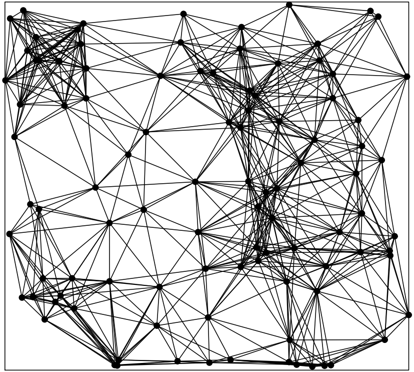
Random geometric graphs

Tobias Müller
Centrum Wiskunde & Informatica

YEP VIII, 18 March 2011

The random geometric graph (RGG)

We construct a random graph $G(n, r)$ as follows. We pick vertices $X_1, \ldots, X_n \in [0, 1]^2$ i.i.d. (independent, identically distributed) uniformly at random and we join X_i, X_j ($i \neq j$) by an edge if $\| X_i - X_j \| \leq r$.



Computer generated example with $n = 100, r = \frac{1}{4}$.

Disclaimer

- Often the RGG is defined in arbitrary dimension d, with the points $X_1, \ldots, X_n$ i.i.d. according to some (general) probability measure on $\mathbb{R}^d$, and where the distance between points is measured by an arbitrary norm $\|.\|$ on $\mathbb{R}^d$ (often the ℓ_p-norm for some p).

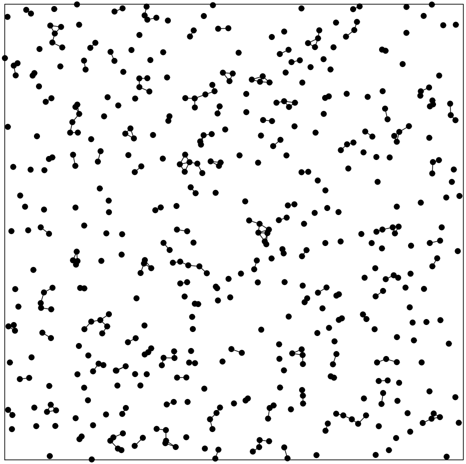
Disclaimer

- Often the RGG is defined in arbitrary dimension d, with the points $X_1, \ldots, X_n$ i.i.d. according to some (general) probability measure on $\mathbb{R}^d$, and where the distance between points is measured by an arbitrary norm $\|.\|$ on $\mathbb{R}^d$ (often the ℓ_p-norm for some p).
- To keep the presentation as light as possible I will state the results only when the dimension is 2, the points are i.i.d. uniform on the unit square and we use the the Euclidean norm to measure distance between points.

Disclaimer

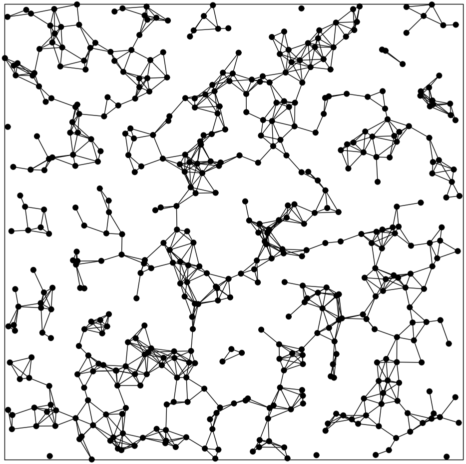
- Often the RGG is defined in arbitrary dimension d, with the points $X_1, \dots, X_n$ i.i.d. according to some (general) probability measure on $\mathbb{R}^d$, and where the distance between points is measured by an arbitrary norm $\|.\|$ on $\mathbb{R}^d$ (often the ℓ_p-norm for some p).
- To keep the presentation as light as possible I will state the results only when the dimension is 2, the points are i.i.d. uniform on the unit square and we use the the Euclidean norm to measure distance between points.
- Feel free to ask me about generalizations.

More pictures: expected degree 1



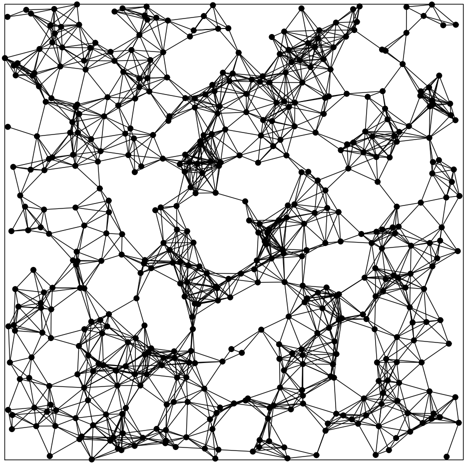
Computer generated example with $n = 500$ and r such that $\pi n r^2 = 1$. (Note that $\pi n r^2$ is roughly the expected degree.)

More pictures: expected degree 5



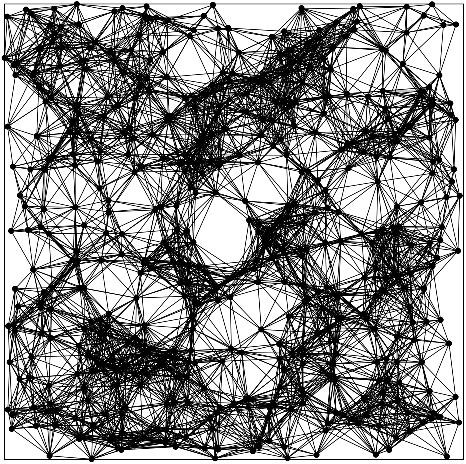
Computer generated example with $n = 500$ and r such that $\pi n r^2 = 5$. (Note that $\pi n r^2$ is roughly the expected degree.)

More pictures: expected degree 10



Computer generated example with $n = 500$ and r such that $\pi n r^2 = 10$. (Note that $\pi n r^2$ is roughly the expected degree.)

More pictures: expected degree 25



Computer generated example with $n = 500$ and r such that $\pi n r^2 = 25$. (Note that $\pi n r^2$ is roughly the expected degree.)

Connectedness of the RGG

Theorem.[Penrose 1997] Let $(r_n)_n$ be a sequence of nonnegative numbers, and write $x_n := \pi n r_n^2 - \ln n$. Then:

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, r_n) \text{ is connected }\right] = \begin{cases} 0 & \text{if } x_n \to -\infty; \\ e^{-e^{-x}} & \text{if } x_n \to x \in \mathbb{R}; \\ 1 & \text{if } x_n \to +\infty. \end{cases}$$

Recall that $\pi n r_n^2$ is (roughly) the average/expected degree.

Some notation

For a graph G, we let $L(G)$ denote the number of vertices of the largest component.

Some notation

For a graph G, we let $L(G)$ denote the number of vertices of the largest component.

If E is an event then E holds *almost surely (a.s.)* means that $\mathbb{P}(E) = 1$.

Some notation

For a graph G, we let $L(G)$ denote the number of vertices of the largest component.

If E is an event then E holds *almost surely (a.s.)* means that $\mathbb{P}(E) = 1$.

If $Z_1, Z_2, \dots$ are random variables and $c \in \mathbb{R}$ a constant then we say that "Z_n converges to c almost surely", denoted:

$$Z_n \to c \quad \text{a.s.}$$

if $\mathbb{P}(Z_n \to c) = 1$.

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.
Let the sequence $(r_n)_n$ be defined by $r_n := \sqrt{\lambda/\pi n}$ with $\lambda \geq 0$ fixed (observe that $\pi n r_n^2 = \lambda$).

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.
Let the sequence $(r_n)_n$ be defined by $r_n := \sqrt{\lambda/\pi n}$ with $\lambda \geq 0$ fixed (observe that $\pi n r_n^2 = \lambda$).
Then

$$\frac{L(G(n, r_n))}{n} \to f(\lambda) \text{ a.s.}$$

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.
Let the sequence $(r_n)_n$ be defined by $r_n := \sqrt{\lambda/\pi n}$ with $\lambda \geq 0$ fixed (observe that $\pi n r_n^2 = \lambda$).
Then

$$\frac{L(G(n, r_n))}{n} \to f(\lambda) \text{ a.s.}$$

Furthermore, there is a constant $\lambda_{\text{crit}} > 0$ such that $f(\lambda) = 0$ for $\lambda \leq \lambda_{\text{crit}}$, and $f(\lambda) > 0$ for $\lambda > \lambda_{\text{crit}}$.

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.
Let the sequence $(r_n)_n$ be defined by $r_n := \sqrt{\lambda/\pi n}$ with $\lambda \geq 0$ fixed (observe that $\pi n r_n^2 = \lambda$).
Then

$$\frac{L(G(n, r_n))}{n} \to f(\lambda) \text{ a.s.}$$

Furthermore, there is a constant $\lambda_{\text{crit}} > 0$ such that $f(\lambda) = 0$ for $\lambda \leq \lambda_{\text{crit}}$, and $f(\lambda) > 0$ for $\lambda > \lambda_{\text{crit}}$.

A similar, but slightly weaker result was already obtained by E. N. Gilbert in 1961.

The largest component of the RGG

Theorem[Penrose 2003] There exists a non-decreasing continuous function $f : [0, \infty) \to [0, 1)$ such that the following holds.
Let the sequence $(r_n)_n$ be defined by $r_n := \sqrt{\lambda/\pi n}$ with $\lambda \geq 0$ fixed (observe that $\pi n r_n^2 = \lambda$).
Then

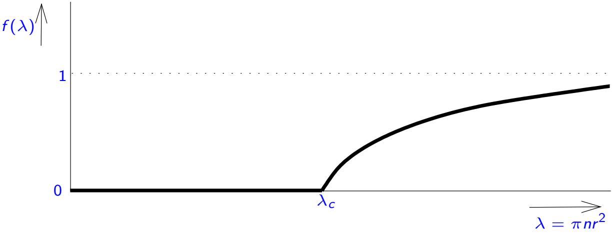
$$\frac{L(G(n, r_n))}{n} \to f(\lambda) \text{ a.s.}$$

Furthermore, there is a constant $\lambda_{\text{crit}} > 0$ such that $f(\lambda) = 0$ for $\lambda \leq \lambda_{\text{crit}}$, and $f(\lambda) > 0$ for $\lambda > \lambda_{\text{crit}}$.

A similar, but slightly weaker result was already obtained by E. N. Gilbert in 1961.

The precise values of the λ_{crit} and $f(\lambda)$ for $\lambda > \lambda_{\text{crit}}$ are unknown, but experimentally $\lambda_{\text{crit}} \approx 4.51$.

A cartoon plot of $f(\lambda)$



Other aspects of the RGG that have been considered. (A selection)

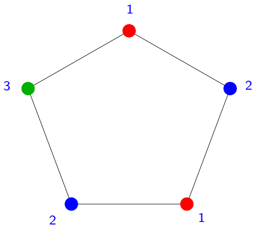
- Cover and mixing times of a random walk on the graph. [Avin+Ercal 07, Cooper+Frieze '09];
- Eigenvalues of the adjacency matrix [BEJ 06, Rai 09];
- Monotone properties [McColm 04, GRK 05];
- First order expressible properties [McColm 99, Agarwal+Spencer 05];
- Min and max bisection [DPPS 99, DGM 06];
- Graph diameter [CKE 05];
- Small components [Penrose 03, DPM 08];
- Broadcasting algorithms on the graph [BEFSS 09];
- Chromatic number [McDiarmid 03, Penrose 03, DSS 07, MM 07+, Müller 08];
- Hamilton cycles [Petit 01,DPM 07,BBMW 09+,MPW 09+].

Part II: Colouring.

chromatic number

Let $G = (V, E)$ be a graph.

A k-colouring of G is a map $f : V \to \{1, \ldots, k\}$ that satisfies $f(v) \neq f(w)$ whenever $vw \in E$



The *chromatic number* $\chi(G)$ is the least k such that G is k-colourable.

Why colour (random geometric) graphs?

- Graph colouring goes back to the famous four-colour conjecture by Francis Guthrie in 1852 (now a theorem by Appel+Haken 1976);

Why colour (random geometric) graphs?

- Graph colouring goes back to the famous four-colour conjecture by Francis Guthrie in 1852 (now a theorem by Appel+Haken 1976);
- Many discrete optimization problems, such as scheduling, are ultimately graph colouring problems;

Why colour (random geometric) graphs?

- Graph colouring goes back to the famous four-colour conjecture by Francis Guthrie in 1852 (now a theorem by Appel+Haken 1976);
- Many discrete optimization problems, such as scheduling, are ultimately graph colouring problems;
- Frequency assignment: A random geometric graph might model a network of radio transmitters.
  - Each transmitter needs to transmit its signal on some frequency;
  - But, if two transmitters use the same frequency and they are (too) close, then there is interference between the signals;
  - We want to minimise the number of distinct frequencies used and keep the interference at an acceptable level.

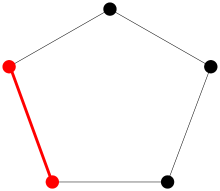
Why colour (random geometric) graphs?

- Graph colouring goes back to the famous four-colour conjecture by Francis Guthrie in 1852 (now a theorem by Appel+Haken 1976);
- Many discrete optimization problems, such as scheduling, are ultimately graph colouring problems;
- Frequency assignment: A random geometric graph might model a network of radio transmitters.
  - Each transmitter needs to transmit its signal on some frequency;
  - But, if two transmitters use the same frequency and they are (too) close, then there is interference between the signals;
  - We want to minimise the number of distinct frequencies used and keep the interference at an acceptable level.
- Great fun.

clique number

Let $G = (V, E)$ be a graph.

A *clique* in G is a complete subgraph of G, ie. a set of vertices $C \subseteq V$ such that $vw \in E$ for all $v, w \in C$.

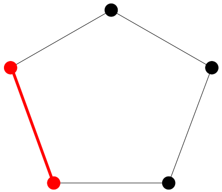


The *clique number* $\omega(G)$ is the cardinality of a largest clique.

clique number

Let $G = (V, E)$ be a graph.

A *clique* in G is a complete subgraph of G, ie. a set of vertices $C \subseteq V$ such that $vw \in E$ for all $v, w \in C$.



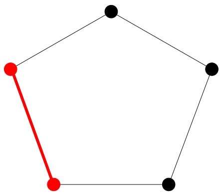
The *clique number* $\omega(G)$ is the cardinality of a largest clique.

Observe that $\chi(G) \geq \omega(G)$ for all G.

clique number

Let $G = (V, E)$ be a graph.

A *clique* in G is a complete subgraph of G, ie. a set of vertices $C \subseteq V$ such that $vw \in E$ for all $v, w \in C$.



The *clique number* $\omega(G)$ is the cardinality of a largest clique.

Observe that $\chi(G) \geq \omega(G)$ for all G.

In general, the ratio $\chi(G)/\omega(G)$ can be arbitrarily large [Mycielski 1955].

A result on the ratio χ/ω

Theorem.[McDiarmid+M 2007+] There exists a $t_0 > 0$ and a continuous, strictly increasing function $x : [t_0, \infty) \to [1, 2\sqrt{3}/\pi)$ such that, for any sequence $(r_n)_n$:

(i) If $\pi n r_n^2 \leq t_0 \ln n$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to 1 \quad \text{a.s.};$$

A result on the ratio χ/ω

Theorem.[McDiarmid+M 2007+] There exists a $t_0 > 0$ and a continuous, strictly increasing function $x : [t_0, \infty) \to [1, 2\sqrt{3}/\pi)$ such that, for any sequence $(r_n)_n$:

(i) If $\pi n r_n^2 \leq t_0 \ln n$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to 1 \quad \text{a.s.};$$

(ii) If $\frac{\pi n r^2}{\ln n} \to t$ with $t_0 \leq t < \infty$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to x(t) \quad \text{a.s.};$$

A result on the ratio χ/ω

Theorem.[McDiarmid+M 2007+] There exists a $t_0 > 0$ and a continuous, strictly increasing function $x : [t_0, \infty) \to [1, 2\sqrt{3}/\pi)$ such that, for any sequence $(r_n)_n$:

(i) If $\pi n r_n^2 \leq t_0 \ln n$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to 1 \quad \text{a.s.};$$

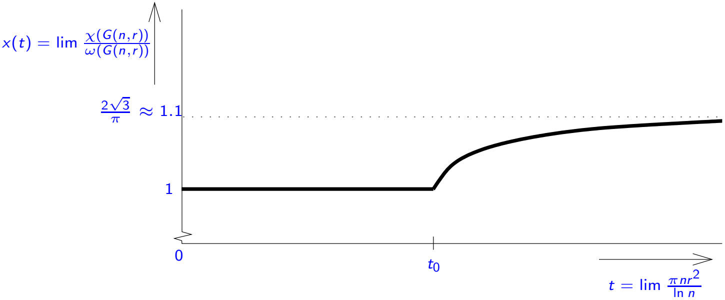
(ii) If $\frac{\pi n r^2}{\ln n} \to t$ with $t_0 \leq t < \infty$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to x(t) \quad \text{a.s.};$$

(iii) If $\frac{\pi n r^2}{\ln n} \to \infty$ then

$$\chi(G(n, r_n))/\omega(G(n, r_n)) \to \frac{2\sqrt{3}}{\pi} \quad \text{a.s.}$$

Another cartoon



Note : This is very different from other graph models

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;
- The precise value of t_0 is not known;

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;
- The precise value of t_0 is not known;
- While we have an 'explicit' expression for $x(t)$, it is not easy to extract information it. It is not even clear how to numerically approximate $x(t)$;

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;
- The precise value of t_0 is not known;
- While we have an 'explicit' expression for $x(t)$, it is not easy to extract information it. It is not even clear how to numerically approximate $x(t)$;
- However, we know $x(t_0) = 1$ and $\lim_{t\to\infty} x(t) = \frac{2\sqrt{3}}{\pi}$.

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;
- The precise value of t_0 is not known;
- While we have an 'explicit' expression for $x(t)$, it is not easy to extract information it. It is not even clear how to numerically approximate $x(t)$;
- However, we know $x(t_0) = 1$ and $\lim_{t\to\infty} x(t) = \frac{2\sqrt{3}}{\pi}$.
- It takes several (technical) slides to state the definition of x so I'll skip it today.

Remarks

- The behaviour of $\omega(G(n, r_n))$ and $\chi(G(n, r_n))$ is also described separately in the paper but we skip this today;
- The precise value of t_0 is not known;
- While we have an 'explicit' expression for $x(t)$, it is not easy to extract information it. It is not even clear how to numerically approximate $x(t)$;
- However, we know $x(t_0) = 1$ and $\lim_{t\to\infty} x(t) = \frac{2\sqrt{3}}{\pi}$.
- It takes several (technical) slides to state the definition of x so I'll skip it today.
- Where does the constant $2\sqrt{3}/\pi$ come from? I will try to explain in the next few slides.

The clique number for large(ish) expected degree.

Theorem.[McDiarmid 2003] If $\pi n r_n^2 \gg \ln n$ then $\omega(G(n, r_n))/nr_n^2 \to \frac{\pi}{4}$ a.s.

Proof sketch:

- A clique is a set of points of diameter $\leq r_n$;

The clique number for large(ish) expected degree.

Theorem.[McDiarmid 2003] If $\pi n r_n^2 \gg \ln n$ then $\omega(G(n, r_n))/nr_n^2 \to \frac{\pi}{4}$ a.s.

Proof sketch:

- A clique is a set of points of diameter $\leq r_n$;
- Fix a disk of radius $r_n/2$ inside the unit square. The clique number is at least the number of points that fall inside this disk. We expect $\frac{\pi}{4} n r_n^2$ points to fall inside it.

The clique number for large(ish) expected degree.

Theorem.[McDiarmid 2003] If $\pi n r_n^2 \gg \ln n$ then $\omega(G(n, r_n))/nr_n^2 \to \frac{\pi}{4}$ a.s.

Proof sketch:

- A clique is a set of points of diameter $\leq r_n$;
- Fix a disk of radius $r_n/2$ inside the unit square. The clique number is at least the number of points that fall inside this disk. We expect $\frac{\pi}{4} n r_n^2$ points to fall inside it. When $\pi n r_n^2 \gg \ln n$ then it contains at least $(1-\varepsilon)\frac{\pi}{4} n r_n^2$ points with probability tending to 1;

The clique number for large(ish) expected degree.

Theorem.[McDiarmid 2003] If $\pi n r_n^2 \gg \ln n$ then $\omega(G(n, r_n))/nr_n^2 \to \frac{\pi}{4}$ a.s.

Proof sketch:

- A clique is a set of points of diameter $\leq r_n$;
- Fix a disk of radius $r_n/2$ inside the unit square. The clique number is at least the number of points that fall inside this disk. We expect $\frac{\pi}{4} n r_n^2$ points to fall inside it. When $\pi n r_n^2 \gg \ln n$ then it contains at least $(1-\varepsilon)\frac{\pi}{4} n r_n^2$ points with probability tending to 1;
- Isodiametric inequality: every set of diameter less than r has area at most $\frac{\pi}{4} r^2$;

The clique number for large(ish) expected degree.

Theorem.[McDiarmid 2003] If $\pi n r_n^2 \gg \ln n$ then $\omega(G(n, r_n))/nr_n^2 \to \frac{\pi}{4}$ a.s.

Proof sketch:

- A clique is a set of points of diameter $\leq r_n$;
- Fix a disk of radius $r_n/2$ inside the unit square. The clique number is at least the number of points that fall inside this disk. We expect $\frac{\pi}{4} n r_n^2$ points to fall inside it. When $\pi n r_n^2 \gg \ln n$ then it contains at least $(1-\varepsilon)\frac{\pi}{4} n r_n^2$ points with probability tending to 1;
- Isodiametric inequality: every set of diameter less than r has area at most $\frac{\pi}{4} r^2$;
- When $nr_n^2 \gg \ln n$ there is a "concentration phenomenon": (with probability tending to 1) every convex set $S \subseteq \mathbb{R}^2$ with diameter $\leq r_n$ contains less than $(1+\varepsilon)\frac{\pi}{4} n r^2$ points.

■

The packing density

For $K > 0$, let $N(K)$ denote the biggest number of points in $[0, K]^2$ that all have pairwise distance 2. (in other words they are the centers of disjoint disks of unit radius). The limit

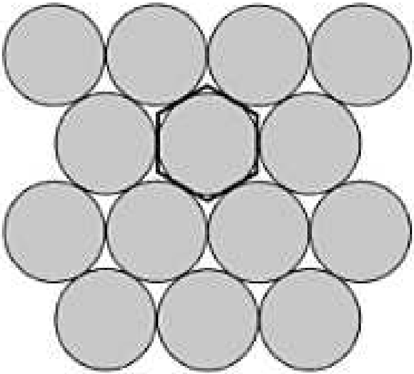
$$\delta := \lim_{K \to \infty} \frac{\pi \cdot N(K)}{K^2}.$$

exists and it equals

$$\delta = \frac{\pi}{2\sqrt{3}},$$

by a theorem of Thue from 1892. The constant δ is the *packing density* of the unit disk and it can be interpreted as the biggest proportion of the plane that can be filled with disjoint unit disks.

(Part of) an optimal packing



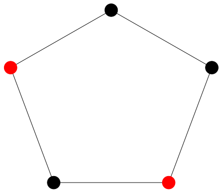
We can cover a proportion of $\delta = \pi/2\sqrt{3}$ of the plane with disks centered on the "hexagonal" lattice.

The circumscribed regular hexagon around a disk of radius 1 has area $2\sqrt{3}$.

independence number

Let $G = (V, E)$ be a graph.

A *independent set* in G is a set of vertices $C \subseteq V$ such that $vw \notin E$ for all $v, w \in C$.

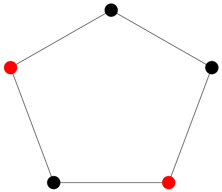


The *independence number* $\alpha(G)$ is the cardinality of a largest independent set.

independence number

Let $G = (V, E)$ be a graph.

A *independent set* in G is a set of vertices $C \subseteq V$ such that $vw \notin E$ for all $v, w \in C$.



The *independence number* $\alpha(G)$ is the cardinality of a largest independent set.

Observe that $\chi(G) \geq |V|/\alpha(G)$ for all G.
(A colouring is a partition into independent sets.)

Why is δ relevant for colouring the RGG – some intuition

Suppose that $X_{i_1}, \ldots, X_{i_m}$ form an independent set.

Why is δ relevant for colouring the RGG – some intuition

Suppose that $X_{i_1}, \dots, X_{i_m}$ form an independent set.

Then $\| X_{i_j} - X_{i_{j'}} \| \geq r$ for all $1 \leq j < j' \leq m$.

Why is δ relevant for colouring the RGG – some intuition

Suppose that $X_{i_1}, \ldots, X_{i_m}$ form an independent set.

Then $\| X_{i_j} - X_{i_{j'}} \| \geq r$ for all $1 \leq j < j' \leq m$.

In other words, the disks $B(X_{i_j}, r/2)$ are disjoint.

Why is δ relevant for colouring the RGG – some intuition

Suppose that $X_{i_1}, \dots, X_{i_m}$ form an independent set.

Then $\| X_{i_j} - X_{i_{j'}} \| \geq r$ for all $1 \leq j < j' \leq m$.

In other words, the disks $B(X_{i_j}, r/2)$ are disjoint.

Since the X_i lie in the unit square, a deterministic bound for the independence number is:

$$\alpha(G(n,r)) \leq N(2/r) \approx \frac{\delta(2/r)^2}{\pi} = \frac{2}{\sqrt{3}} r^{-2},$$

(for small r) by definition of δ.

Why is δ relevant for colouring the RGG – some intuition

Suppose that $X_{i_1}, \dots, X_{i_m}$ form an independent set.

Then $\| X_{i_j} - X_{i_{j'}} \| \geq r$ for all $1 \leq j < j' \leq m$.

In other words, the disks $B(X_{i_j}, r/2)$ are disjoint.

Since the X_i lie in the unit square, a deterministic bound for the independence number is:

$$\alpha(G(n,r)) \leq N(2/r) \approx \frac{\delta(2/r)^2}{\pi} = \frac{2}{\sqrt{3}} r^{-2},$$

(for small r) by definition of δ.

As it turns out, when $nr^2 \gg \ln n$, the deterministic lower bound

$$\chi(G(n,r)) \geq \frac{n}{\alpha(G(n,r))} \geq \frac{n}{N(2/r)} \approx \frac{\sqrt{3}}{2} nr^2,$$

gives (roughly) the right answer.

How we ended up with $2\sqrt{3}/\pi$.

When $\pi n r_n^2 \gg \ln n$ then

$$\frac{\omega(G(n, r_n))}{n r_n^2} \to \frac{\pi}{4} \text{ a.s.}$$

and (although I only sketched the easy half of the proof):

$$\frac{\chi(G(n, r_n))}{n r_n^2} \to \frac{\sqrt{3}}{2} \text{ a.s.}$$

And hence

$$\frac{\chi(G(n, r_n))}{\omega(G(n, r_n))} \to \frac{2\sqrt{3}}{\pi} \text{ a.s.}$$

■

The probability distribution: two-point concentration

When the "expected degree" $\pi n r_n^2$ is not too large then the clique and chromatic numbers are 'quasi-deterministic':

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(k_n)_n$ such that:

$$\mathbb{P}\left[\omega(G(n, r_n)) \in \{k_n, k_n + 1\}\right] \to 1.$$

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(m_n)_n$ such that:

$$\mathbb{P}\left[\chi(G(n, r_n)) \in \{m_n, m_n + 1\}\right] \to 1.$$

The probability distribution: two-point concentration

When the "expected degree" $\pi n r_n^2$ is not too large then the clique and chromatic numbers are 'quasi-deterministic':

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(k_n)_n$ such that:

$$\mathbb{P}\left[\omega(G(n, r_n)) \in \{k_n, k_n + 1\}\right] \to 1.$$

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(m_n)_n$ such that:

$$\mathbb{P}\left[\chi(G(n, r_n)) \in \{m_n, m_n + 1\}\right] \to 1.$$

This proves and extends a conjecture of Penrose.

The probability distribution: two-point concentration

When the "expected degree" $\pi n r_n^2$ is not too large then the clique and chromatic numbers are 'quasi-deterministic':

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(k_n)_n$ such that:

$$\mathbb{P}\left[\omega(G(n, r_n)) \in \{k_n, k_n + 1\}\right] \to 1.$$

Theorem.[M, 2008] If $\frac{\pi n r_n^2}{\ln n} \to 0$ then there is a sequence $(m_n)_n$ such that:

$$\mathbb{P}\left[\chi(G(n, r_n)) \in \{m_n, m_n + 1\}\right] \to 1.$$

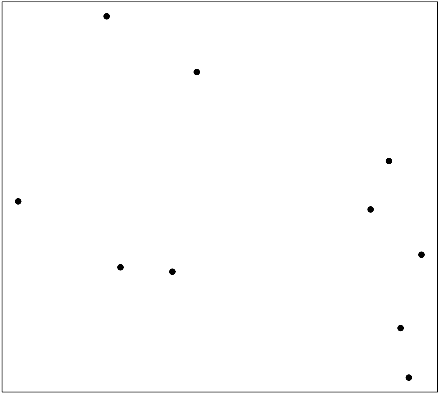
This proves and extends a conjecture of Penrose.

For other choices of $(r_n)_n$ the probability distribution of χ, ω is an open problem.

Part III: Hamilton cycles

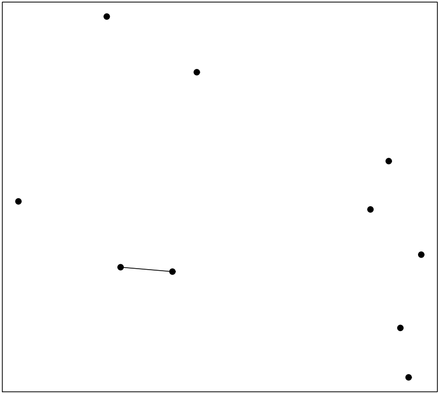
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



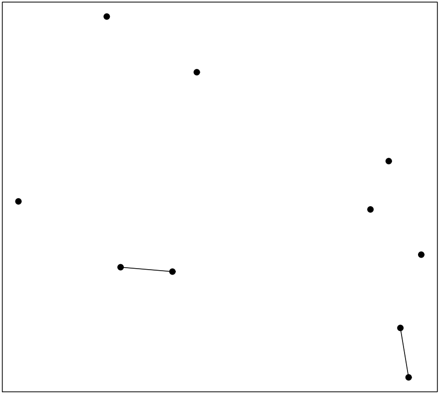
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



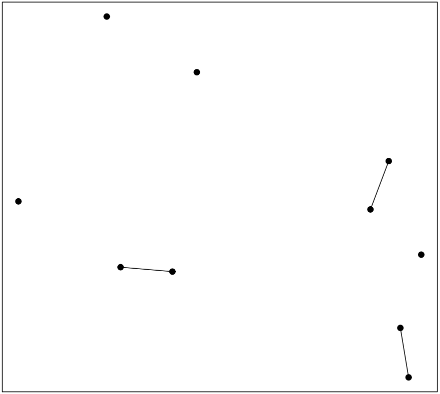
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



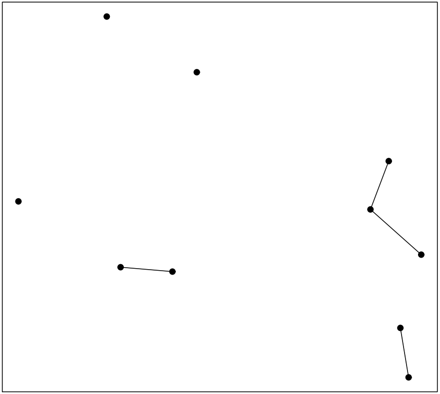
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



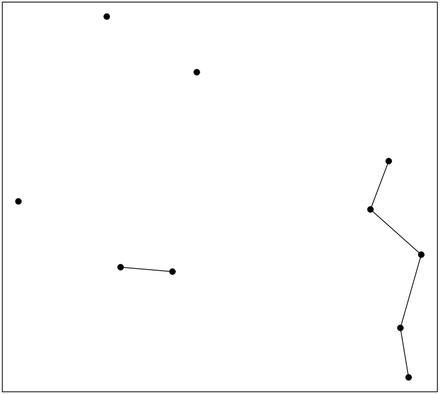
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



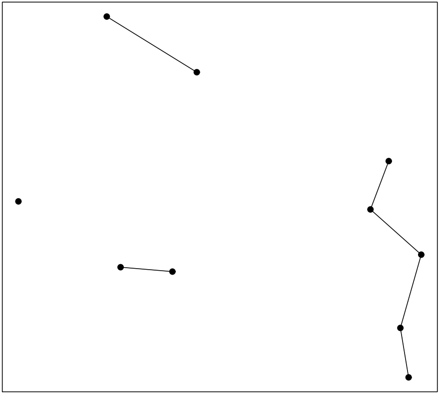
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



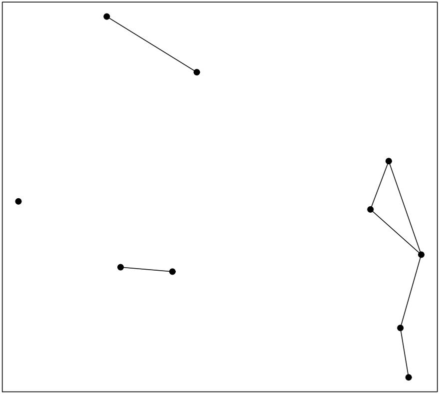
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



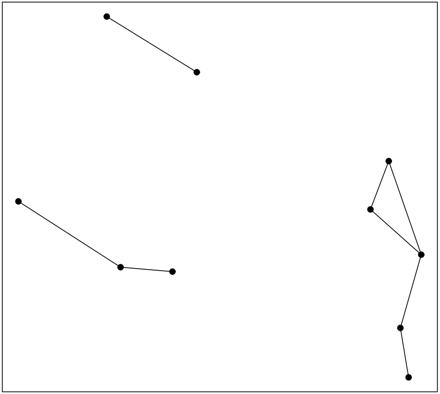
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



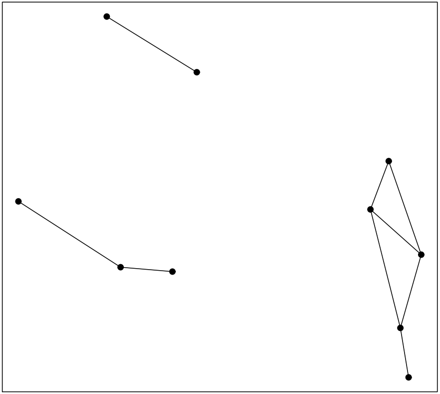
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



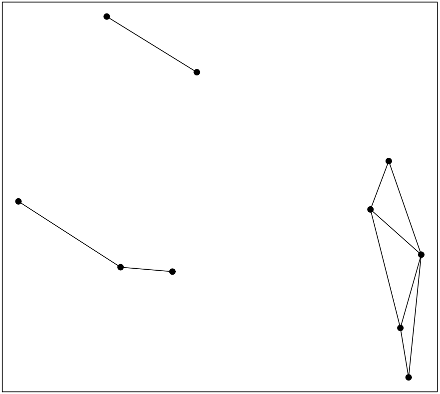
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



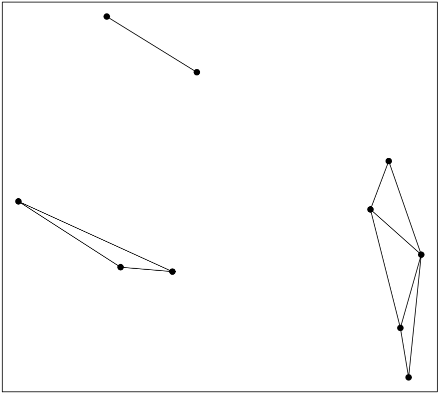
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



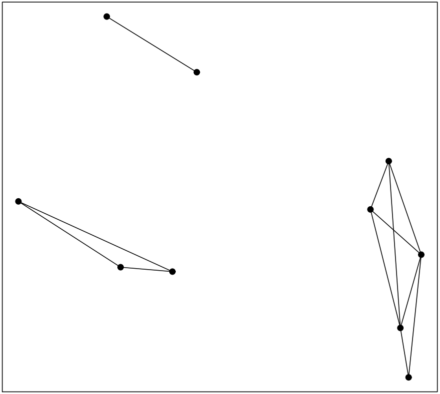
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



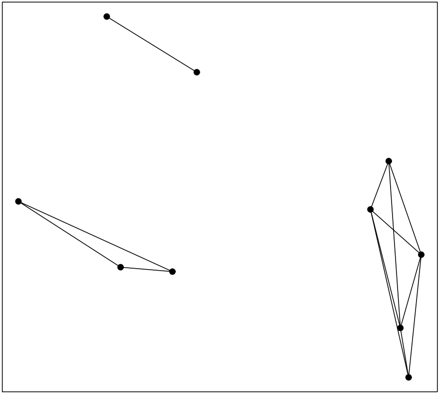
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



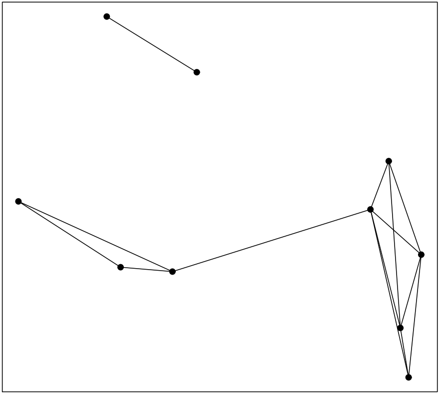
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



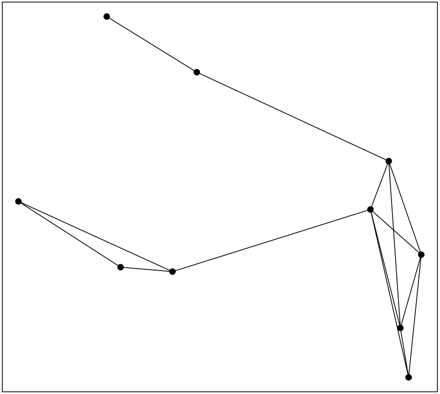
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



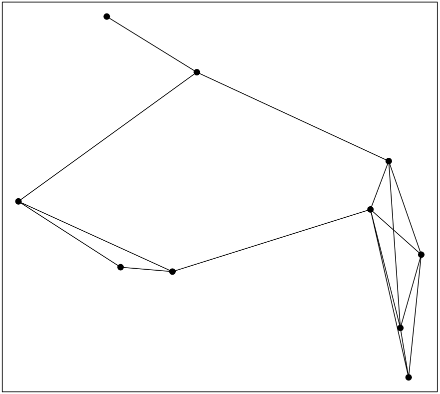
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



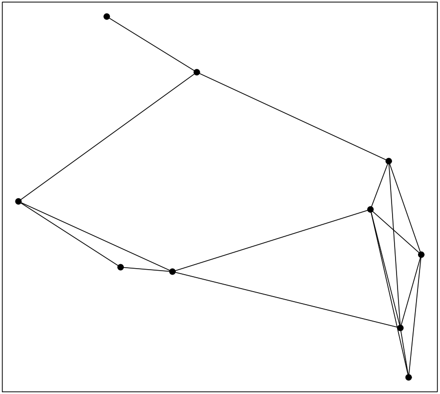
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



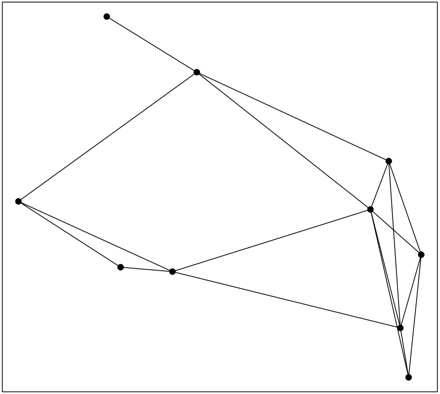
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



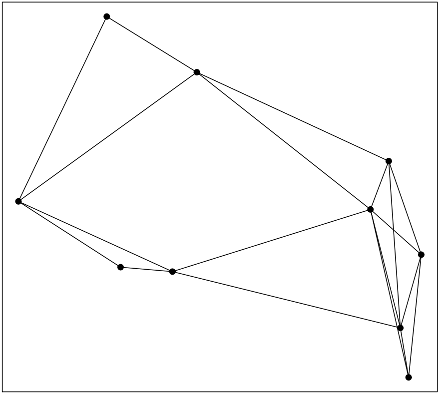
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



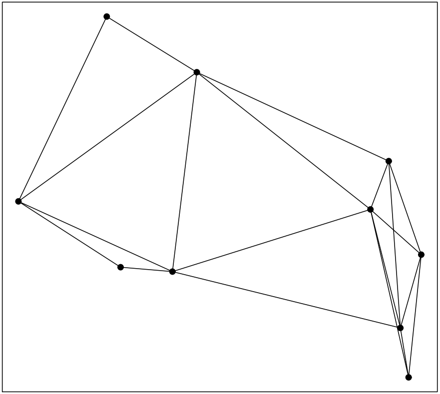
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



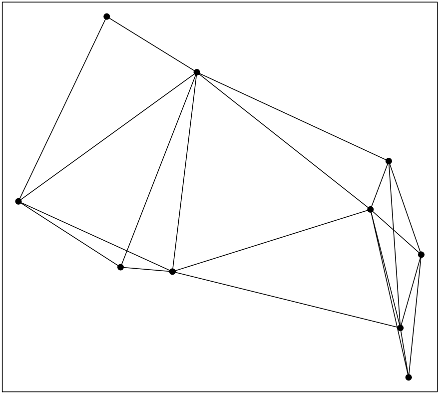
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



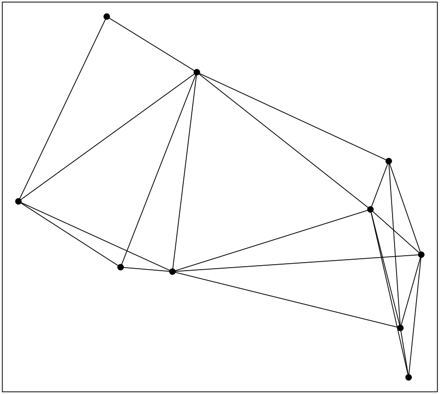
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



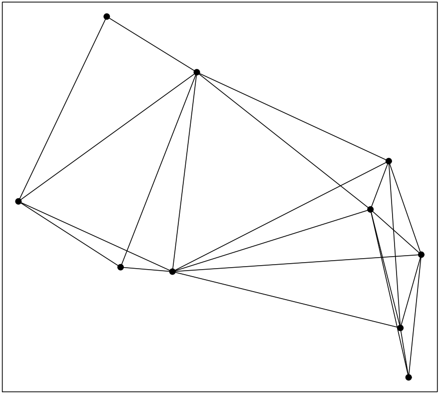
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



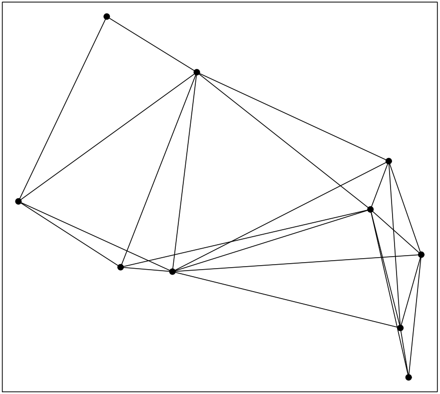
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



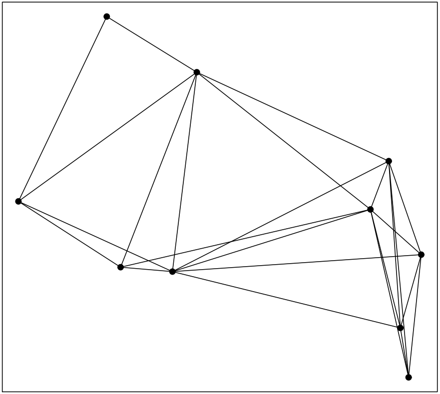
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



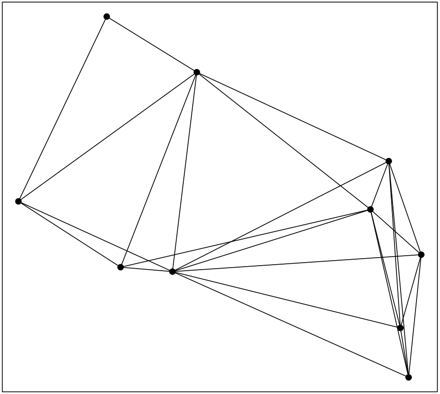
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



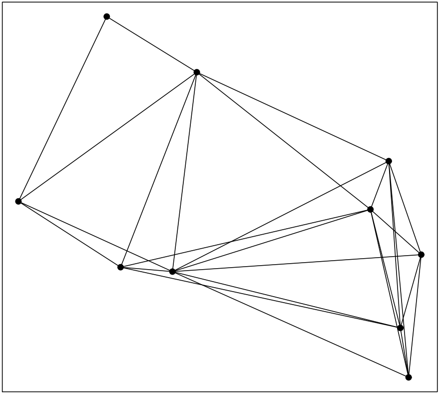
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



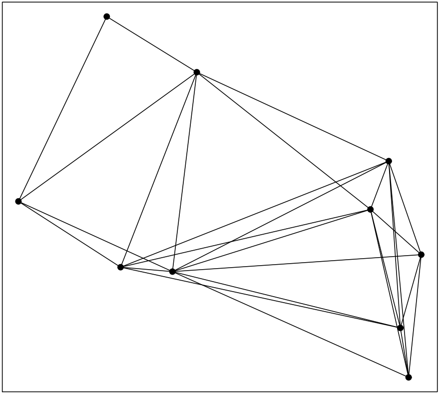
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



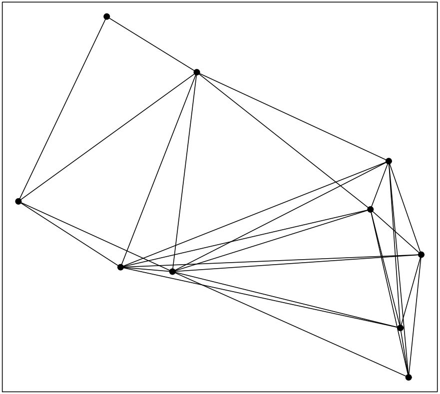
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



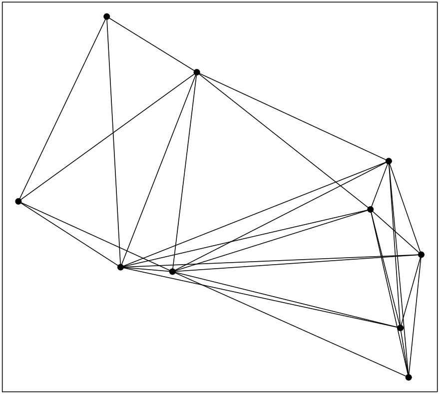
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



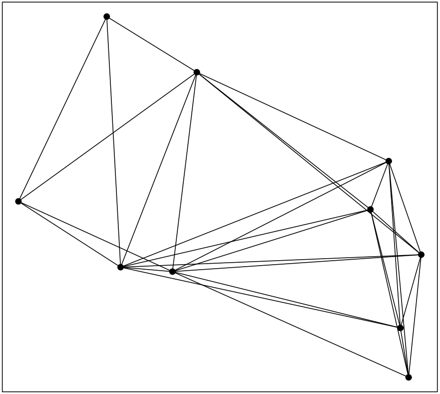
The RGG process

We pick vertices $X_1, \dots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



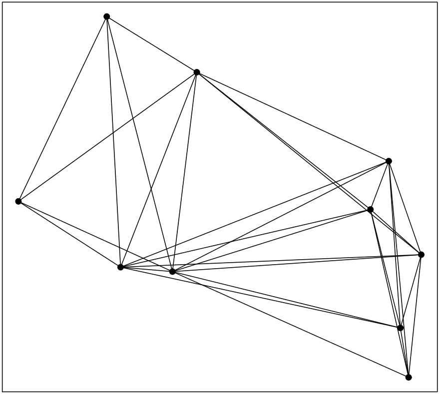
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



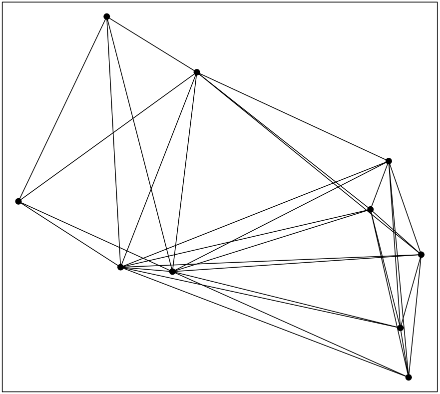
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



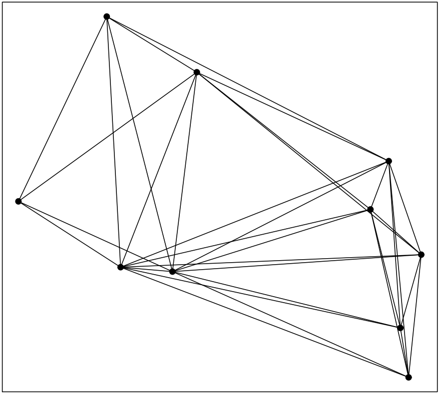
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



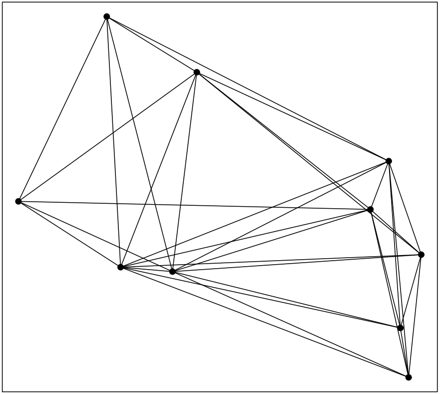
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



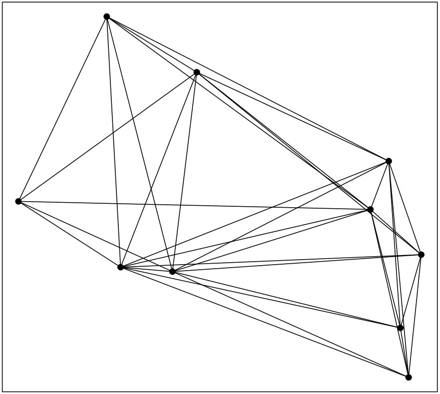
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



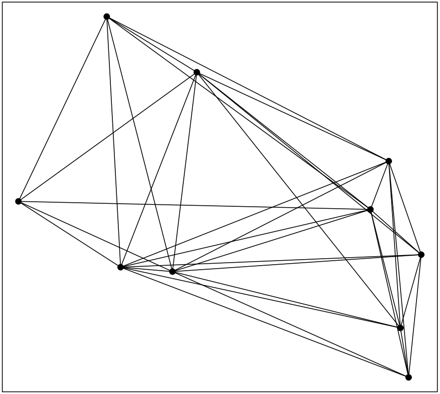
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



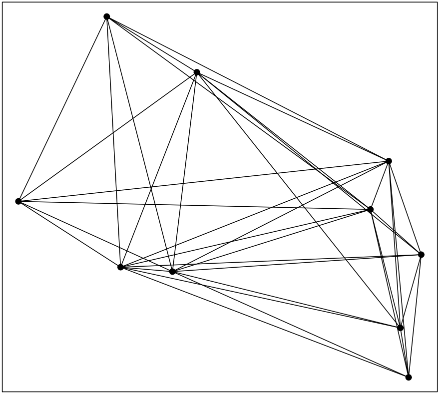
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



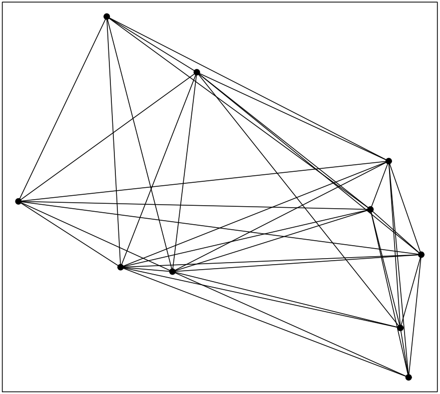
The RGG process

We pick vertices $X_1, \dots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



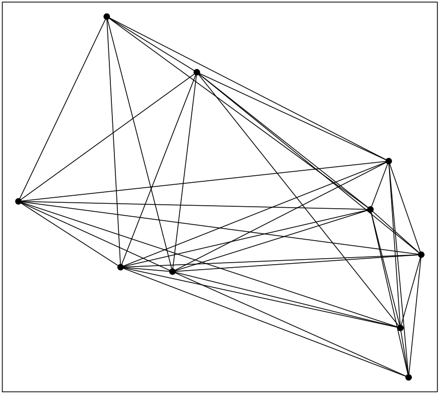
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



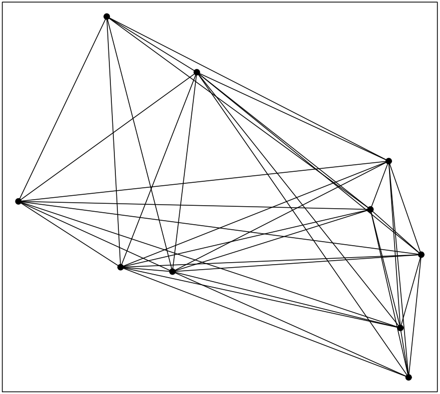
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



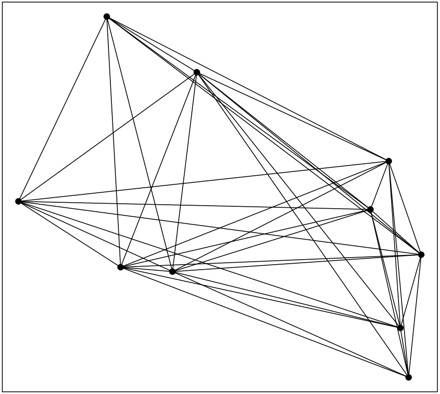
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



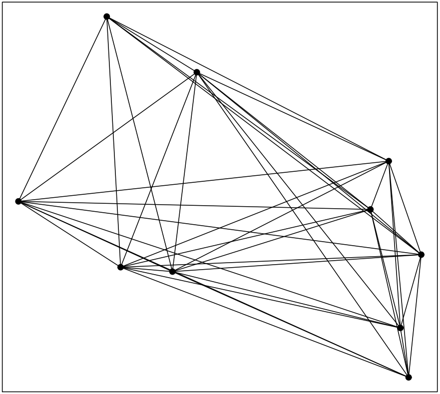
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



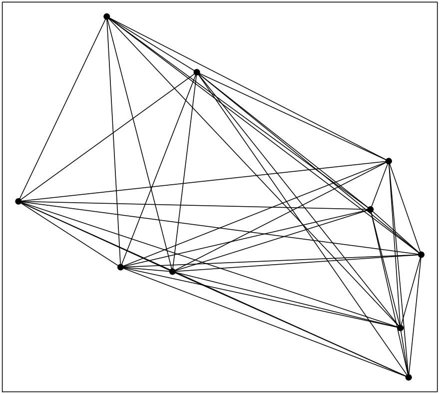
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



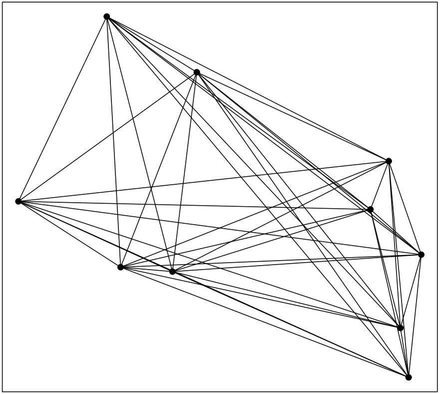
The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



The RGG process

We pick vertices $X_1, \ldots, X_n \in [0,1]^2$ i.i.d. uniformly at random and add the edges one by one in order of increasing edge length.



Definition : the hitting radius

Let $\mathcal{P}$ be an increasing graph property (adding edges cannot destroy the property – e.g. connected, non-planar). The *hitting radius* of $\mathcal{P}$ is defined as:

$$\rho_n(\mathcal{P}) := \min\{r \geq 0 : G(n,r) \text{ satisfies } \mathcal{P}\}.$$

Here we keep the positions of the points $X_1, \ldots, X_n$ fixed as we take the minimum.

Definition : the hitting radius

Let $\mathcal{P}$ be an increasing graph property (adding edges cannot destroy the property – e.g. connected, non-planar). The *hitting radius* of $\mathcal{P}$ is defined as:

$$\rho_n(\mathcal{P}) := \min\{r \geq 0 : G(n,r) \text{ satisfies } \mathcal{P}\}.$$

Here we keep the positions of the points $X_1, \ldots, X_n$ fixed as we take the minimum. So $\rho_n(\mathcal{P})$ is a function of $X_1, \ldots, X_n$ (and hence a random variable).

Definition : the hitting radius

Let $\mathcal{P}$ be an increasing graph property (adding edges cannot destroy the property – e.g. connected, non-planar). The *hitting radius* of $\mathcal{P}$ is defined as:

$$\rho_n(\mathcal{P}) := \min\{r \geq 0 : G(n,r) \text{ satisfies } \mathcal{P}\}.$$

Here we keep the positions of the points $X_1, \ldots, X_n$ fixed as we take the minimum. So $\rho_n(\mathcal{P})$ is a function of $X_1, \ldots, X_n$ (and hence a random variable).

In words, $\rho_n(\mathcal{P})$ is the least r for which $G(n,r)$ satisfies the property $\mathcal{P}$.

Definition : the hitting radius

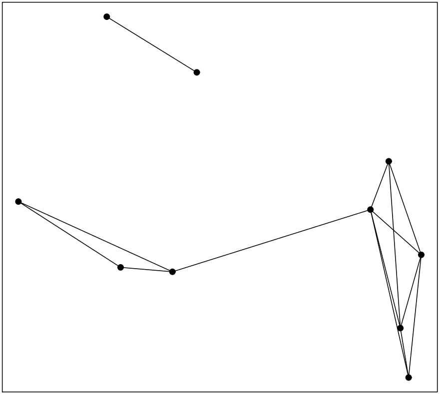
Let $\mathcal{P}$ be an increasing graph property (adding edges cannot destroy the property – e.g. connected, non-planar). The *hitting radius* of $\mathcal{P}$ is defined as:

$$\rho_n(\mathcal{P}) := \min\{r \geq 0 : G(n, r) \text{ satisfies } \mathcal{P}\}.$$

Here we keep the positions of the points $X_1, \ldots, X_n$ fixed as we take the minimum. So $\rho_n(\mathcal{P})$ is a function of $X_1, \ldots, X_n$ (and hence a random variable).

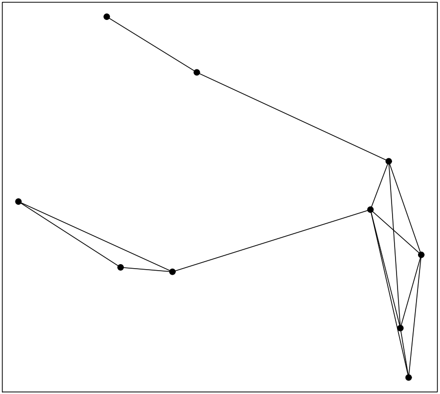
In words, $\rho_n(\mathcal{P})$ is the least r for which $G(n, r)$ satisfies the property $\mathcal{P}$.
In again other words, it is the length of the edge which makes the RGG process satisfy $\mathcal{P}$.

Example: ρ_{10}(connected)



In the particular instance of the RGG process we had two slides ago $\rho_{10}(\text{connected}) \approx 0.4513$.

Example: ρ_{10}(connected)



In the particular instance of the RGG process we had two slides ago $\rho_{10}(\text{connected}) \approx 0.4513$.

Terminology: with high probability

Let $(A_n)_n$ be a sequence of events.

We say that A_n holds with high probability (notation: A_n w.h.p.) if

$$\mathbb{P}(A_n) \to 1,$$

as $n \to \infty$.

A result of Penrose on the hitting radius for connectedness

Theorem.[Penrose'97] $\rho_n(\text{connected}) = \rho_n(\text{min.deg.} \geq 1)$ w.h.p.

This implies the result stated on a earlier slide (I will explain how the corollary follows from the theorem on the next few slides)

Corollary. Let $(r_n)_n$ be a sequence of nonnegative numbers, and write $x_n := \pi n r_n^2 - \ln n$. Then:

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, r_n) \text{ is connected }\right] = \begin{cases} 0 & \text{if } x_n \to -\infty; \\ e^{-e^{-x}} & \text{if } x_n \to x \in \mathbb{R}; \\ 1 & \text{if } x_n \to +\infty. \end{cases}$$

Explanation: why the corollary is a corollary

Note that for any n, r:

$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1) \\ \geq \\ \mathbb{P}(G(n,r) \text{ is connected})$$

(Continued on next slide)

Explanation: why the corollary is a corollary

Note that for any n, r:

$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1)$$
$$\geq$$
$$\mathbb{P}(G(n,r) \text{ is connected})$$
$$=$$
$$\mathbb{P}(G(n,r) \text{ is connected and has min.deg } \geq 1\ )$$

(Continued on next slide)

Explanation: why the corollary is a corollary

Note that for any n, r:

$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1)$$
$$\geq$$
$$\mathbb{P}(G(n,r) \text{ is connected})$$
$$=$$
$$\mathbb{P}(G(n,r) \text{ is connected and has min.deg } \geq 1\ )$$
$$=$$
$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1) -$$
$$\mathbb{P}(G(n,r) \text{ is not connected and has min.deg } \geq 1\ )$$

(Continued on next slide)

Explanation: why the corollary is a corollary

Note that for any n, r:

$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1)$$
$$\geq$$
$$\mathbb{P}(G(n,r) \text{ is connected})$$
$$=$$
$$\mathbb{P}(G(n,r) \text{ is connected and has min.deg } \geq 1\ )$$
$$=$$
$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1) -$$
$$\mathbb{P}(G(n,r) \text{ is not connected and has min.deg } \geq 1\ )$$
$$\geq$$
$$\mathbb{P}(G(n,r) \text{ has min.deg. } \geq 1) -$$
$$\mathbb{P}(\rho_n(\text{connected}) \neq \rho_n(\text{min.deg. } \geq 1)).$$

(Continued on next slide)

Why the corollary is a corollary, continued

By inequalities on the previous slide and the hitting-radius result of Penrose, for any sequence $(r_n)_n$:

$$\mathbb{P}(G(n, r_n) \text{ is connected}) = \mathbb{P}(G(n, r_n) \text{ has min.deg.} \geq 1) - o(1).$$

Why the corollary is a corollary, continued

By inequalities on the previous slide and the hitting-radius result of Penrose, for any sequence $(r_n)_n$:

$$\mathbb{P}(G(n, r_n) \text{ is connected}) = \mathbb{P}(G(n, r_n) \text{ has min.deg.} \geq 1) - o(1).$$

So the probability of being connected is about the same as the probability of having no isolated vertex (= a vertex of degree 0).

(Continued on next slide)

Why the corollary is a corollary, continued

Suppose that $(r_n)_n$ is such that $\pi n r_n^2 - \ln n \to x \in \mathbb{R}$.
Let Z_n denote the number of isolated vertices in $G(n, r_n)$.

Why the corollary is a corollary, continued

Suppose that $(r_n)_n$ is such that $\pi n r_n^2 - \ln n \to x \in \mathbb{R}$.
Let Z_n denote the number of isolated vertices in $G(n, r_n)$.
Observe that

$$\mathbb{E}Z_n \approx n \cdot (1 - \pi r_n^2)^{n-1} = n \cdot \left(1 - \frac{\ln n + x}{n}\right)^{n-1} \to e^{-x}.$$

Why the corollary is a corollary, continued

Suppose that $(r_n)_n$ is such that $\pi n r_n^2 - \ln n \to x \in \mathbb{R}$.
Let Z_n denote the number of isolated vertices in $G(n, r_n)$.
Observe that

$$\mathbb{E}Z_n \approx n \cdot (1 - \pi r_n^2)^{n-1} = n \cdot \left(1 - \frac{\ln n + x}{n}\right)^{n-1} \to e^{-x}.$$

It is relatively straightforward to show that

$$\mathbb{P}(Z_n = 0) \to e^{-e^{-x}}.$$

In fact Z_n is approximately distributed like a $\text{Poisson}(e^{-x})$. $\square$

Why the corollary is a corollary, continued

Suppose that $(r_n)_n$ is such that $\pi n r_n^2 - \ln n \to x \in \mathbb{R}$.
Let Z_n denote the number of isolated vertices in $G(n, r_n)$.
Observe that

$$\mathbb{E}Z_n \approx n \cdot (1 - \pi r_n^2)^{n-1} = n \cdot \left(1 - \frac{\ln n + x}{n}\right)^{n-1} \to e^{-x}.$$

It is relatively straightforward to show that

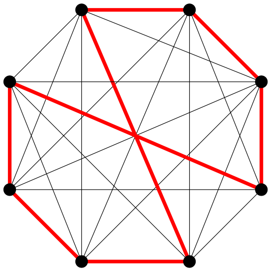
$$\mathbb{P}(Z_n = 0) \to e^{-e^{-x}}.$$

In fact Z_n is approximately distributed like a $\text{Poisson}(e^{-x})$. □

(By monotonicity we only need to consider the range when $\pi n r^2 - \ln n$ is constant to prove the corollary.)

Definition: Hamilton cycle

A *Hamilton cycle* in a graph $G = (V, E)$ is a closed walk that vists every vertex exactly once



G is *Hamiltonian* = G has a Hamilton cycle

Hamiltonicity = having a Hamilton cycle.

Earlier work on Hamiltonicity

Theorem. [Petit'01] there exists a $C > 0$ such that if $\pi n r_n^2 \geq C \ln n$ then

$$\lim_{n \to \infty} \mathbb{P}(G(n, r_n) \text{ has a Hamilton cycle }) = 1.$$

This was later improved to:

Theorem. [Diaz+Mitsche+Perez'07] For any fixed $\varepsilon > 0$, if $\pi n r_n^2 \geq (1 + \varepsilon) \ln n$ then

$$\lim_{n \to \infty} \mathbb{P}(G(n, r_n) \text{ has a Hamilton cycle }) = 1.$$

An observation

Note that being connected is a necessary condition for having a Hamilton cycle.

An observation

Note that being connected is a necessary condition for having a Hamilton cycle.

So by Penrose's result on connectivity
$\mathbb{P}(G(n, r_n) \text{ is Hamiltonian }) \to 0$ whenever $\pi n r_n^2 \leq (1-\varepsilon)\ln n$ for some fixed $\varepsilon > 0$.

An observation

Note that being connected is a necessary condition for having a Hamilton cycle.

So by Penrose's result on connectivity
$\mathbb{P}(G(n, r_n)$ is Hamiltonian $) \to 0$ whenever $\pi n r_n^2 \leq (1-\varepsilon) \ln n$ for some fixed $\varepsilon > 0$.

In other words, the Diaz-Mitsche-Perez result shows that

$$\frac{\rho_n(\text{Hamiltonian})}{\sqrt{\frac{\ln n}{\pi n}}} \to 1 \quad \text{in probability.}$$

Penrose's question

Penrose [2003] asked whether

$$\rho_n(\text{Hamiltonian}) = \rho_n(\text{min.deg} \geq 2) \quad \text{w.h.p.}$$

This would establish an analogue of a celebrated theorem by Ajtai+Komlós+Szemerédi'85 and independently Bollobás'84 on the Erdős-Rényi random graph.

The answer to Penrose's question is yes

Theorem. [KM, BBW, PW 09+]
$\rho_n(\text{Hamiltonian}) = \rho_n(\text{min.deg.} \geq 2)$ w.h.p.

Corollary. Let $(r_n)_n$ be a sequence of nonnegative numbers, and write $x_n := \pi n r_n^2 - (\ln n + \ln \ln n)$. Then:

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, r_n) \text{ is Hamiltonian } \right] = \begin{cases} 0 & \text{if } x_n \to -\infty; \\ e^{-\sqrt{\pi e^{-x}} - e^{-x}} & \text{if } x_n \to x \in \mathbb{R}; \\ 1 & \text{if } x_n \to +\infty. \end{cases}$$

An extension

Let us write

k-EDHs := there exist k edge-disjoint Hamilton cycles.

Theorem. [M+Perez+Wormald, 09+] For each fixed k we have:

$$\rho_n(k\text{-EDHs}) = \rho_n(\text{min.deg.} \geq 2k) \quad \text{w.h.p.}$$

More extensions

pancyclic := $\exists$ cycles of all lengths between 3 and n.
Hamilton connected := between any two points $\exists$ a Hamilton path.

Theorem. $\rho_n(\text{pancyclic}) = \rho_n(\text{min.deg.} \geq 2)$ w.h.p.

Theorem. $\rho_n(\text{Hamilton connected}) = \rho_n(\text{min.deg.} \geq 3)$ w.h.p.

All are analogues of classical results for the Erdős-Rényi random graph.

Part IV: the power of two choices.
(Work in progress)

The power of two choices

There are n *rounds*, and in each round *two* random points arrive.

The power of two choices

There are n *rounds*, and in each round *two* random points arrive.

There is a player who has to decide which of the two points two points to keep. (He does not know the future).

The power of two choices

There are n *rounds*, and in each round *two* random points arrive.

There is a player who has to decide which of the two points two points to keep. (He does not know the future).

The player wants to delay or speed up some property / event, such as having a linear size component.

The birth of a "giant component"

Let us denote by $G(n, r, S)$ the (random geometric) graph we end up after n rounds, when the player plays strategy S.

The birth of a "giant component"

Let us denote by $G(n, r, S)$ the (random geometric) graph we end up after n rounds, when the player plays strategy S.

Suppose the player plays randomly.
If $(r_n)_n$ is such that $\pi n r_n^2 \geq \lambda_{\text{crit}} + \varepsilon$ then there is a linear size component, i.e.

$$L(G(n, r_n, S) \geq c \cdot n \quad \text{w.h.p.}$$

(This follows from a result of Penrose.)

The birth of a "giant component"

Let us denote by $G(n, r, S)$ the (random geometric) graph we end up after n rounds, when the player plays strategy S.

Suppose the player plays randomly.
If $(r_n)_n$ is such that $\pi n r_n^2 \geq \lambda_{\text{crit}} + \varepsilon$ then there is a linear size component, i.e.

$$L(G(n, r_n, S) \geq c \cdot n \quad \text{w.h.p.}$$

(This follows from a result of Penrose.)

What happens if we play optimally rather than randomly?

Birth control

Theorem.[M+Spöhel, 11+] There exist functions $f, g : [0, \infty) \to (0, 1)$ such that the following hold. If $(r_n)_n$ is such that

$$\pi n r_n^2 = c \cdot n^{\frac{1}{3}} / (\ln \ln n)^{\frac{2}{3}},$$

for some fixed $c > 0$ then

(i) There exists a strategy S such that $L(G(n, r, S)) \leq f(c)n$ w.h.p.;

(ii) For every strategy S, we have that $L(G(n, r, S)) \geq g(c)n$ w.h.p.

Moreover, $f(c) \to 0$ as $c \downarrow 0$ and $g(c) \to 1$ as $c \to \infty$.

Thank you for your attention!