{i \geqslant 1 : X_i^* \in W\}$




**Theorem**

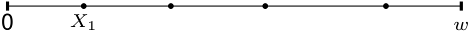
$\mathbb{P}(E_0 \,|\mathcal{F}^*) = e^{-\lambda((I_*-1)_+ + (w - X_*)\wedge 1)}$, *where* $\mathcal{F}^* := \sigma\big(X_1^*, X_2^*, \dots\big)$.

- $X = \{X_i\}_i =$ PPP on $[0, \infty)$ with intensity $\lambda$
- $G(X \cap W) =$ edges in RGG at level 1 in window $W = [0, w]$

***Few edges.*** $\mathbb{P}(E_{\leqslant k}) := \mathbb{P}(|G(X \cap W) \leqslant k|) =?$

$k = 0$.

- $X_1^* := X_1$; $X_{m+1}^* := \inf\{X_i \in X : X_i \geqslant X_m^* + 1\}$
- $X_* := X_{I_*}$; $I_* := \sup\{i \geqslant 1 : X_i^* \in W\}$



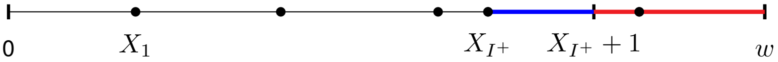


**Theorem**

$\mathbb{P}(E_0 \,|\mathcal{F}^*) = e^{-\lambda((I_*-1)_+ + (w - X_*)\wedge 1)}$, *where* $\mathcal{F}^* := \sigma\big(X_1^*, X_2^*, \dots\big)$.

$k = 1$.

- $I^+ := \inf\{i \geqslant 2 : X_i - X_{i-1} \leqslant 1\}$
- $X^+ := (X - X_{I^+}) \cap [1, \infty)$




**Theorem ($k = 1$)**

$$\mathbb{P}(E_{\leq 1} \,|\, \mathcal{F}^+) = e^{-\lambda((w - X_{I^+})_+ \wedge 1)} \mathbb{1}\{G(X^+ \cap [1, w - X_{I^+}]) = \emptyset\},$$

*where* $\mathcal{F}^+ := \sigma(X_1, X_2, \ldots, X_{I^+}, X^+)$

- $p_{\leqslant i} := \mathbb{P}(E_{\leqslant i})$
- $w \in \{5, 7.5, 10\}$, $\lambda = 2$, $N = 10^6$ simulations

- $p^{\mathsf{CMC}}_{\leq k} := \frac{1}{N} \sum_{i \leqslant N} \mathbb{1}\{|G(X^{((i)} \cap [0, w])| \leqslant k\}$

| $w$ | $p_0^{\mathsf{CMC}}$ | $p_{\leq 1}^{\mathsf{CMC}}$ |
|---|---|---|
| 5 | $4.056 \times 10^{-3} \pm 6.36 \times 10^{-5}$ | $1.676 \times 10^{-2} \pm 1.28 \times 10^{-4}$ |
| 7.5 | $2.410 \times 10^{-4} \pm 1.55 \times 10^{-5}$ | $1.354 \times 10^{-4} \pm 3.68 \times 10^{-5}$ |
| 10 | $1.100 \times 10^{-5} \pm 3.32 \times 10^{-6}$ | $8.500 \times 10^{-5} \pm 9.23 \times 10^{-6}$ |

- $p^{\mathsf{Cond}}_{\leq k} := \frac{1}{N} \sum_{i \leqslant N} P^{(i)}$

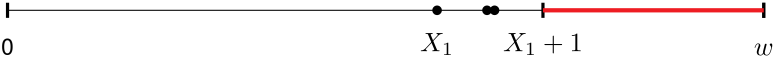
| $w$ | $p_0^{\mathsf{Cond}}$ | $p_{\leq 1}^{\mathsf{Cond}}$ |
|---|---|---|
| 5 | $4.148 \times 10^{-3} \pm 1.41 \times 10^{-5}$ (20.30) | $1.670 \times 10^{-2} \pm 7.63 \times 10^{-4}$ (2.8) |
| 7.5 | $2.244 \times 10^{-4} \pm 1.06 \times 10^{-6}$ (216.3) | $1.304 \times 10^{-4} \pm 1.87 \times 10^{-5}$ (3.9) |
| 10 | $1.296 \times 10^{-5} \pm 1.06 \times 10^{-7}$ (984.9) | $9.637 \times 10^{-5} \pm 4.56 \times 10^{-6}$ (4.1) |

***Few missing edges.***

▷ $M_w := \binom{|X\cap[0,w]|}{2} - |G(X\cap[0,w])|$= number of missing edges.

**Theorem (No missing edges)**

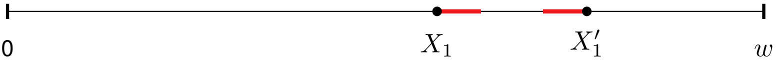
$\mathbb{P}(M_w = 0) = e^{-\lambda(w-1)} + (w-1)\lambda e^{-\lambda(w-1)}.$




**Theorem (At most one missing edge)**

$\mathbb{P}(M_w = 1) = \frac{\lambda^2(w-2)^2}{2}e^{-\lambda w} + (w-3/2)\lambda^2 e^{-\lambda(w-1)}$

▷ $X_1' :=$ first Poisson point to the left of $w$




**Case 1.** $X_1' \in [X_1+2, w] \rightsquigarrow X \cap (X_1, w] = \{X_1'\}$

**Case 2.** $X_1' \in [X_1+1, X_1+2] \rightsquigarrow X \cap ((X_1, X_1'-1] \cup [X_1+1, X_1')) = \emptyset$

$W$ = full-dimensional sampling window in $\mathbb{R}^d$

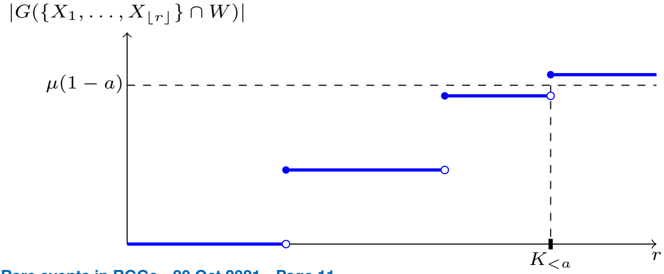
***Rare event.*** $F_{<a} := \{|G(X \cap W)| < (1-a)\mu\}$

***Idea.*** $X = \{X_n\}_{n \leqslant K}$, where

- $X_\infty := \{X_n\}_{n \geqslant 1}$ iid $X_i \sim \mathrm{Unif}(W)$
- $K \sim \mathrm{Poi}(\lambda|W|)$

$$\mathbb{P}(F_{<a}) = \mathbb{E}[\mathbb{P}(F_{<a} \mid X_\infty)] = \mathbb{E}\big[\mathrm{Poi}(K_{<a}(X_\infty) - 1)\big],$$

where $K_{<a}(X_\infty) := \inf\{k \geqslant 1 : |G(\{X_1, \ldots, X_k\} \cap W)| \geqslant (1-a)\mu\}$



- $W = [0, w]^2$, $w \in \{20, 25, 30\}$
- $\lambda = 2$
- $N = 10^5$ simulations

- $p^{\mathsf{Cond}}_{<0.2} := \frac{1}{N} \sum_{i \leqslant N} \mathsf{Poi}\big(K_{<a}(X_\infty(i)) - 1\big)$,

| | $p^{\mathsf{Cond}}_{<0.2}$ | $p^{\mathsf{Cond}}_{>0.2}$ |
|---|---|---|
| $X^{\mathsf{S}}$ | $2.02 \times 10^{-3} \pm 6.98 \times 10^{-6}$ $(414.8)$ | $5.12 \times 10^{-3} \pm 1.63 \times 10^{-5}$ $(193.7)$ |
| $X^{\mathsf{M}}$ | $1.54 \times 10^{-4} \pm 7.05 \times 10^{-7}$ $(3,106.4)$ | $6.76 \times 10^{-4} \pm 2.77 \times 10^{-6}$ $(878.6)$ |
| $X^{\mathsf{L}}$ | $6.91 \times 10^{-6} \pm 4.19 \times 10^{-8}$ $(39,415.8)$ | $6.24 \times 10^{-5} \pm 3.24 \times 10^{-7}$ $(5,911.1)$ |

Refined estimator by ***importance sampling*** of $X_\infty$?

***Lower tail***

***Idea.*** Find point-process law $\mathbb{Q}$ such that

- ▷ Rare event is typical under $\mathbb{Q}$ and $h(\mathbb{Q}|\text{Poi})$ is minimal
- ⇝ ***Edge-tilt/Strauss process***

***Sequential algorithm.***

- ▷ ***Initialize.*** $X = \{X_1, \dots, X_{n_0}\}$ iid points $n_0 \approx \lambda|W|$
- ▷ ***Iterate.*** Remove point $X_i$ from $X$ with prob. $\propto \gamma^{\mathsf{deg}(X_i)}$, $\gamma > 1$
- ▷ ***Likelihood ratio per step.***

$$\frac{|X|\gamma^{\mathsf{deg}(X_i)}}{\sum_j \gamma^{\mathsf{deg}(X_j)}}.$$

- ▷ ***Stop.*** If $|G(X \cap W)| \leqslant \mu(1-a)$

# Importance sampling

**Algorithm 4.1:** Importance sampling estimator $q^{\mathsf{IS}}_{<a}$ for $q_{<a}$

**input:** Number $N \geqslant 1$ of MC runs; parameter $\gamma$; Poisson intensity $\lambda$; sampling window $W$.

**output:** Importance sampling estimator $q^{\mathsf{IS}}_{<a}$ for $q_{<a}$.

$n_0 \leftarrow \lfloor \lambda |W| \rfloor$

sum $\leftarrow 0$

**for** $i \leqslant N$ **do**

  $X \leftarrow \{X_1, \ldots, X_{n_0}\}$ iid uniform in $W$

  $\rho \leftarrow 1$

  **while** $|G(X \cap W)| \geqslant \mu(1-a)$ **do**

    draw $X_i \in X$ with probability $r_i = \gamma^{\deg(X_i)} / \sum_{X_j \in X} \gamma^{\deg(X_j)}$

    $\rho \leftarrow \rho / (|X| r_i)$ $X \leftarrow X \setminus \{X_i\}$

  sum $\leftarrow$ sum $+ \rho\mathsf{Poi}(|X|)$

**return** sum$/N$

## *Upper tail*

### *Sequential algorithm.*

- ***Initialize.*** $X = \{X_1, \ldots, X_{n_0}\}$ iid points $n_0 \approx \lambda |W|$
- ***Discretize*** $W$ into bins $W_i$
- ***Iterate.*** Add point in $W_i$ prob. $\propto \gamma^{n(W_i)}$,
  - $n(W_i) = \#$ points in adjacent bins
- ***Likelihood ratio per step.***

$$\frac{\gamma^{n(W_i)} \cdot \#\text{bins}}{\sum_j \gamma^{n(W_j)}}.$$

- ***Stop.*** If $|G(X \cap W)| \geqslant \mu(1 - a)$

## Simulation study

- $W = [0, w]^2, w \in \{20, 25, 30\}$
- $\lambda = 2$
- $N = 10^5$ simulations

| | $q^{\mathsf{IS}}_{<0.2}$ | $q^{\mathsf{IS}}_{>0.2}$ |
|---|---|---|
| $X^{\mathsf{S}}$ | $2.02 \times 10^{-3} \pm 6.22 \times 10^{-6} (523.3)$ | $5.12 \times 10^{-3} \pm 1.57 \times 10^{-5} (207.9)$ |
| $X^{\mathsf{M}}$ | $1.54 \times 10^{-4} \pm 6.16 \times 10^{-7} (4,071.0)$ | $6.74 \times 10^{-4} \pm 2.66 \times 10^{-6} (951.8)$ |
| $X^{\mathsf{L}}$ | $6.93 \times 10^{-6} \pm 3.63 \times 10^{-8} (52,665.8)$ | $6.24 \times 10^{-5} \pm 3.09 \times 10^{-7} (6,537.22)$ |

# Summary and Outlook

- ▷ ***1D.*** Probabilities of few edges/few missing edges ✓
- ▷ ***2D.*** Conditional MC ✓
- ▷ ***2D.*** Sequential importance sampling ✓
- ▷ ***Rigorous results in 2D !***
- ▷ ***Optimality !***
- ▷ ***Triangles, edge length, . . . !***

# Thanks!

# Literature

[1] S. Chatterjee and M. Harel.
Localization in random geometric graphs with too many edges.
*Ann. Probab.*, 48(1):574–621, 2020.

[2] C. Hirsch, S. B. Moka, D. Taimre, and D. P. Kroese.
Rare events in random geometric graphs.
*Methodol. Comput. Appl. Probab.*, 2021+.