Phase Transitions in Random Graphs

Mihyun Kang

TU Graz

RandNET Summer School and Workshop on Random Graphs

Eindhoven, August 2022

Outline of the minicourse

I. Prelude

II. Basic probabilistic tools

III. Erdős-Rényi random graphs

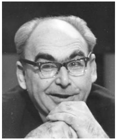
IV. Higher-dimensional analogues

V. Random subgraphs of the hypercube

VI. Topological aspects of random graphs

The Beginning

Erdős (1913 – 1996)



Rényi (1921 – 1970)

On random graphs I.

Dedicated to O. Varga, at the occasion of his 50^{th} birthday.

By P. ERDŐS and A. RÉNYI (Budapest).

Let us consider a "random graph" $\Gamma_{n,N}$ having n possible (labelled) vertices and N edges; in other words, let us choose at random (with equal probabilities) one of the $\binom{\binom{n}{2}}{N}$ possible graphs which can be formed from the n (labelled) vertices $P_1, P_2, \dots, P_n$ by selecting N edges from the $\binom{n}{2}$ possible edges $\widehat{P_iP_j}$ $(1 \leqq i < j \leqq n)$. Thus the effective number of vertices of $\Gamma_{n,N}$ may be less than n, as some points P_i may be not connected in $\Gamma_{n,N}$ with any other point P_j; we shall call such points P_i *isolated* points. We consider the isolated points also as belonging to $\Gamma_{n,N}$. $\Gamma_{n,N}$ is called completely connected if it effectively contains all points $P_1, P_2, \dots, P_n$ (i. e. if it has no isolated points) and is connected in the ordinary sense. In the present paper we consider asymptotic statistical properties of random graphs for $n \to +\infty$. We shall deal with the following questions:

1. What is the probability of $\Gamma_{n,N}$ being completely connected?
2. What is the probability that the greatest connected component (sub-graph) of $\Gamma_{n,N}$ should have effectively $n-k$ points? $(k = 0, 1, \dots)$.
3. What is the probability that $\Gamma_{n,N}$ should consist of exactly $k+1$ connected components? $(k = 0, 1, \dots)$.
4. If the edges of a graph with n vertices are chosen successively so that after each step every edge which has not yet been chosen has the same probability to be chosen as the next, and if we continue this process until the graph becomes completely connected, what is the probability that the number of necessary steps ν will be equal to a given number l?

ON THE EVOLUTION OF RANDOM GRAPHS

by

P. ERDŐS and A. RÉNYI

Dedicated to Professor P. Turán at his 50th birthday.

§ 9. On the growth of the greatest component

We prove in this § (see Theorem 9b) that the size of the greatest component of $\Gamma_{n,N(n)}$ is for $N(n) \sim cn$ with $c > 1/2$ with probability tending to 1 approximately $G(c)n$ where

$$G(c) = 1 - \frac{x(c)}{2c} \tag{9.1}$$

and $x(c)$ is defined by (6.4). (The curve $y = G(c)$ is shown on Fig. 2b).

Thus by Theorem 6 for $N(n) \sim cn$ with $c > 1/2$ almost all points of $\Gamma_{n,N(n)}$ (i. e. all but $o(n)$ points) belong either to some small component which is a tree (of size at most $1/\alpha$ $(\log n - \frac{5}{2}\log\log n) + O(1)$ where $\alpha = 2c - 1 - \log 2c$ by Theorem 7a) or to the single "giant" component of the size $\sim G(c)n$.

Thus the situation can be summarized as follows: the largest component of $\Gamma_{n,N(n)}$ is of order $\log n$ for $\frac{N(n)}{n} \sim c < 1/2$, of order $n^{2/3}$ for $\frac{N(n)}{n} \sim \frac{1}{2}$ and of order n for $\frac{N(n)}{n} \sim c > 1/2$. This double "jump" of the size of the largest component when $\frac{N(n)}{n}$ passes the value $1/2$ is one of the most striking facts concerning random graphs. We prove first the following

Main topics of the minicourse

- What is the probability of a random graph being completely connected?
- Emergence of the giant component
- Double jump vs smooth phase transition

On random graphs I.

Dedicated to O. Varga, at the occasion of his 50^{th} birthday.

By P. ERDŐS and A. RÉNYI (Budapest).

Let us consider a "random graph" $\Gamma_{n,N}$ having n possible (labelled) vertices and N edges; in other words, let us choose at random (with equal probabilities) one of the $\binom{\binom{n}{2}}{N}$ possible graphs which can be formed from the n (labelled) vertices $P_1, P_2, \ldots, P_n$ by selecting N edges from the $\binom{n}{2}$ possible edges $\overline{P_iP_j}$ ($1 \leqq i < j \leqq n$). Thus the effective number of vertices of $\Gamma_{n,N}$ may be less than n, as some points P_i may be not connected in $\Gamma_{n,N}$ with any other point P_j; we shall call such points P_i *isolated* points. We consider the isolated points also as belonging to $\Gamma_{n,N}$. $\Gamma_{n,N}$ is called completely connected if it effectively contains all points $P_1, P_2, \ldots, P_n$ (i. e. if it has no isolated points) and is connected in the ordinary sense. In the present paper we consider asymptotic statistical properties of random graphs for $n \to +\infty$. We shall deal with the following questions:

1. What is the probability of $\Gamma_{n,N}$ being completely connected?
2. What is the probability that the greatest connected component (sub-graph) of $\Gamma_{n,N}$ should have effectively $n-k$ points? ($k = 0, 1, \ldots$).
3. What is the probability that $\Gamma_{n,N}$ should consist of exactly $k+1$ connected components? ($k = 0, 1, \ldots$).
4. If the edges of a graph with n vertices are chosen successively so that after each step every edge which has not yet been chosen has the same probability to be chosen as the next, and if we continue this process until the graph becomes completely connected, what is the probability that the number of necessary steps ν will be equal to a given number l?

ON THE EVOLUTION OF RANDOM GRAPHS

by

P. ERDŐS and A. RÉNYI

Dedicated to Professor P. Turán at his 50th birthday.

§ 9. On the growth of the greatest component

We prove in this § (see Theorem 9b) that the size of the greatest component of $\Gamma_{n,N(n)}$ is for $N(n) \sim cn$ with $c > {}^1\!/_2$ with probability tending to 1 approximately $G(c)n$ where

$$G(c) = 1 - \frac{x(c)}{2c} \tag{9.1}$$

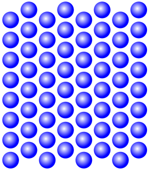
and $x(c)$ is defined by (6.4). (The curve $y = G(c)$ is shown on Fig. 2b).

Thus by Theorem 6 for $N(n) \sim cn$ with $c > {}^1\!/_2$ almost all points of $\Gamma_{n,N(n)}$ (i. e. all but $o(n)$ points) belong either to some small component which is a tree (of size at most $1/\alpha$ ($\log n - \frac{5}{2}\log\log n) + O(1)$ where $\alpha = 2c - 1 - \log 2c$ by Theorem 7a) or to the single "giant" component of the size $\sim G(c)n$.

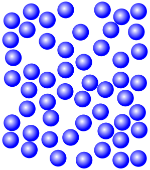
Thus the situation can be summarized as follows: the largest component of $\Gamma_{n,N(n)}$ is of order $\log n$ for $\frac{N(n)}{n} \sim c < {}^1\!/_2$, of order $n^{2/3}$ for $\frac{N(n)}{n} \sim \frac{1}{2}$ and of order n for $\frac{N(n)}{n} \sim c > {}^1\!/_2$. This double "jump" of the size of the largest component when $\frac{N(n)}{n}$ passes the value ${}^1\!/_2$ is one of the most striking facts concerning random graphs. We prove first the following

Phase transition

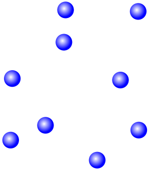
The phase transition deals with a sudden change in the properties of a large structure, caused by altering a critical parameter.



low temperature
dense, ordered

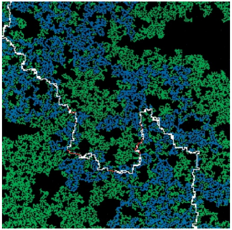


mid temperature
short-range order



high temperature
sparse, irregular

Percolation in physics, materials science, . . .

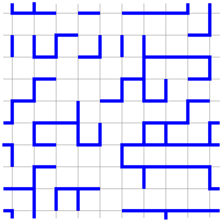





Mathematical models of percolation

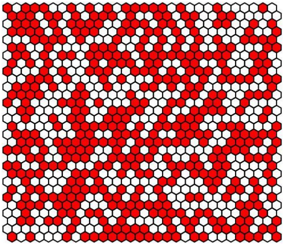
- Bond percolation: each bond (or edge) is either open with prob. p or closed with prob. $1-p$, independently
- Site percolation: each site (or vertex) is either occupied with prob. p or empty with prob. $1-p$, independently

$p < p_c$



Bond percolation on square lattice

$p > p_c$

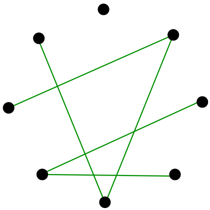


Site percolation on hexagonal lattice

Erdős-Rényi random graphs

Let $G(n,p)$ denote a binomial random graph:

a graph on vertex set $[n]$, in which each pair of vertices is joined by an edge with probability $p = p(n)$, independently

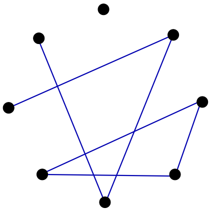


$*$ Bond percolation on the complete graph K_n

Erdős-Rényi random graphs – cont'd

Let $G(n,m)$ denote a uniform random graph:

a graph taken uniformly at random from the set $\mathcal{G}(n,m)$ of all graphs on vertex set $[n] := \{1,\ldots,n\}$ with $m = m(n)$ edges

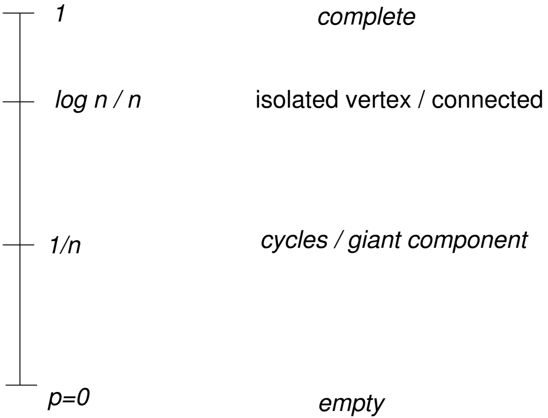


* $G(n,p)$ and $G(n,m)$ are 'essentially equivalent' when $m \sim \binom{n}{2} p$

  and they are called Erdős-Rényi random graphs

Phase transitions in $G(n, p)$

Let $p = p(n) \in [0, 1]$



Part II.

Basic probabilistic tools

- First moment method
- Second moment method
- Chernoff bounds
- Sprinkling argument
- Galton-Watson branching process

First moment method

Markov's inequality

Let X be a non-negative integer-valued random variable. Then for every $a > 0$

$$\mathbb{P}\left[X \geq a\right] \quad \leq \quad \frac{\mathbb{E}\left[X\right]}{a}$$

In particular, $$\mathbb{P}\left[X \geq 1\right] \quad \leq \quad \mathbb{E}\left[X\right]$$

First moment method

Markov's inequality

Let X be a non-negative integer-valued random variable. Then for every $a > 0$

$$\mathbb{P}\left[X \geq a\right] \quad \leq \quad \frac{\mathbb{E}\left[X\right]}{a}$$

In particular, $$\mathbb{P}\left[X \geq 1\right] \quad \leq \quad \mathbb{E}\left[X\right]$$

For example, let $X_n \; = \; \#$ isolated vertices in $G(n,p)$ for $p = \ldots$

IF

$$\mathbb{E}\left[X_n\right] \quad \xrightarrow{n\to\infty} \quad 0,$$

THEN

$$\mathbb{P}\left[G(n,p) \text{ contains an isolated vertex}\right]$$
$$= \quad \mathbb{P}\left[X_n \geq 1\right] \quad \leq \quad \mathbb{E}\left[X_n\right] \quad \xrightarrow{n\to\infty} \quad 0$$

Second moment method

Chebyshev's inequality

Let X be a random variable with $\mathbb{E}[X] > 0$. Then

$$\mathbb{P}[X = 0] \quad \leq \quad \mathbb{P}\big[\, |X - \mathbb{E}[X]| \geq \mathbb{E}[X] \,\big] \quad \leq \quad \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2}$$

Let $X = X_1 + X_2 + \ldots$ be a sum of indicator random variables with $\mathbb{E}[X] > 0$.

Then

$$\mathbb{P}[X = 0] \quad \leq \quad \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2} \quad \leq \quad \frac{1}{\mathbb{E}[X]} + \frac{\sum_{i \neq j} \mathrm{Cov}[X_i, X_j]}{\mathbb{E}[X]^2},$$

where $\mathrm{Cov}[X_i, X_j] \; = \; \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$.

Second moment method

Chebyshev's inequality

Let X be a random variable with $\mathbb{E}[X] > 0$. Then

$$\mathbb{P}[X=0] \quad \leq \quad \mathbb{P}\big[\, |X - \mathbb{E}[X]| \geq \mathbb{E}[X] \,\big] \quad \leq \quad \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2}$$

Let $X = X_1 + X_2 + \ldots$ be a sum of indicator random variables with $\mathbb{E}[X] > 0$.

Then

$$\mathbb{P}[X=0] \quad \leq \quad \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2} \quad \leq \quad \frac{1}{\mathbb{E}[X]} + \frac{\sum_{i \neq j} \mathrm{Cov}[X_i, X_j]}{\mathbb{E}[X]^2},$$

where $\mathrm{Cov}[X_i, X_j] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$.

For example, let $X_n = \#$ isolated vertices in $G(n,p)$ for $p = \ldots$.

IF

$$\mathbb{E}[X_n] \xrightarrow{n\to\infty} \infty \quad \text{and} \quad \frac{\sum_{i\neq j} \mathrm{Cov}[X_{n,i}, X_{n,j}]}{\mathbb{E}[X_n]^2} \xrightarrow{n\to\infty} 0,$$

THEN $\mathbb{P}[G(n,p) \text{ contains no isolated vertex}] = \mathbb{P}[X_n = 0] \xrightarrow{n\to\infty} 0$

Method of moments

Let $(X_n)_{n\geq 1}$ be a sequence of sums of indicator random variables.

Suppose $\exists\, \lambda > 0$ such that $\forall\, k \in \mathbb{N}$, the k-th binomial moment satisfies

$$\lim_{n\to\infty} \mathbb{E}\binom{X_n}{k} = \frac{\lambda^k}{k!}.$$

Then

$$X_n \xrightarrow{D} \mathrm{Po}(\lambda)$$

i.e., X_n converges in distribution to a Poisson random variable with mean λ.

In particular,

$$\lim_{n\to\infty} \mathbb{P}\big[X_n = 0\big] = e^{-\lambda}$$

Method of moments

Let $(X_n)_{n\geq 1}$ be a sequence of sums of indicator random variables.

Suppose $\exists\ \lambda > 0$ such that $\forall\ k \in \mathbb{N}$, the k-th binomial moment satisfies

$$\lim_{n\to\infty} \mathbb{E}\binom{X_n}{k} = \frac{\lambda^k}{k!}.$$

Then

$$X_n \xrightarrow{D} \mathrm{Po}(\lambda)$$

i.e., X_n converges in distribution to a Poisson random variable with mean λ.

In particular,
$$\lim_{n\to\infty} \mathbb{P}\left[X_n = 0\right] = e^{-\lambda}$$

For example, let $X_n = \#$ isolated vertices in $G(n,p)$ for $p = \ldots$.

IF

$$\mathbb{E}\binom{X_n}{k} \xrightarrow{n\to\infty} \frac{\left(e^{-c}\right)^k}{k!}, \qquad \forall k \in \mathbb{N},$$

THEN

$$\mathbb{P}\left[G(n,p) \text{ contains no isolated vertex}\right] = \mathbb{P}\left[X_n = 0\right] \xrightarrow{n\to\infty} e^{-e^{-c}}$$

Chernoff bounds

Let $N \in \mathbb{N}$, let $p \in [0,1]$, and let $X \sim \text{Bi}(N,p)$.

- For every $a > 0$,

$$\begin{aligned} \mathbb{P}\left[\, X \;\geq\; Np + a \,\right] &\quad\leq\quad \exp\left(-\frac{a^2}{2(Np + a/3)}\right) \\ \mathbb{P}\left[\, X \;\leq\; Np - a \,\right] &\quad\leq\quad \exp\left(-\frac{a^2}{2Np}\right) \end{aligned}$$

- For every $0 < a \leq \frac{Np}{2}$,

$$\mathbb{P}\left[\, |X - Np| \;\geq\; a \,\right] \quad\leq\quad 2\exp\left(-\frac{a^2}{4Np}\right)$$

Sprinkling argument

- Two round exposure:
  - let $p,\ p_1,\ p_2 \in (0,1)$ satisfy

$$(1-p) = (1-p_1)(1-p_2)$$

  - generate $G(n,p)$, $G(n,p_1)$, and $G(n,p_2)$ independently

Then $G(n,p)$ and $G(n,p_1) \cup G(n,p_2)$ have the same distribution

Sprinkling argument

- Two round exposure:
  - let $p,\ p_1,\ p_2 \in (0,1)$ satisfy

$$(1-p) = (1-p_1)(1-p_2)$$

  - generate $G(n,p),\ G(n,p_1)$, and $G(n,p_2)$ independently

Then $G(n,p)$ and $G(n,p_1) \cup G(n,p_2)$ have the same distribution

- Multi-round exposure:
  - let $p,\ p_1,\ \ldots,\ p_r \in (0,1)$ satisfy

$$(1-p) = \prod_{i=1}^{r}(1-p_i)$$

  - generate $G(n,p),\ G(n,p_1),\ \ldots,\ G(n,p_r)$ independently

Then $G(n,p)$ and $\bigcup_{i=1}^{r} G(n,p_i)$ have the same distribution

Galton-Watson branching process

$\mathrm{GW(Z)}$ $=$ Galton-Watson tree with offspring distribution Z

$=$ a random tree constructed by the Galton-Watson process with offspring distribution Z

- start with a single root vertex
- each vertex has a random number of children with distribution Z
- $\#$ children are independent of each other and of the history

Galton-Watson branching process

$\mathrm{GW(Z)}$ = Galton-Watson tree with offspring distribution Z

= a random tree constructed by the Galton-Watson process with offspring distribution Z

- start with a single root vertex
- each vertex has a random number of children with distribution Z
- $\#$ children are independent of each other and of the history

f_Z = probability generating function of Z

Theorem

- If $\mathbb{E}(Z) > 1$, the GW process survives (i.e., $\mathrm{GW(Z)}$ is infinite) with probability $\rho \in (0,1)$ satisfying $1-\rho = f_Z(1-\rho)$
- If $\mathbb{E}(Z) < 1$, the survival probability of the GW process is zero

Poisson branching process

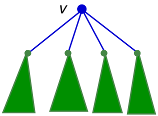
$$\mathrm{GW(Z)} \;=\; \text{Galton-Watson tree with offspring distribution } Z$$

$$f_Z \;=\; \text{probability generating function of } Z$$

$$\rho \;=\; \text{survival probability of GW process with offspring distribution } Z$$

Assume $Z \sim \mathrm{Po}(d)$

$$\begin{aligned}
f_Z(x) &= \sum_{\ell \geq 0} \mathbb{P}\,[\,\text{vertex } v \text{ generates } \ell \text{ children}\,]\, x^\ell \\
&= \sum_{\ell \geq 0} \mathbb{P}\,[\,\mathrm{Po(d)} = \ell\,]\, x^\ell \\
&= \sum_{\ell \geq 0} \exp(-d)\, \frac{d^\ell}{\ell!}\, x^\ell \\
&= \exp(-d) \sum_{\ell \geq 0} \frac{(dx)^\ell}{\ell!} \\
&= \exp\Big(-d(1-x)\Big)
\end{aligned}$$



ℓ subtrees

Binomial branching process

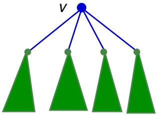
$$\mathrm{GW(Z)} = \text{Galton-Watson tree with offspring distribution } Z$$

$$f_Z = \text{probability generating function of } Z$$

$$\rho = \text{survival probability of GW process with offspring distribution } Z$$

Assume $Z \sim \mathrm{Bi}(n,p)$

$$\begin{aligned}
f_Z(x) &= \sum_{\ell \geq 0} \mathbb{P}\,[\,\text{vertex } v \text{ generates } \ell \text{ children}\,]\, x^\ell \\
&= \sum_{\ell=0}^{n} \mathbb{P}\,[\,\mathrm{Bi}(n,p) = \ell\,]\, x^\ell \\
&= \sum_{\ell=0}^{n} \binom{n}{\ell} p^\ell \,(1-p)^{n-\ell}\, x^\ell \\
&= \big(px + (1-p)\big)^n \\
&= \big(1 - p(1-x)\big)^n
\end{aligned}$$



ℓ subtrees

Extinction probability of branching processes

$\mathrm{GW(Z)}$ $=$ Galton-Watson tree with offspring distribution Z

f_Z $=$ probability generating function of Z

ρ $=$ survival probability of GW process with offspring distribution Z

$\eta = 1 - \rho =$ extinction probability of $\mathrm{GW(Z)}$

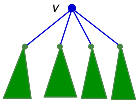
Extinction probability of branching processes

$$\mathrm{GW(Z)} \;=\; \text{Galton-Watson tree with offspring distribution } Z$$

$$f_Z \;=\; \text{probability generating function of } Z$$

$$\rho \;=\; \text{survival probability of GW process with offspring distribution } Z$$

$$\eta \;=\; 1 - \rho \;=\; \text{extinction probability of } \mathrm{GW(Z)}$$



ℓ subtrees

We have

$$\eta \;=\; \sum_{\ell \geq 0} \mathbb{P}\left[\, v \text{ generates } \ell \text{ children } \right] \eta^{\ell} \;=\; f_Z(\eta)$$

In other words,

$$1 - \rho = f_Z(1 - \rho)$$

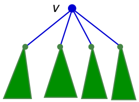
Extinction probability of branching processes

$\mathrm{GW(Z)}$ $=$ Galton-Watson tree with offspring distribution Z

f_Z $=$ probability generating function of Z

ρ $=$ survival probability of GW process with offspring distribution Z

$\eta \;=\; 1 \;-\; \rho \;=$ extinction probability of $\mathrm{GW(Z)}$



ℓ subtrees

We have

$$\eta \;=\; \sum_{\ell \geq 0} \mathbb{P}\left[\, v \text{ generates } \ell \text{ children } \right] \eta^{\ell} \;=\; f_Z(\eta)$$

In other words,

$$1 \;-\; \rho \;=\; f_Z(1 \;-\; \rho)$$

Assume that $d \;:=\; np \;>\; 1$

- If $Z \;\sim\; \mathrm{Po}(d)$, then $\quad 1-\rho \;=\; \exp\left(-d\,\rho\right)$
- If $Z \;\sim\; \mathrm{Bi}(n,p)$, then $\quad 1-\rho \;=\; (1-p\,\rho)^{n} \;\sim\; \exp\left(-d\,\rho\right)$

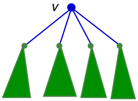
Extinction probability of branching processes

$\mathrm{GW(Z)}$ = Galton-Watson tree with offspring distribution Z

f_Z = probability generating function of Z

ρ = survival probability of GW process with offspring distribution Z

$\eta = 1 - \rho =$ extinction probability of $\mathrm{GW(Z)}$



ℓ subtrees

We have

$$\eta = \sum_{\ell \geq 0} \mathbb{P}[\, v \text{ generates } \ell \text{ children} \,]\, \eta^{\ell} = f_Z(\eta)$$

In other words,

$$1 - \rho = f_Z(1 - \rho)$$

Assume that $d := np > 1$

- If $Z \sim \mathrm{Po}(d)$, then $\quad 1 - \rho = \exp(-d\,\rho)$
- If $Z \sim \mathrm{Bi}(n,p)$, then $\quad 1 - \rho = (1 - p\,\rho)^n \sim \exp(-d\,\rho)$

Thus, if $\varepsilon := d - 1 \xrightarrow{n \to \infty} 0$ with $\varepsilon > 0$ then $\rho \sim 2\varepsilon$

Part III.

Erdős-Rényi random graphs

- Thresholds
- Connectedness threshold
- Percolation threshold
- Coupling with Galton-Watson branching trees
- More about the giant component

Thresholds and sharp thresholds

Let $\mathcal{A}$ be a monotone increasing property.

Threshold

A function $p^* = p^*(n)$ is called a threshold for $\mathcal{A}$ if

$$\mathbb{P}\left[\, G(n,p) \text{ satisfies } \mathcal{A} \,\right] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if} \quad p \ll p^* \\ 1 & \text{if} \quad p \gg p^* \end{cases}$$

Sharp threshold

A function $p^* = p^*(n)$ is called a sharp threshold for $\mathcal{A}$ if $\forall \varepsilon > 0$,

$$\mathbb{P}\left[\, G(n,p) \text{ satisfies } \mathcal{A} \,\right] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if} \quad p \leq (1-\varepsilon)\, p^* \\ 1 & \text{if} \quad p \leq (1+\varepsilon)\, p^* \end{cases}$$

$\ast$ Every monotone property has a threshold [Bollobás–Thomason 87]

Sharp thresholds – two toy examples

- Containment of isolated vertices in $G(n, p)$
- Connectedness in $G(n, p)$

Sharp threshold for isolated vertices

A sharp threshold for the property that $G(n, p)$ contains no isolated vertex is

$$p^* = \frac{\log n}{n}$$

Sharp threshold for isolated vertices

A sharp threshold for the property that $G(n, p)$ contains no isolated vertex is

$$p^* = \frac{\log n}{n}$$

Theorem

Let
$$p = \frac{\log n + c(n)}{n}$$

where $|c(n)| \to \infty$ arbitrarily slowly as $n \to \infty$. Then

$$\mathbb{P}\big[\, G(n,p) \text{ contains no isolated vertex } \big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

Proof sketch

To prove the statement, we may assume without loss of generality that $|c(n)| \ll \log n$.

* The function $\phi : [0,1] \to [0,1]$ defined by

  $$\phi(p) := \mathbb{P}[\ G(n,p) \text{ contains NO isolated vertex }]$$

  is monotone increasing in p.

* The function $\xi : [0,1] \to [0,1]$ defined by

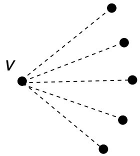
  $$\xi(p) := \mathbb{P}[\ G(n,p) \text{ contains at least one isolated vertex }]$$

  is monotone decreasing in p.

Proof sketch - cont'd

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$

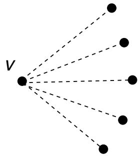


Then $\mathbb{E}\left[X_v\right] = (1-p)^{n-1}$.

Proof sketch - cont'd

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$



Then $\mathbb{E}\left[X_v\right] = (1-p)^{n-1}$.

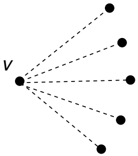
Set $X := \sum_{v \in [n]} X_v$. By linearity expectation we have

$$\mathbb{E}\left[X\right] = \sum_{v \in [n]} \mathbb{E}\left[X_v\right]$$

Proof sketch - cont'd

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$



Then $\mathbb{E}\left[X_v\right] = (1-p)^{n-1}$.

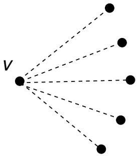
Set $X := \sum_{v \in [n]} X_v$. By linearity expectation we have

$$\begin{aligned} \mathbb{E}\left[X\right] &= \sum_{v \in [n]} \mathbb{E}\left[X_v\right] \\ &= n\,(1-p)^{n-1} \end{aligned}$$

Proof sketch - cont'd

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$



Then $\mathbb{E}[X_v] = (1-p)^{n-1}$.

Set $X := \sum_{v \in [n]} X_v$. By linearity expectation we have

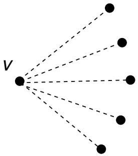
$$\begin{aligned} \mathbb{E}[X] &= \sum_{v \in [n]} \mathbb{E}[X_v] \\ &= n\,(1-p)^{n-1} \\ &= \exp\big(\log n - pn + p + O(p^2 n)\big), \end{aligned}$$

using $\quad 1 - x = \exp(1 - x + O(x^2)) \quad$ for $\quad x = o(1)$.

Proof sketch - cont'd

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$



Then $\mathbb{E}\big[X_v\big] = (1-p)^{n-1}$.

Set $X := \sum_{v\in[n]} X_v$. By linearity expectation we have

$$\begin{aligned} \mathbb{E}\big[X\big] &= \sum_{v\in[n]} \mathbb{E}\big[X_v\big] \\ &= n\ (1-p)^{n-1} \\ &= \exp\big(\log n - pn + p + O(p^2 n)\big), \end{aligned}$$

using $\quad 1 - x = \exp(1 - x + O(x^2)) \quad$ for $\quad x = o(1)$.

Taking $p = \frac{\log n + c(n)}{n}$ with $|c(n)| \ll \log n$, we have

$$\mathbb{E}\big[X\big] = (1+o(1)) \exp\big(-c(n)\big).$$

Proof sketch - cont'd

Note $X := \sum_{v \in [n]} X_v$ is equal to the number of isolated vertices in $G(n,p)$ and we have

$$\mathbb{E}[X] = (1+o(1)) \exp\big(-c(n)\big).$$

Proof sketch - cont'd

Note $X := \sum_{v\in[n]} X_v$ is equal to the number of isolated vertices in $G(n,p)$ and we have

$$\mathbb{E}\big[X\big] \;= (1+o(1))\,\exp\big(-c(n)\big).$$

Case (1): assume that $p \;=\; \frac{\log n + c(n)}{n}$ with $c(n) \to \infty$.

Using the first moment method, we have

$$\mathbb{P}\big[X \geq 1\big] \;\leq\; \mathbb{E}\big[X\big] \;=\; (1+o(1))\,\exp\big(-c(n)\big) \;\to\; 0,$$

and

$$\mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \;=\; \mathbb{P}\big[X \geq 1\big] \;\to\; 0.$$

Proof sketch - cont'd

Note $X := \sum_{v\in[n]} X_v$ is equal to the number of isolated vertices in $G(n,p)$ and we have

$$\mathbb{E}\big[X\big] = (1+o(1))\exp\big(-c(n)\big).$$

Case (1): assume that $p = \frac{\log n + c(n)}{n}$ with $c(n) \to \infty$.

Using the first moment method, we have

$$\mathbb{P}\big[X \geq 1\big] \leq \mathbb{E}\big[X\big] = (1+o(1))\exp\big(-c(n)\big) \to 0,$$

and

$$\mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] = \mathbb{P}\big[X \geq 1\big] \to 0.$$

Summing up, if $p = \frac{\log n + c(n)}{n}$ with $c(n) \to \infty$, then

$$\mathbb{P}\big[G(n,p) \text{ contains no isolated vertex}\big] = \mathbb{P}\big[X = 0\big] \to 1.$$

Proof sketch - cont'd

Case (2): assume that $p = \frac{\log n + c(n)}{n}$ with $c(n) \to -\infty$.

We have $$\mathbb{E}[X] = (1 + o(1)) \exp\big(-c(n)\big) \to \infty$$

Proof sketch - cont'd

Case (2): assume that $p = \frac{\log n + c(n)}{n}$ with $c(n) \to -\infty$.

We have $$\mathbb{E}\left[X\right] = (1 + o(1)) \exp\left(-c(n)\right) \to \infty$$

For $v \neq w$,

$$\begin{aligned} \mathrm{Cov}\left[X_v, X_w\right] &= \mathbb{E}\left[X_v X_w\right] - \mathbb{E}\left[X_v\right]\mathbb{E}\left[X_w\right] \\ &= (1-p)^{2n-3} - (1-p)^{2n-2} = p(1-p)^{2n-3} \end{aligned}$$

and therefore

$$\frac{\sum_{v \neq w} \mathrm{Cov}\left[X_v, X_w\right]}{\mathbb{E}\left[X\right]^2} = \frac{n(n-1)p(1-p)^{2n-3}}{n^2(1-p)^{2n-2}} \sim \frac{p}{1-p} \to 0$$

Proof sketch - cont'd

Case (2): assume that $p = \frac{\log n + c(n)}{n}$ with $c(n) \to -\infty$.

We have $\mathbb{E}\left[X\right] = (1 + o(1)) \exp\left(-c(n)\right) \to \infty$

For $v \neq w$,

$$\begin{aligned} \mathrm{Cov}\left[X_v, X_w\right] &= \mathbb{E}\left[X_v X_w\right] - \mathbb{E}\left[X_v\right]\mathbb{E}\left[X_w\right] \\ &= (1-p)^{2n-3} - (1-p)^{2n-2} = p(1-p)^{2n-3} \end{aligned}$$

and therefore

$$\frac{\sum_{v \neq w} \mathrm{Cov}\left[X_v, X_w\right]}{\mathbb{E}\left[X\right]^2} = \frac{n(n-1)p(1-p)^{2n-3}}{n^2(1-p)^{2n-2}} \sim \frac{p}{1-p} \to 0$$

Using the second moment method, we have

$$\mathbb{P}\left[X = 0\right] \leq \frac{1}{\mathbb{E}\left[X\right]} + \frac{\sum_{v \neq w} \mathrm{Cov}\left[X_v, X_w\right]}{\mathbb{E}\left[X\right]^2} \to 0.$$

Proof sketch - cont'd

Case (2): assume that $p = \frac{\log n + c(n)}{n}$ with $c(n) \to -\infty$.

We have $$\mathbb{E}\left[X\right] = (1 + o(1)) \exp\left(-c(n)\right) \to \infty$$

For $v \neq w$,

$$\begin{aligned}\mathrm{Cov}\left[X_v, X_w\right] &= \mathbb{E}\left[X_v X_w\right] - \mathbb{E}\left[X_v\right]\mathbb{E}\left[X_w\right] \\ &= (1-p)^{2n-3} - (1-p)^{2n-2} = p(1-p)^{2n-3}\end{aligned}$$

and therefore

$$\frac{\sum_{v \neq w} \mathrm{Cov}\left[X_v, X_w\right]}{\mathbb{E}\left[X\right]^2} = \frac{n(n-1)p(1-p)^{2n-3}}{n^2(1-p)^{2n-2}} \sim \frac{p}{1-p} \to 0$$

Using the second moment method, we have

$$\mathbb{P}\left[X = 0\right] \leq \frac{1}{\mathbb{E}\left[X\right]} + \frac{\sum_{v \neq w} \mathrm{Cov}\left[X_v, X_w\right]}{\mathbb{E}\left[X\right]^2} \to 0.$$

Summing up, if $p = \frac{\log n + c(n)}{n}$ with $c(n) \to -\infty$, then

$$\mathbb{P}\left[G(n,p) \text{ contains no isolated vertex}\right] = \mathbb{P}\left[X = 0\right] \to 0.$$

Sharp threshold for isolated vertices

Theorem

Let
$$p = \frac{\log n + c(n)}{n}$$

where $|c(n)| \to \infty$ arbitrarily slowly as $n \to \infty$. Then

$$\mathbb{P}\big[\, G(n,p) \text{ contains no isolated vertex } \big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

$*$ What happens when $c(n) \to c \in \mathbb{R}$?

Isolated vertices in critical window

Theorem

Let

$$p = \frac{\log n + c(n)}{n}, \quad \text{where} \quad c(n) \to c \in \mathbb{R}.$$

Let $X = X(n)$ be $\#$ isolated vertices in $G(n,p)$. Then

$$X \xrightarrow{D} \mathrm{Po}(e^{-c}).$$

It means, for every $\ell = 0, 1, 2, \ldots$

$$\lim_{n\to\infty} \mathbb{P}[X = \ell] = \exp\left(-e^{-c}\right) \frac{\left(e^{-c}\right)^{\ell}}{\ell!}$$

In particular,

$$\mathbb{P}\left[G(n,p) \text{ contains no isolated vertex }\right] \xrightarrow{n\to\infty} \exp\left(-e^{-c}\right)$$

Proof sketch

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$

Then $X = \sum_{v \in [n]} X_v$ denotes the number of isolated vertices.

Assume $p = \frac{\log n + c(n)}{n}$ with $c(n) \to c$. Then

$$\mathbb{E}[X] = (1 + o(1)) \exp\big(-c(n)\big) \xrightarrow{n \to \infty} e^{-c}$$

Proof sketch

For each $v \in [n]$, let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated in } G(n,p) \\ 0 & \text{otherwise.} \end{cases}$$

Then $X = \sum_{v \in [n]} X_v$ denotes the number of isolated vertices.

Assume $p = \frac{\log n + c(n)}{n}$ with $c(n) \to c$. Then

$$\mathbb{E}[X] = (1 + o(1)) \exp\left(-c(n)\right) \xrightarrow{n\to\infty} e^{-c}$$

For each $k \geq 2$,

$$\begin{aligned} \mathbb{E}\binom{X}{k} &= \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \mathbb{P}\left[X_{i_1} = 1, X_{i_2} = 1, \ldots, X_{i_k} = 1\right] \\ &= \binom{n}{k} (1-p)^{k(n-k) + \binom{k}{2}} = \ldots \xrightarrow{n\to\infty} \frac{\left(e^{-c}\right)^k}{k!} \end{aligned}$$

By the method of moments, we have $X \xrightarrow{D} \mathrm{Po}(e^{-c})$

Phase transition in $\#$ isolated vertices in $G(n,p)$

Theorem

Let

$$p \;=\; \frac{\log n \;+\; c(n)}{n}.$$

Then

$$\mathbb{P}\big[\, G(n,p) \text{ contains no isolated vertex } \big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

Sharp threshold for connectedness

A sharp threshold for the property that $G(n,p)$ is connected is

$$p^* = \frac{\log n}{n}$$

Theorem [Erdős–Rényi 59; Stepanov 69]

Let

$$p = \frac{\log n + c(n)}{n}.$$

Then

$$\mathbb{P}\big[\, G(n,p) \text{ is connected} \,\big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

Sharp threshold for connectedness

A sharp threshold for the property that $G(n,p)$ is connected is

$$p^* = \frac{\log n}{n}$$

Theorem [ERDŐS–RÉNYI 59; STEPANOV 69]

Let

$$p = \frac{\log n + c(n)}{n}.$$

Then

$$\mathbb{P}\big[\, G(n,p) \text{ is connected} \,\big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

$*$ If $G(n,p)$ contains an isolated vertex, then it is not connected.

But, the converse is not true.

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n, p)$.

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}\mathbb{P}\big[C_1 \geq 1\,\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \\ &\leq \mathbb{P}\big[\,G(n,p) \text{ is not connected}\,\big]\end{aligned}$$

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}
\mathbb{P}\big[C_1 \geq 1\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \\
&\leq \mathbb{P}\big[\, G(n,p) \text{ is not connected} \,\big] \\
&\leq \mathbb{P}\big[C_1 \geq 1\big] + \sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\big]
\end{aligned}$$

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}
\mathbb{P}\big[C_1 \geq 1\,\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \\
&\leq \mathbb{P}\big[\,G(n,p) \text{ is not connected}\,\big] \\
&\leq \mathbb{P}\big[C_1 \geq 1\,\big] + \sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\,\big]
\end{aligned}$$

For $k \geq 2$, a component of order k contains a tree of order k

$$\mathbb{P}\big[C_k \geq 1\,\big] \leq \mathbb{E}\big[C_k\big] \leq \binom{n}{k} k^{k-2} p^{k-1}$$

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}\mathbb{P}\big[C_1 \geq 1\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \\ &\leq \mathbb{P}\big[\, G(n,p) \;\text{ is not connected}\,\big] \\ &\leq \mathbb{P}\big[C_1 \geq 1\big] + \sum_{2\leq k\leq n/2} \mathbb{P}\big[C_k \geq 1\big]\end{aligned}$$

For $k \geq 2$, a component of order k contains a tree of order k

whose vertices are joined to no vertex outside the tree, and thus

$$\mathbb{P}\big[C_k \geq 1\big] \leq \mathbb{E}\big[C_k\big] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}.$$

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}\mathbb{P}\big[C_1 \geq 1\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big]\\ &\leq \mathbb{P}\big[\ G(n,p) \ \text{ is not connected}\ \big]\\ &\leq \mathbb{P}\big[C_1 \geq 1\big] + \sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\big]\end{aligned}$$

For $k \geq 2$, a component of order k contains a tree of order k

whose vertices are joined to no vertex outside the tree, and thus

$$\mathbb{P}\big[C_k \geq 1\big] \leq \mathbb{E}\big[C_k\big] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}.$$

If $p \sim \frac{\log n}{n}$, then

$$\sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\big] = \sum_{2 \leq k \leq n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} = \ldots = o(1)$$

Proof sketch

For $k \in [n]$ let C_k denote $\#$ components of order k in $G(n,p)$.

$$\begin{aligned}\mathbb{P}\big[C_1 \geq 1\big] &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\big] \\ &\leq \mathbb{P}\big[\, G(n,p) \text{ is not connected}\,\big] \\ &\leq \mathbb{P}\big[C_1 \geq 1\big] + \sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\big]\end{aligned}$$

For $k \geq 2$, a component of order k contains a tree of order k

whose vertices are joined to no vertex outside the tree, and thus

$$\mathbb{P}\big[C_k \geq 1\big] \leq \mathbb{E}\big[C_k\big] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}.$$

If $p \sim \frac{\log n}{n}$, then

$$\sum_{2 \leq k \leq n/2} \mathbb{P}\big[C_k \geq 1\big] = \sum_{2 \leq k \leq n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} = \ldots = o(1)$$

we have

$$\mathbb{P}\big[\, G(n,p) \text{ is not connected}\,\big] = \mathbb{P}\big[C_1 \geq 1\big] + o(1)$$

Proof sketch – cont'd

Summing up, If $p \sim \frac{\log n}{n}$, then we have

$$\begin{aligned}\mathbb{P}\big[\ G(n,p) \text{ is not connected}\ \big] &= \mathbb{P}\big[C_1 \geq 1\ \big] + o(1)\\ &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\ \big] + o(1)\end{aligned}$$

Proof sketch – cont'd

Summing up, If $p \sim \frac{\log n}{n}$, then we have

$$\begin{aligned}\mathbb{P}\left[\ G(n,p) \text{ is not connected}\ \right] &= \mathbb{P}\left[C_1 \geq 1\ \right] + o(1) \\ &= \mathbb{P}\left[G(n,p) \text{ contains an isolated vertex}\ \right] + o(1)\end{aligned}$$

Thus

$$\begin{aligned}&\mathbb{P}\left[\ G(n,p) \text{ is connected}\ \right] \\ &\qquad = \mathbb{P}\left[G(n,p) \text{ contains no isolated vertex}\ \right] + o(1) \\ &\qquad \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}\end{aligned}$$

$\square$

Proof sketch – cont'd

Summing up, If $p \sim \frac{\log n}{n}$, then we have

$$\begin{aligned}\mathbb{P}\big[\ G(n,p) \text{ is not connected}\ \big] &= \mathbb{P}\big[C_1 \geq 1\ \big] + o(1) \\ &= \mathbb{P}\big[G(n,p) \text{ contains an isolated vertex}\ \big] + o(1)\end{aligned}$$

Thus

$$\begin{aligned}&\mathbb{P}\big[\ G(n,p) \text{ is connected}\ \big] \\ &\qquad = \mathbb{P}\big[G(n,p) \text{ contains no isolated vertex}\ \big] + o(1) \\ &\qquad \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}\end{aligned}$$

$\square$

$*$ Hitting time result: [Bollobás–Thomason 83]

the last 'minimum obstruction' for connectedness is an isolated vertex

Outline of the minicourse – Day 2

I. Prelude

II. Basic probabilistic tools

III. Erdős-Rényi random graphs

IV. Higher-dimensional analogues

V. Random subgraphs of the hypercube

VI. Topological aspects of random graphs

Part II.

Erdős-Rényi random graphs

- ✓ Thresholds
- ✓ Connectedness threshold
- Percolation threshold
- Coupling with Galton-Watson branching trees
- More about the giant component

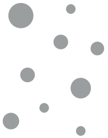
Emergence of giant component

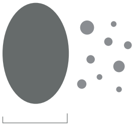
$d = p\,(n-1)$

$|L_1|$ = # vertices in the largest component in $G(n,p)$

Theorem [Erdős–Rényi 60]

- If $d < 1$, whp $|L_1| = O(\log n)$
- If $d > 1$, whp $|L_1| = \Theta(n)$





* whp = with high probability = with prob tending to one as $n \to \infty$

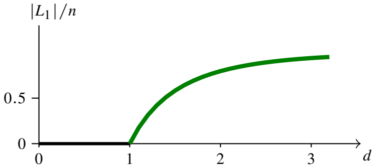
Giant component

$d = p\,(n-1)$ and $\rho \in (0,1)$ with $1-\rho = \exp(-d\,\rho)$

$|L_i| = \#$ vertices in the i-th largest component in $G(n,p)$

Theorem

- If $d < 1$, whp $\quad |L_1| \leq \frac{3}{(d-1)^2} \log n$
- If $d > 1$, whp $\quad |L_1| = (\rho + o(1))\, n$ and $|L_2| \leq \frac{20}{(d-1)^2} \log n$



Component exploration process via BFS

[KARP 1991]

- Given a vertex v,

  construct a spanning tree T_v

  by exploring the component C_v that contains v

  using Breadth-First Search

Component exploration process via BFS

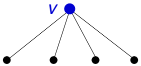
[KARP 1991]

- Given a vertex v,

  construct a spanning tree T_v

  by exploring the component C_v that contains v

  using Breadth-First Search



- $\#$ children of v $\sim$ $\mathrm{Bi}(n-1, p)$

Component exploration process via BFS

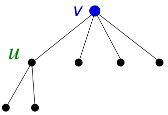
[KARP 1991]

- Given a vertex v,

  construct a spanning tree T_v

  by exploring the component C_v that contains v

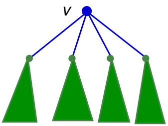
  using Breadth-First Search



- $\#$ children of $v \sim \mathrm{Bi}(n-1, p)$
- $\#$ children of $u \sim \mathrm{Bi}(n-5, p)$

Coupling with Galton-Watson trees

$T(N,p) \;=\;$ Galton-Watson tree with offspring distribution $\mathrm{Bi}(N,p)$



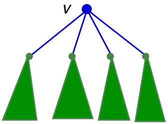
- start with a single vertex v
- number of children of each vertex

  is an i.i.d random variable

  with distribution $\mathrm{Bi}(N,p)$

Coupling with Galton-Watson trees

$T(N,p) \ = \ $ Galton-Watson tree with offspring distribution $\mathrm{Bi}(N,p)$



- start with a single vertex v
- number of children of each vertex is an i.i.d random variable with distribution $\mathrm{Bi}(N,p)$

- Upper coupling:

  couple $T(n,p)$ and a spanning tree T_v so that $T_v \subset T(n,p)$

- Lower coupling:

  couple $T(n-k,p)$ and a tree T'_v such that

  either $\min\{|T(n-k,p)|,\ |T'_v|\} \geq k$ or $T(n-k,p) \subset T'_v$

Proof sketch – Galton-Watson tree

$(Z_1, Z_2, \ldots)$ = a sequence of i.i.d. random variables with

$Z_t \sim \mathrm{Bi}(n-1, p) \sim \mathrm{Po}(d)$ = number of vertices born at time t

Proof sketch – Galton-Watson tree

$(Z_1, Z_2, \ldots)$ = a sequence of i.i.d. random variables with

$Z_t \sim \mathrm{Bi}(n-1, p) \sim \mathrm{Po}(d)$ = number of vertices born at time t

$Y_0 = 1$ and $Y_t := Y_{t-1} + Z_t - 1$ = queue size at time $t \geq 1$

T = minimal integer $t \geq 1$ with $Y_t = 0$ = total size of GW tree $T(n, p)$

Proof sketch – Galton-Watson tree

$(Z_1, Z_2, \ldots) \;=\;$ a sequence of i.i.d. random variables with

$Z_t \;\sim\; \mathrm{Bi}(n-1,p) \;\sim\; \mathrm{Po}(d) \quad = \;$ number of vertices born at time t

$Y_0 \;=\; 1$ and $Y_t \;:=\; Y_{t-1} + Z_t - 1 \quad = \;$ queue size at time $t \geq 1$

$T \;=\;$ minimal integer $t \geq 1$ with $Y_t = 0 \quad = \;$ total size of GW tree $T(n,p)$

Note that for any $\ell \geq 1$

- $Y_\ell \;=\; 1 - \ell + \sum_{t=1}^{\ell} Z_t$
- if $|T(n,p)| > \ell$, equiv, $T > \ell$, then $Y_\ell > 0$ and so $X_\ell := \sum_{t=1}^{\ell} Z_t \geq \ell - 1$
- since $X_\ell = \sum_{t=1}^{\ell} Z_t \;\sim\; \mathrm{Po}(\ell\, d)$, we have

$$\mathbb{E}(X_\ell) = \ell\, d \quad \text{and} \quad \ell = \mathbb{E}(X_\ell) + \ell(1-d) - 1$$

Proof sketch – subcritical regime

Assume $d < 1$.

By applying Chernoff bounds, we obtain, for any $\ell \geq 1$

$$\mathbb{P}[X_\ell \geq \ell - 1] = \mathbb{P}[X_\ell \geq \mathbb{E}(X_\ell) + \ell(1-d) - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

and

$$\mathbb{P}[|C_v| \geq \ell] = \mathbb{P}[|T_v| \geq \ell] \leq \mathbb{P}[|T(n,p)| \geq \ell] = \mathbb{P}[T \geq \ell]$$
$$\leq \mathbb{P}[X_\ell \geq \ell - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

Proof sketch – subcritical regime

Assume $d < 1$.

By applying Chernoff bounds, we obtain, for any $\ell \geq 1$

$$\mathbb{P}[X_\ell \geq \ell - 1] = \mathbb{P}[X_\ell \geq \mathbb{E}(X_\ell) + \ell(1-d) - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

and

$$\mathbb{P}[|C_v| \geq \ell] = \mathbb{P}[|T_v| \geq \ell] \leq \mathbb{P}[|T(n,p)| \geq \ell] = \mathbb{P}[T \geq \ell]$$
$$\leq \mathbb{P}[X_\ell \geq \ell - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

Taking $\ell = \frac{3}{(1-d)^2}\log n$ we have

$$\mathbb{P}\left[|C_v| \geq \frac{3}{(1-d)^2}\log n\right] \leq \exp\left(-\frac{(1-d)^2}{2} \cdot \frac{3}{(1-d)^2}\log n\right) = n^{-3/2}$$

Proof sketch – subcritical regime

Assume $d < 1$.

By applying Chernoff bounds, we obtain, for any $\ell \geq 1$

$$\mathbb{P}[X_\ell \geq \ell - 1] = \mathbb{P}[X_\ell \geq \mathbb{E}(X_\ell) + \ell(1-d) - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

and
$$\mathbb{P}[|C_v| \geq \ell] = \mathbb{P}[|T_v| \geq \ell] \leq \mathbb{P}[|T(n,p)| \geq \ell] = \mathbb{P}[T \geq \ell]$$
$$\leq \mathbb{P}[X_\ell \geq \ell - 1] \leq \exp\left(-\frac{(1-d)^2}{2}\ell\right)$$

Taking $\ell = \frac{3}{(1-d)^2}\log n$ we have

$$\mathbb{P}\left[|C_v| \geq \frac{3}{(1-d)^2}\log n\right] \leq \exp\left(-\frac{(1-d)^2}{2} \cdot \frac{3}{(1-d)^2}\log n\right) = n^{-3/2}$$

and thus

$$\mathbb{P}\left[G(n,p) \text{ contains a component of order } \geq \frac{3}{(1-d)^2}\log n\right]$$
$$\leq \sum_{v\in[n]} \cdot \mathbb{P}\left[|C_v| \geq \frac{3}{(1-d)^2}\log n\right] \leq n \cdot n^{-3/2} = o(1)$$

Proof sketch – supercritical regime

Assume $d > 1$.

(1) No middle ground:

whp $\nexists$ component of order between $k_* := \frac{20}{(1-d)^2} \log n$ and $k^* := n^{2/3}$

(using lower/upper couplings and Chernoff bounds)

Proof sketch – supercritical regime

Assume $d > 1$.

(1) No middle ground:

whp $\nexists$ component of order between $k_* := \frac{20}{(1-d)^2} \log n$ and $k^* := n^{2/3}$

(using lower/upper couplings and Chernoff bounds)

(2) First moment argument

– let $L :=$ total number of vertices in 'large' components,

each containing $\geq n^{2/3}$ vertices

– lower/upper couplings: $\mathbb{E}(L) \sim \rho n$

Proof sketch – supercritical regime

Assume $d > 1$.

(1) No middle ground:

whp $\nexists$ component of order between $k_* := \frac{20}{(1-d)^2} \log n$ and $k^* := n^{2/3}$

(using lower/upper couplings and Chernoff bounds)

(2) First moment argument

– let $L :=$ total number of vertices in 'large' components,

each containing $\geq n^{2/3}$ vertices

– lower/upper couplings: $\mathbb{E}(L) \sim \rho n$

(3) Second moment argument : $\mathrm{Var}[L] \ll (\mathbb{E}(L))^2$

By (2) and (3) we have $L \sim \mathbb{E}(L) \sim \rho n$

Proof sketch – supercritical regime

Assume $d > 1$.

(1) No middle ground:

whp $\nexists$ component of order between $k_* := \frac{20}{(1-d)^2} \log n$ and $k^* := n^{2/3}$

(using lower/upper couplings and Chernoff bounds)

(2) First moment argument

– let $L :=$ total number of vertices in 'large' components,

each containing $\geq n^{2/3}$ vertices

– lower/upper couplings: $\mathbb{E}(L) \sim \rho n$

(3) Second moment argument : $\mathrm{Var}\left[L\right] \ll (\mathbb{E}(L))^2$

By (2) and (3) we have $L \sim \mathbb{E}(L) \sim \rho n$

(4) Sprinkling argument:

almost all vertices in 'large' components lie in a single giant component

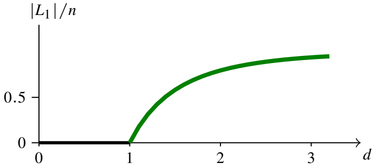
giant component

$d = p\,(n-1)$ and $\rho = 1 - \exp(-d\,\rho)$

$|L_i| = \#$ vertices in the i-th largest component in $G(n,p)$

Theorem

- If $d < 1$, whp $\quad |L_1| \leq \frac{3}{(d-1)^2} \log n$
- If $d > 1$, whp $\quad |L_1| = (\rho + o(1))\, n$ and $|L_2| \leq \frac{20}{(d-1)^2} \log n$



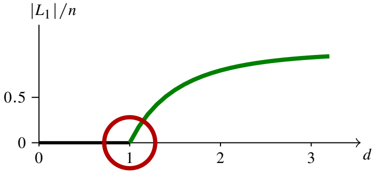
More on the giant component

$d = p\,(n-1)$ and $\rho = 1 - \exp(-d\,\rho)$

$|L_i| = \#$ vertices in the i-th largest component in $G(n,p)$

Theorem

- If $d < 1$, whp $\quad |L_1| \leq \frac{3}{(d-1)^2} \log n$
- If $d > 1$, whp $\quad |L_1| = (\rho + o(1))\,n$ and $|L_2| \leq \frac{20}{(d-1)^2} \log n$



How does the component structure look when $d \to 1$? [Bollobás 84]

Double jump ?

$d \;=\; p\,(n-1)$

$|L_1| \;=\;$ # vertices in the largest component in $G(n,p)$

Theorem [Erdős–Rényi 60]

- If $d < 1$, whp $\quad |L_1| \;=\; O(\log n)$
- If $d \to 1$, whp $\quad |L_1| \;=\; \Theta(n^{2/3})$
- If $d > 1$, whp $\quad |L_1| \;=\; \Theta(n)$

Smooth transition as the giant emerges

$d \; = \; p\,(n-1) \; = \; 1+\varepsilon$

$\varepsilon = \varepsilon(n)$ with $n^{-1/3} \ll |\varepsilon| \ll 1$

$|L_1| \; = \; \#$ vertices in the largest component in $G(n,p)$

[Bollobás 84; Łuczak 90; Bollobás–Riordan 12]

- If $\varepsilon < 0$, then whp $\quad |L_1| \; \sim \; \frac{2}{\varepsilon^2}\,\log\left(|\varepsilon|^3\,n\right)$
- If $\varepsilon > 0$, then whp $\quad |L_1| \; \sim \; 2\varepsilon\,n$

Smooth transition as the giant emerges

$d \ = \ p\,(n-1) \ = \ 1+\varepsilon$

$\varepsilon = \varepsilon(n)$ with $n^{-1/3} \ll |\varepsilon| \ll 1$

$|L_1| \ = \ \#$ vertices in the largest component in $G(n,p)$

[Bollobás 84; Łuczak 90; Bollobás–Riordan 12]

- If $\varepsilon < 0$, then whp $\qquad |L_1| \ \sim \ \frac{2}{\varepsilon^2}\log\left(|\varepsilon|^3\, n\right)$
- If $\varepsilon > 0$, then whp $\qquad |L_1| \ \sim \ 2\varepsilon\, n$

$*$ In the critical regime when $\varepsilon = O(n^{-1/3})$ [Aldous1997]

$|L_i| \ = \ \#$ vertices in the i-th largest component in $G(n,p)$

$\left(|L_i|n^{-2/3}\right)_{i\geq 1} \ \to$ (lengths of excursions of reflecting Brownian motion)

Asymptotic normality of the giant component

Assume $d = p\,(n-1) > 1$ and $1-\rho = \exp(-d\,\rho)$

$|L_1| = \#$ vertices in the largest component in $G(n,p)$

Let $\mu := \rho\, n$ and $\sigma := \sqrt{\frac{\rho(1-\rho)}{(1-d(1-\rho))^2}\, n}$

Central limit theorem

Let $N(0,1)$ denote the standard normal distribution. Then

$$\frac{|L_1| - \mu}{\sigma} \xrightarrow{D} N(0,1)$$

for $1 \ll (d-1)^3\, n \ll \frac{\log n}{\log\log n}$ [Karoński–Łuczak 02]

for constant d [Behrisch–Coja-Oghlan–K. 09]

for $d = d(n) \to 1$ and $(d-1)^3\, n \gg 1$ [Bollobás–Riordan 12]

Limit theorems for the giant

Assume $d = p\,(n-1) > 1$ and $1-\rho = \exp(-d\,\rho)$

$|L_1| = \#$ vertices in the largest component in $G(n,p)$

Let $\mu := \rho\, n$ and $\sigma := \sqrt{\frac{\rho(1-\rho)}{(1-d(1-\rho))^2}\, n}$

$\#$ vertices in L_1 [STEPANOV 70; PITTEL–WORMALD 05; BEHRISCH–COJA-OGHLAN–K. 09]

For any integer k with $k = \rho n + x$ where $x = O(\sqrt{n}\,) = O(\sigma)$

$$\mathbb{P}\big[\,|L_1| = k\,\big] \quad \sim \quad \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Limit theorems for the giant

Assume $d = p\,(n-1) > 1$ and $1-\rho = \exp(-d\,\rho)$

$|L_1| = \#$ vertices in the largest component in $G(n,p)$

Let $\mu := \rho\, n$ and $\sigma := \sqrt{\frac{\rho(1-\rho)}{(1-d(1-\rho))^2}\, n}$

$\#$ vertices in L_1 [Stepanov 70; Pittel–Wormald 05; Behrisch–Coja-Oghlan–K. 09]

For any integer k with $k = \rho n + x$ where $x = O(\sqrt{n}\,) = O(\sigma)$

$$\mathbb{P}\big[\,|L_1| = k\,\big] \quad \sim \quad \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

$\#$ vertices and $\#$ edges in L_1

Joint distribution for the number of vertices and edges

for constant d [Behrisch–Coja-Oghlan–K. 14]

for $d = d(n) \to 1$ and $(d-1)^3\, n \to \infty$ [Bollobás–Riordan 18]

Part IV.

High-dimensional analogues

- Random k-uniform hypergraphs
  - Sharp threshold for high-order connectedness
  - High-order giant component

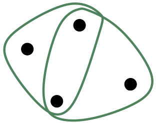
- Random k-dimensional simplicial complexes
  - Sharp threshold for cohomologically connectedness

k-uniform hypergraphs

Given $k \in \mathbb{N}_{\geq 2}$, a k-uniform hypergraph is a pair $H = (V, E)$ of

- a set V of vertices and
- a set $E \subset \binom{V}{k}$ of hyperedges,

  i.e., each hyperedge is a k-(element sub)set of vertex set V



$*$ 2-uniform hypergraph is a graph

$*$ the notion of a component in hypergraphs is ambiguous

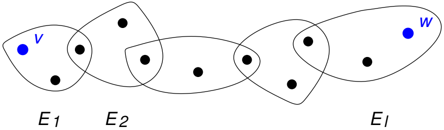
Classical notion of components

- Given a k-uniform hypergraph H,

  a vertex v is said to be reachable from another vertex w

  if there is a sequence $E_1, \ldots, E_\ell$ of hyperedges such that

  $v \in E_1$, $w \in E_\ell$ and $|E_i \cap E_{i+1}| \geq 1$ for each $i = 1, \ldots, \ell - 1$.



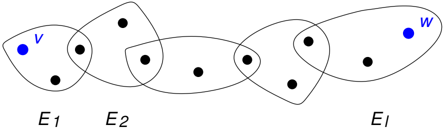
Classical notion of components

- Given a k-uniform hypergraph H,

  a vertex v is said to be reachable from another vertex w

  if there is a sequence $E_1, \ldots, E_\ell$ of hyperedges such that

  $v \in E_1$, $w \in E_\ell$ and $|E_i \cap E_{i+1}| \geq 1$ for each $i = 1, \ldots, \ell - 1$.



- The reachability is an equivalence relation on vertices, and

  the equivalence classes are called components of H.

- If H consists of a single component, it is connected.

High-order components

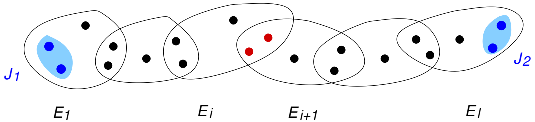
[Bollobás–Riordan 12]

- Given a k-uniform hypergraph H and $1 \le j \le k-1$,

  a j-(element sub)set J_1 is reachable from another j-set J_2 if

  $\exists\ E_1, \ldots, E_\ell$ of edges such that $J_1 \subseteq E_1$, $J_2 \subseteq E_\ell$ and

$$|E_i \cap E_{i+1}| \ge j, \quad i \in [\ell - 1].$$



High-order components

[BOLLOBÁS–RIORDAN 12]

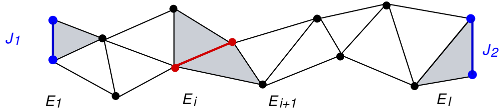
- Given a k-uniform hypergraph H and $1 \leq j \leq k-1$,

  a j-(element sub)set J_1 is reachable from another j-set J_2 if

  $\exists\ E_1, \ldots, E_\ell$ of edges such that $J_1 \subseteq E_1$, $J_2 \subseteq E_\ell$ and

$$|E_i \cap E_{i+1}| \geq j, \quad i \in [\ell - 1].$$

  e.g., $k = 3,\ j = 2$



- Reachability is an equivalence relation on j-sets, and equivalence classes are called j-(tuple)component.

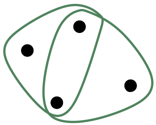
- If H consists of a single j-component, it is j-connected.

Random binomial k-uniform hypergraphs

$H_k(n, p)$ denotes a random binomial k-uniform hypergraph

- on vertex set $[n] := \{1, 2, \ldots, n\}$,
- in which each k-(element sub)set of vertex set $[n]$ is

  an hyperedge with probability p, independently

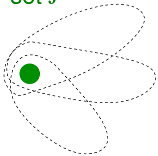


$*$ note $H_2(n, p) \ = \ G(n, p)$

Number of isolated j-sets in $H_k(n,p)$

For each j-set $J \in \binom{[n]}{j}$, let $X_J = \begin{cases} 1 & \text{if } J \text{ is isolated in } H_k(n,p) \\ 0 & \text{otherwise.} \end{cases}$

Then $X = \sum_{J \in \binom{[n]}{j}} X_J$ counts # isolated j-sets and



$$\mathbb{E}[X] = \sum_{J \in \binom{[n]}{j}} \mathbb{E}[X_J] = \binom{n}{j} (1-p)^{\binom{n-j}{k-j}}$$

$$\sim \frac{1}{j!} \exp\left(j \log n - p \binom{n-j}{k-j} \right)$$

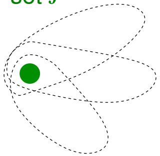
Number of isolated j-sets in $H_k(n,p)$

For each j-set $J \in \binom{[n]}{j}$, let $X_J = \begin{cases} 1 & \text{if } J \text{ is isolated in } H_k(n,p) \\ 0 & \text{otherwise.} \end{cases}$

Then $X = \sum_{J \in \binom{[n]}{j}} X_J$ counts $\#$ isolated j-sets and

$$\mathbb{E}[X] = \sum_{J \in \binom{[n]}{j}} \mathbb{E}[X_J] = \binom{n}{j} (1-p)^{\binom{n-j}{k-j}}$$

$$\sim \frac{1}{j!} \exp\left(j \log n - p \binom{n-j}{k-j} \right)$$



If $p = \frac{j \log n + c(n)}{\binom{n-j}{k-j}}$, then

$$\mathbb{E}[X] \sim \frac{1}{j!} \exp\left(-c(n) \right) \quad \to \quad \begin{cases} \infty & \text{if } c(n) \to -\infty \\ \frac{1}{j!} e^{-c} & \text{if } c(n) \to c \in \mathbb{R} \\ 0 & \text{if } c(n) \to \infty \end{cases}$$

Sharp threshold for j-connectedness

Theorem [COOLEY–K.–KOCH 16]

Given $k \in \mathbb{N}_{\geq 2}$ and $1 \leq j \leq k-1$, let $p = \frac{j \log n + c(n)}{\binom{n-j}{k-j}}$. Then

$$\mathbb{P}\big[\, H_k(n,p) \text{ is } j\text{-connected} \,\big] \xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-\frac{1}{j!}\, e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

* $j = 1$ [POOLE 15]
* $j = k-1$ [KAHLE –PITTEL 16]
* Proof methods
  - last ‘minimum obstruction’ for j-connectedness is an isolated j-set
  - $\#$ j-set of degree s when $p = \frac{j \log n + s \log\log n + c(n)}{\binom{n-j}{k-j}}$ for $s \in \mathbb{N} \cup \{0\}$
  - $\exists$ j-component containing a reasonably large subset which is smooth

Heuristics for percolation threshold

Breadth-First Search & Galton-Watson tree

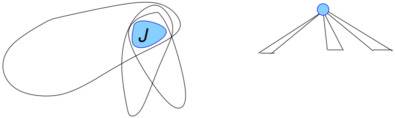
Heuristics for percolation threshold

Breadth-First Search & Galton-Watson tree

- Begin with a j-set J

Heuristics for percolation threshold

Breadth-First Search & Galton-Watson tree

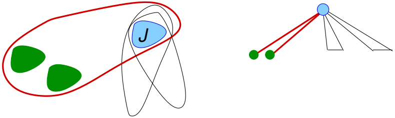


- Begin with a j-set J
- Discover all edges that contain that j-set J

  $\exists \binom{n-j}{k-j}$ such edges containing J, each with prob. p

Heuristics for percolation threshold

Breadth-First Search & Galton-Watson tree



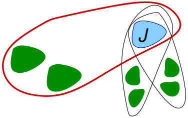
- Begin with a j-set J
- Discover all edges that contain that j-set J

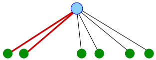
  $\exists$ $\binom{n-j}{k-j}$ such edges containing J, each with prob. p
- For each edge E containing J,

  discover $(\binom{k}{j} - 1)$ new j-sets in E

Heuristics for percolation threshold

Breadth-First Search & Galton-Watson tree





- Begin with a j-set J
- Discover all edges that contain that j-set J

  $\exists$ $\binom{n-j}{k-j}$ such edges containing J, each with prob. p
- For each edge E containing J,

  discover $(\binom{k}{j} - 1)$ new j-sets in E

$$\mathbb{E}[\ \#\ j\text{-sets discovered from } J\ ] \;=\; \left(\binom{k}{j} - 1\right)\binom{n-j}{k-j}\, p \;=:\; d$$

Giant j-component

Given $k \in \mathbb{N}_{\geq 2}$ and $1 \leq j \leq k-1$, let $d = \left(\binom{k}{j} - 1\right) \binom{n-j}{k-j} p$

Assume $\varepsilon = d - 1$ satisfy $\varepsilon \to 0,\ |\varepsilon|^3 n^j \gg 1, \ldots$

$|L^{(j)}| = \#$ j-sets in largest j-component in $H_k(n,p)$

Theorem [Cooley–K.–Koch 18; Cooley–Fang–Del Giudice–K. 19]

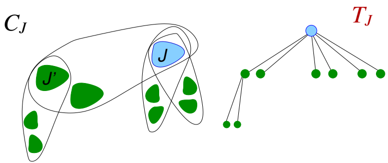
- If $\varepsilon < 0$, whp $\quad |L^{(j)}| \sim \frac{2\left(\binom{k}{j}-1\right)}{\varepsilon^2} \log\left(|\varepsilon|^3 \binom{n}{j}\right)$
- If $\varepsilon > 0$, whp $\quad |L^{(j)}| \sim \frac{2\varepsilon}{\binom{k}{j}-1} \binom{n}{j}$

The simplest case when $j = 1$

- for constant ε [Schmidt-Pruzan–Schamir 85]
- for $\varepsilon > 0$ and $1 \ll \varepsilon^3 n \ll \frac{\log n}{\log \log n}$ [Karoński–Łuczak 02]
- for $\varepsilon > 0$, $\varepsilon \to 0$ and $\varepsilon^3 n \gg 1$ [Bollobás–Riordan 14]

Proof sketch when $\varepsilon > 0$

(1) Breadth-First Search



Given j-set J

construct a spanning tree T_J

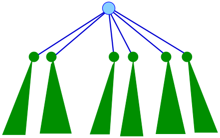
(representing j-component C_J)

consisting of j-sets as vertices

(2) Couple T_J and

Galton-Watson branching tree in which each vertex has

$\left(\binom{k}{j} - 1\right) \cdot \mathrm{Bi}\left(\binom{n-j}{k-j}, p\right)$ many children independently



$\rho := \mathbb{P}\,[\text{ process survives }]$

$$1 - \rho = \sum_\ell \mathbb{P}\left[\mathrm{Bi}\left(\binom{n-j}{k-j}, p\right) = \ell\right] \cdot (1-\rho)^{\ell\,\left(\binom{k}{j}-1\right)}$$

$$\longrightarrow \quad \rho \sim \frac{2\varepsilon}{\binom{k}{j}-1}$$

Proof sketch – cont'd

(3) First moment argument:

- $L := \#\, j$-sets in 'large' j-components,

  each containing $\geq n^{2j/3}$ many j-sets

- using upper and lower couplings with Galton-Watson trees,

$$\mathbb{E}[L] \sim \frac{2\varepsilon}{\binom{k}{j} - 1} \binom{n}{j}$$

(4) Second moment argument:

IF

$$\mathrm{Var}[L] \ll (\mathbb{E}(L))^2,$$

THEN

$$L \sim \mathbb{E}[L] \sim \frac{2\varepsilon}{\binom{k}{j} - 1} \binom{n}{j}$$

(5) Sprinkling argument:

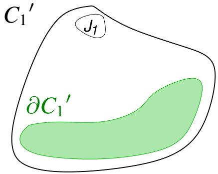
almost all j-sets in 'large' j-components are in a single j-component

More on second moment argument

Need to consider pairs of j-sets in ’large’ j-components

More on second moment argument

Need to consider pairs of j-sets in 'large' j-components

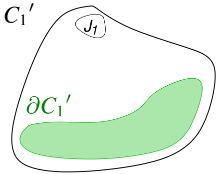


- Fix j-set J_1 and grow its j-component $C_1{}'$ until hit stopping conditions

$$S_1 \;=\; \{\; |C_1{}'| \;\geq\; n^{2j/3} \quad \text{or} \quad |\partial C_1{}'| \;\geq\; \epsilon\, n^{2j/3} \;\}$$

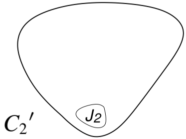
More on second moment argument

Need to consider pairs of j-sets in 'large' j-components



- Fix j-set J_1 and grow its j-component $C_1{}'$ until hit stopping conditions

$$S_1 = \{\ |C_1{}'| \geq n^{2j/3} \quad \text{or} \quad |\partial C_1{}'| \geq \epsilon\, n^{2j/3}\ \}$$

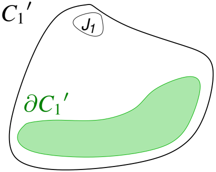


- Delete all the vertices in $C_1{}'$

& fix a j-set J_2, grow component $C_2{}'$

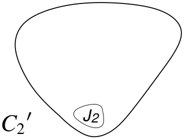
More on second moment argument

Need to consider pairs of j-sets in 'large' j-components



- Fix j-set J_1 and grow its j-component $C_1{}'$ until hit stopping conditions

$$S_1 = \{ \, |C_1{}'| \geq n^{2j/3} \quad \text{or} \quad |\partial C_1{}'| \geq \epsilon \, n^{2j/3} \, \}$$



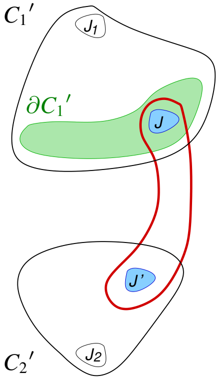
- Delete all the vertices in $C_1{}'$

& fix a j-set J_2, grow component $C_2{}'$

- Need to show $\mathbb{P}\left(e(\partial C_1{}', \; C_2{}') \geq 1 \right)$ is small

More on second moment argument

Need to consider pairs of j-sets in 'large' j-components



- Fix j-set J_1 and grow its j-component C_1' until hit stopping conditions

$$S_1 = \{ |C_1'| \geq n^{2j/3} \quad \text{or} \quad |\partial C_1'| \geq \epsilon n^{2j/3} \}$$

- Delete all the vertices in C_1'

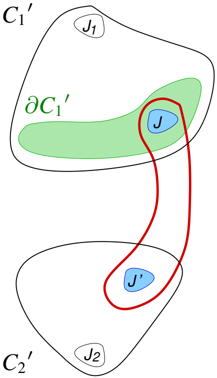
& fix a j-set J_2, grow component C_2'

- Need to show $\mathbb{P}(\, e(\partial C_1',\, C_2') \geq 1)$ is small

Naive approach: $\mathbb{P}(\, e(\partial C_1',\, C_2') \geq 1) \leq p \cdot |\partial C_1'| \cdot |C_2'|$

More on second moment argument

Need to consider pairs of j-sets in 'large' j-components



- Fix j-set J_1 and grow its j-component C_1' until hit stopping conditions

$$S_1 = \{ |C_1'| \geq n^{2j/3} \quad \text{or} \quad |\partial C_1'| \geq \epsilon n^{2j/3} \}$$

- Delete all the vertices in C_1'

& fix a j-set J_2, grow component C_2'

- Need to show $\mathbb{P}(\, e(\partial C_1',\, C_2') \geq 1)$ is small

Naive approach: $\mathbb{P}(\, e(\partial C_1',\, C_2') \geq 1) \leq p \cdot |\partial C_1'| \cdot |C_2'|$ too big !!

More on second moment argument – cont.

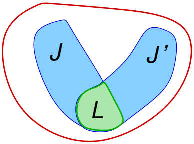
Instead we need

- $\mathbb{P}\,(\ e(\partial C_1{}',\ C_2{}') \geq 1)$

  $\leq\ \mathbb{E}\,(\#\ k$-sets containing

  a pair of j-sets, J, J', intersecting at an ℓ-set L

  for some $0 \leq \ell \leq j-1)$



More on second moment argument – cont.

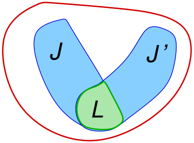
Instead we need

- $\mathbb{P}\,(\;e(\partial C_1{}',\;C_2{}') \geq 1)$

  $\leq\;\mathbb{E}\,(\#\;k$-sets containing

  a pair of j-sets, J, J', intersecting at an ℓ-set L

  for some $0 \leq \ell \leq j-1)$



- smooth boundary lemma: 'reasonably large' boundary is smooth

Summary and open problems

Given $k \in \mathbb{N}_{\geq 2}$ and $1 \leq j \leq k-1$, let $d = \left(\binom{k}{j} - 1\right)\binom{n-j}{k-j} p$

Assume $\varepsilon = d - 1$ satisfy $\varepsilon \to 0,\ |\varepsilon|^3 n^j \gg 1, \ldots$

$|L^{(j)}| = \#$ j-sets in largest j-component in $H_k(n,p)$

Theorem [Cooley–K.–Koch 18; Cooley–Fang–Del Giudice–K. 19]

- If $\varepsilon < 0$, whp $\quad |L^{(j)}| \sim \frac{2\left(\binom{k}{j}-1\right)}{\varepsilon^2} \log\left(|\varepsilon|^3 \binom{n}{j}\right)$
- If $\varepsilon > 0$, whp $\quad |L^{(j)}| \sim \frac{2\varepsilon}{\binom{k}{j}-1} \binom{n}{j}$

$Q1.$ Width of critical window

- $\varepsilon = O(n^{-1/3})$ in $G(n,p)$
- $\varepsilon = O(n^{-j/3})$ in $H_k(n,p)$?

$Q2.$ Asymptotic normality of the giant j-component

Summary and open problems – cont'd

$Q3.$ Structural symmetry

- Fragment $R = G(n,p) \setminus L_1$

  behaves like subcritical $G(n',p')$ with $d' = n'p' < 1$

Summary and open problems – cont'd

$Q3.$ Structural symmetry

- Fragment $R = G(n,p) \setminus L_1$

  behaves like subcritical $G(n',p')$ with $d' = n'p' < 1$

- Does the fragment $R = H_k(n,p) \setminus L^{(j)}$

  behave like subcritical $H_k(n',p')$ with $d' = \left(\binom{k}{j} - 1\right)\binom{n'-j}{k-j}p' < 1$?

- What does subcritical $H_k(n,p)$ look like?

  What about hypertrees, high-order 'cycles', etc

Part IV.

High-dimensional analogues

- ✓ Random k-uniform hypergraphs
  - ✓ Sharp threshold for high-order connectedness
  - ✓ High-order giant component

- Random k-dimensional simplicial complexes
  - Sharp threshold for cohomologically connectedness

Simplicial complexes

- Given a set V,

  a family X of subsets of V is called a simplicial complex if

  - $\{x\} \in X, \quad \forall x \in V$
  - X is downward-closed, i.e., if $A \in X, \emptyset \neq B \subset A$, then $B \in X$

- Given a simplicial complex X,

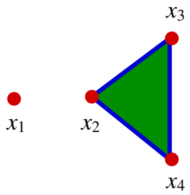
  - $A \in X$ is called an ℓ-simplex if $|A| = \ell + 1$
  - X is said to be k-kimensional if it contains no $(k+1)$-simplex

For example, given $V = \{x_1, x_2, x_3, x_4\}$, let

$$S_0 = \Big\{ \{x_1\}, \{x_2\}, \{x_3\}, \{x_4\} \Big\}$$

$$S_1 = \Big\{ \{x_2, x_3\}, \{x_2, x_4\}, \{x_3, x_4\} \Big\}$$

$$S_2 = \Big\{ \{x_2, x_3, x_4\} \Big\}$$



Then $X = S_0 \cup S_1 \cup S_2$ is a 2-dimensional simplex

Random k-dimensional simplicial complexes

They arise from random binomial $(k+1)$-uniform hypergraph $H_p = ([n], E_p)$

- 0-simplices are singletons of $[n]$
- k-simplices are the hyperedges of H_p

Random k-dimensional simplicial complexes

They arise from random binomial $(k+1)$-uniform hypergraph $H_p = ([\,n\,], E_p)$

- 0-simplices are singletons of $[\,n\,]$
- k-simplices are the hyperedges of H_p

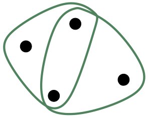
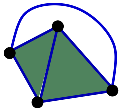
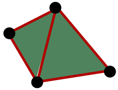
(a) the full $(k-1)$-skeleton on $[\,n\,]$ is included

$$\Delta_p := \binom{[\,n\,]}{1} \cup \binom{[\,n\,]}{2} \cup \ldots \cup \binom{[\,n\,]}{k} \cup E_p$$

(b) $\forall \ell \in [\,k-1\,]$,

ℓ-simplices are $(\ell+1)$-subsets contained in the hyperedges of H_p

$$\mathcal{G}_p := \binom{[\,n\,]}{1} \cup \ \ldots \cup \partial(\partial E_p) \cup \partial E_p \cup E_p$$

$\mathbb{F}_2$-Cohomologically connectedness

$\Delta_p := \binom{[n]}{1} \cup \binom{[n]}{2} \cup \ldots \cup \binom{[n]}{k} \cup E_p$ is said to be

$\mathbb{F}_2$-cohomologically connected if $H^{k-1}(\Delta_p;\ \mathbb{F}_2) = 0$

Theorem [Linial–Meshulam 06; Meshulam–Wallach 09; Kahle–Pittel 16]

Let

$$p = \frac{k \log n \ + \ c(n)}{n}.$$

Then

$$\mathbb{P}\Big(\ \Delta_p \text{ is } \mathbb{F}_2\text{-cohomologically connected}\ \Big)$$

$$\xrightarrow{n \to \infty} \begin{cases} 0 & \text{if } \ c(n) \ \to \ -\infty \\ \exp\left(-\frac{1}{k!}\, e^{-c}\right) & \text{if } \ c(n) \ \to \ c \in \mathbb{R} \\ 1 & \text{if } \ c(n) \ \to \ \infty \end{cases}$$

$*$ last 'minimum obstruction' is an isolated $(k-1)$-simplex

e.g., an isolated 1-simplex ($=$ an isolated edge) when $k = 2$

$\mathbb{F}_2$-Cohomologically j-connectedness

$k \in \mathbb{N}_{\geq 2}$ and $1 \leq j \leq k-1$

$\mathcal{G}_p := \binom{[n]}{1} \cup \ldots \cup \partial(\partial E_p) \cup \partial E_p \cup E_p$ is said to be

$\mathbb{F}_2$-cohomologically j-connected if $H^i(\mathcal{G}_p; \mathbb{F}_2) = 0, \quad \forall\, i \in [j]$.

Theorem [Cooley–del Giudice–K.–Sprüssel 20]

Let

$$p = \frac{(j+1)\log n + \log\log n + c(n)}{(k-j+1)\binom{n}{k-j}}.$$

Then

$$\mathbb{P}\Big(\mathcal{G}_p \text{ is } \mathbb{F}_2\text{-cohomologically } j\text{-connected} \Big)$$

$$\xrightarrow{n\to\infty} \begin{cases} 0 & \text{if } c(n) \to -\infty \\ \exp\left(-\frac{(j+1)}{(k-j+1)^2\, j!}\, e^{-c}\right) & \text{if } c(n) \to c \in \mathbb{R} \\ 1 & \text{if } c(n) \to \infty \end{cases}$$

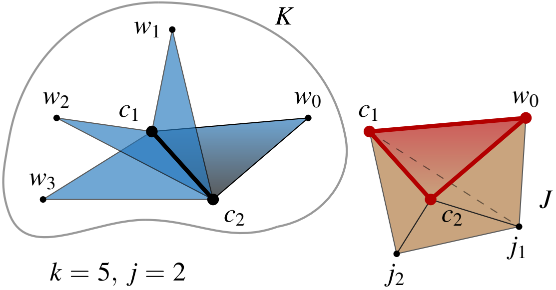
Minimal obstruction

M_j = triple (K, C, J) where K is a k-simplex in $\mathcal{G}_p$ and

- C = $(j-1)$-simplex in K such that for each $w \in K \setminus C$,

  j-simplex $C \cup \{w\}$ is contained in no other k-simplex of $\mathcal{G}_p$

- J = set of j-simplices such that
  - every $(j-1)$-simplex is in even number of j-simplices in J
  - it contains exactly one $C \cup \{w_0\}$, $w_0 \in K \setminus C$



Outline of the minicourse – Day 3

I. Prelude

II. Basic probabilistic tools

III. Erdős-Rényi random graphs

IV. Higher-dimensional analogues

V. Random subgraphs of the hypercube

VI. Topological aspects of random graphs

Part V.

Random subgraphs of the hypercube

- Random subgraphs
- Emergence of the giant component
- Expansion properties of the giant component
- Consequences of expansion properties

The hypercube

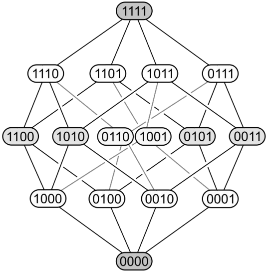
Given $d \in \mathbb{N}$, the d-dimensional hypercube Q^d is the graph with

- vertex set

$$V\left(Q^d\right) = \{0,1\}^d = \left\{x = (x_1, \ldots, x_d) \;:\; x_i \in \{0,1\},\ 1 \le i \le d\right\}$$

- edge set $E\left(Q^d\right)$: $\forall\, x = (x_1, \ldots, x_d),\ y = (y_1, \ldots, y_d) \in V\left(Q^d\right)$,

$\{x, y\} \in E\left(Q^d\right)$ iff x and y differ in exactly one coordinate



Hasse diagram

The hypercube

Given $d \in \mathbb{N}$, the d-dimensional hypercube Q^d is the graph with

- vertex set

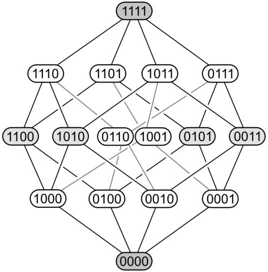
$$V\left(Q^d\right) \;=\; \{0,1\}^d \;=\; \left\{x = (x_1, \ldots, x_d) \;:\; x_i \in \{0,1\},\; 1 \le i \le d\right\}$$

- edge set $E\left(Q^d\right)$: $\forall\, x = (x_1, \ldots, x_d),\, y = (y_1, \ldots, y_d) \in V\left(Q^d\right)$,

$$\{x, y\} \in E\left(Q^d\right) \quad \text{iff} \quad x \text{ and } y \text{ differ in exactly one coordinate}$$

Basic facts:

- $|V\left(Q^d\right)| \;=\; 2^d$
- Q^d is d-regular
- Q^d is bipartite
- diameter of Q^d is d
- . . .



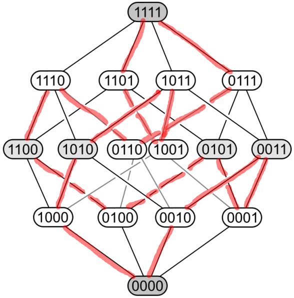
Hasse diagram

A random subgraph of the hypercube

Given $p \in (0,1)$

Q^d_p $=$ a graph obtained by retaining each edge of Q^d independently with probability p

$=$ bond percolation on Q^d with edge probability p



Typical properties of Q_p^d around $p = \frac{1}{2}$

Connectivity [Sapoženko 67; Burtin 67; Erdős–Spencer 79; Bollobás 83]

$p = \frac{1}{2}$ is a sharp threshold for connectedness: $\forall \varepsilon > 0$

$$\mathbb{P}\left[Q_p^d \text{ is connected} \right] \xrightarrow{d \to \infty} \begin{cases} 0 & \text{if} \quad p < \frac{1-\varepsilon}{2} \\ 1 & \text{if} \quad p > \frac{1+\varepsilon}{2} \end{cases}$$

Typical properties of Q_p^d around $p = \frac{1}{2}$

Connectivity [Sapoženko 67; Burtin 67; Erdős–Spencer 79; Bollobás 83]

$p = \frac{1}{2}$ is a sharp threshold for connectedness: $\forall \varepsilon > 0$

$$\mathbb{P}\left[Q_p^d \text{ is connected}\right] \xrightarrow{d\to\infty} \begin{cases} 0 & \text{if} \quad p < \frac{1-\varepsilon}{2} \\ 1 & \text{if} \quad p > \frac{1+\varepsilon}{2} \end{cases}$$

Perfect matching [Bollobás 90]

$p = \frac{1}{2}$ is a sharp threshold for the existence of a perfect matching

Is $p = \frac{1}{2}$ a sharp threshold for Hamiltonicity? [Bollobás 80's; Frieze 14]

Hamiltonicity [Condon-Espuny-Díaz-Girão-Kühn-Osthus 21]

$p = \frac{1}{2}$ is a sharp threshold the existence of a Hamiltonian cycle.

Emergence of the giant component in Q_p^d

[Erdős–Spencer 79]

Does the component structure of Q_p^d undergo a phase transition at $p = \frac{1}{d}$?

Giant component [Ajtai-Komlós- Szemerédi 81]

$p = \frac{1}{d}$ is a sharp threshold: $\forall \varepsilon > 0$

- whp all components are of order $O(d)$ if $p < \frac{1-\varepsilon}{d}$
- whp $\exists$ a unique largest component of order $\Theta(2^d)$ if $p > \frac{1+\varepsilon}{d}$

Phase transitions in Q_p^d

- More detailed component structure for a wider range of p

[Bollobás-Kohayakawa-Łuczak 92]

[Borgs–Chayes–van der Hofstad–Slade–Spencer 2006]

[Hulshof–Nachmias 2020]

[McDiarmid–Scott–Withers 2021]

- Correct width of critical window [van der Hofstad–Nachmias 2017]

- Diameter of components ? [Bollobás-Kohayakawa-Łuczak 92]

[Heydenreich–van der Hofstad 2011]

[van der Hofstad–Nachmias 2014]

[Hulshof–Nachmias 2020]

Two open problems on the giant component

$L_1 \;=\;$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

$Q1.$ What is the diameter of L_1 ? [Bollobás-Kohayakawa-Łuczak 92]

$Q2.$ What is the mixing time of the lazy simple random walk on L_1 ?

[Pete 08; van der Hofstad-Nachmias 17]

Diameter and mixing time

$L_1 \;=\;$ largest component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem [Erde-K.-Krivelevich 22]

whp the diameter of L_1 is $O\left(d^3\right)$

whp the mixing time of the lazy simple random walk on L_1 is $O\left(d^{11}\right)$

Diameter and mixing time

$L_1 \;=\;$ largest component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem [ERDE-K.-KRIVELEVICH 22]

whp the diameter of L_1 is $O\left(d^3\right)$

whp the mixing time of the lazy simple random walk on L_1 is $O\left(d^{11}\right)$

Theorem [ERDE-K.-KRIVELEVICH 22]

whp L_1

- is $c\, d^{-5}-$expander
- contains a $c'\, d^{-2}\,(\log d)^{-1}-$expander on $\geq 0.99\,|L_1|$ vertices
- has Cheeger constant $\Omega\left(d^{-5}\right)$

Expanders

[ALON 86; HOORY-LINIAL-WIGDERSON 06; KRIVELEVICH 19; KRIVELEVICH-SUDAKOV 09; SARNAK 04; . . .]

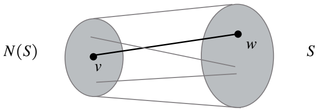
Given a graph G

- $N(S)$ = external neighbourhood of a subset $S \subseteq V(G)$

  $= \{v \in V(G) \setminus S : \exists\, w \in S \text{ with } \{v,w\} \in E(G)\}$

- G is an α−expander if

$$|N(S)| \geq \alpha |S|, \qquad \forall\, S \subseteq V(G) \quad \text{with} \quad |S| \leq \frac{|V(G)|}{2}$$



Expanders

[ALON 86; HOORY-LINIAL-WIGDERSON 06; KRIVELEVICH 19; KRIVELEVICH-SUDAKOV 09; SARNAK 04; . . .]

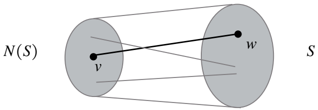
Given a graph G

- $N(S) =$ external neighbourhood of a subset $S \subseteq V(G)$

  $= \{v \in V(G) \setminus S : \exists\, w \in S$ with $\{v, w\} \in E(G)\}$

- G is an $\alpha-$expander if

$$|N(S)| \geq \alpha |S|, \qquad \forall\, S \subseteq V(G) \quad \text{with} \quad |S| \leq \frac{|V(G)|}{2}$$



Properties of an expander

- small diameter, long paths/cycles, large complete minor, . . .
- edge-expansion for graphs with bounded max degree

Proof sketch

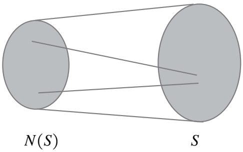
Theorem

whp L_1 is a $\frac{1}{\text{poly}(d)}$ $-$expander

We will prove that whp

for an arbitrary subset S of $V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$,

$$|N(S)| \geq \frac{|S|}{\text{poly}(d)}$$



Sprinkling argument

Sprinkling

$$p = \frac{1+\varepsilon}{d} \text{ for } \varepsilon > 0$$

$$q_1 = \frac{1+\delta_1}{d} \text{ and } q_2 = \frac{\delta_2}{d} \text{ s.t. } 1-p = (1-q_1)(1-q_2) \text{ and } 0 < \delta_2 \ll \delta_1$$

$$Q_p^d \sim Q_{q_1}^d \cup Q_{q_2}^d$$

Sprinkling argument

Sprinkling

$$p = \frac{1+\varepsilon}{d} \text{ for } \varepsilon > 0$$

$$q_1 = \frac{1+\delta_1}{d} \text{ and } q_2 = \frac{\delta_2}{d} \text{ s.t. } 1-p = (1-q_1)(1-q_2) \text{ and } 0 < \delta_2 \ll \delta_1$$

$$Q_p^d \sim Q_{q_1}^d \cup Q_{q_2}^d$$

Largest components before and after sprinkling

$$L_1' = \text{largest component in } Q_{q_1}^d \quad \text{(before sprinkling)}$$

$$L_1 = \text{largest component in } Q_p^d \quad \text{(after sprinkling)}$$

$$\gamma(x) = \text{survival probability of } \mathrm{Po}(1+x) \text{ branching process}$$

Lemma [Ajtai-Komlós- Szemerédi 81]

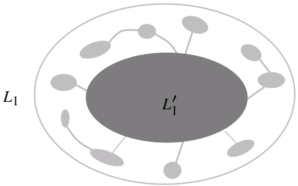
- whp $L_1' \sim \gamma(\delta_1)\, 2^d$
- whp $L_1 \sim \gamma(\epsilon)\, 2^d$

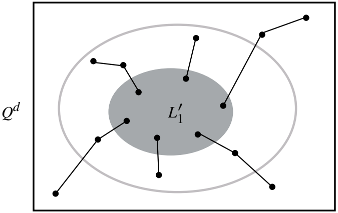
Giant component before and after sprinkling

$$L_1' \;=\; \text{largest component in } Q^d_{q_1} \quad \text{(before sprinkling)}$$
$$L_1 \;=\; \text{largest component in } Q^d_p \quad \text{(after sprinkling)}$$

Lemma

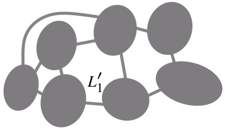
- whp $\forall$ connected component in $Q^d_p\,[L_1 - L_1']$ is of order $O(d)$
- whp $\forall$ vertex in $V(Q^d)$ is within distance two from $\geq c\,d^2$ vertices in L_1'





Splitting the largest component into pieces

$$\begin{aligned} L_1' &= \text{largest component} \quad \text{(before sprinkling)} \\ &= \text{split into a family } \mathcal{C} \text{ of vertex-disjoint connected subgraphs} \\ &\quad \text{('pieces'), each of order poly}(d) \end{aligned}$$

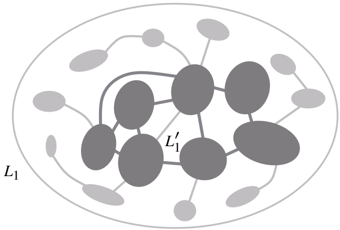


Splitting the largest component into pieces

L'_1 = largest component (before sprinkling)

= split into a family $\mathcal{C}$ of vertex-disjoint connected subgraphs ('pieces'), each of order poly(d)

L_1 = largest component (after sprinkling)



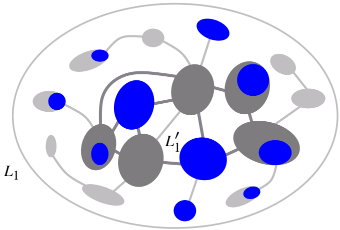
Splitting the largest component into pieces

L_1' $=$ largest component (before sprinkling)

$=$ split into a family $\mathcal{C}$ of vertex-disjoint connected subgraphs ('pieces'), each of order poly(d)

L_1 $=$ largest component (after sprinkling)

S $=$ arbitrary subset of $V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$



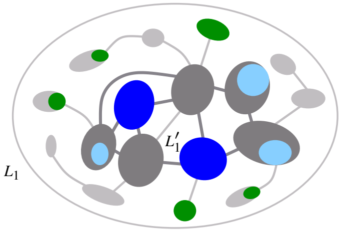
Splitting the largest component into pieces

L_1' = largest component (before sprinkling)

= split into a family $\mathcal{C}$ of vertex-disjoint connected subgraphs ('pieces'), each of order poly(d)

L_1 = largest component (after sprinkling)

S = arbitrary subset of $V(L_1)$ with $|S| \leq \frac{|V(L_1)|}{2}$



$S_1 = S - L_1'$

S_2 = vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

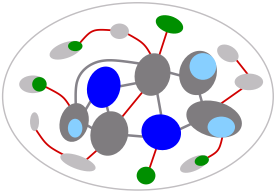
S_3 = vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$

Contribution of S_1 to $N(S)$

$$S_1 = S - L_1'$$

$$S_2 = \text{vertices in pieces } C \in \mathcal{C} \text{ with } C \cap S \neq \emptyset \text{ and } S - C \neq \emptyset$$

$$S_3 = \text{vertices in pieces } C \in \mathcal{C} \text{ with } C \subseteq S$$



With sprinkling, each component in $Q_p^d[L_1 - L_1']$ which intersects with S_1

- contributes at least one edge to $N(S)$
- or is connected to $S_2 \cup S_3$

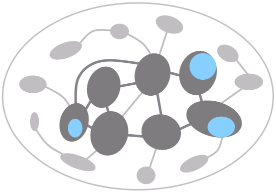
Thus $|N(S)| \geq \frac{c|S_1|}{d}$ or $e(S_1, S_2 \cup S_3) \geq \frac{c|S_1|}{d}$ and thus $|S_2 \cup S_3| \geq \frac{c|S_1|}{d^2}$

Contribution of S_2 to $N(S)$

$$L_1' \;=\; \text{split into a family } \mathcal{C} \text{ of pieces, each of order poly}(d)$$

$$S_2 \;=\; \text{vertices in pieces } C \in \mathcal{C} \text{ with } C \cap S \neq \emptyset \text{ and } S - C \neq \emptyset$$

$$S_3 \;=\; \text{vertices in pieces } C \in \mathcal{C} \text{ with } C \subseteq S$$



Each piece $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

- contributes at least one edge to $N(S)$

and each piece is of order poly(d)

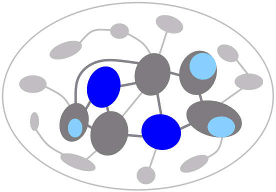
Thus $\quad |N(S)| \geq \frac{|S_2|}{\text{poly}(d)}$

Contribution of S_3 to $N(S)$

$L_1' \;=\;$ split into a family $\mathcal{C}$ of pieces, each of order $\mathrm{poly}(d)$

$S_2 \;=\;$ vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

$S_3 \;=\;$ vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$



(1) Partition the family $\mathcal{C}$ of pieces into two disjoint families $\{\mathcal{A}, \mathcal{B}\}$

$$\mathcal{A} := \{C \in \mathcal{C} \colon C \subseteq S\} \quad \text{and} \quad \mathcal{B} := \mathcal{C} - \mathcal{A}$$

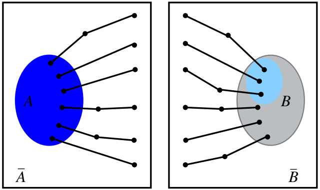
This partitions $V(L_1')$ into two sets A, B where

$$A := V(\mathcal{A}) = S_3 \quad \text{and} \quad B := V(\mathcal{C} - \mathcal{A})$$

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L_1') = A \,\dot\cup\, B$ to a partition $V(Q^d) = \bar{A} \,\dot\cup\, \bar{B}$ s.t.

- every vertex in $\bar{A}$ is within distance 2 of A
- every vertex in $\bar{B}$ is within distance 2 of B

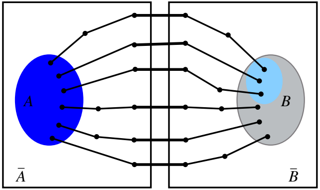


whp every vertex in $V(Q^d)$ is within distance two from $\geq c\, d^2$ vertices in L_1'

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L_1') = A \,\dot\cup\, B$ to a partition $V(Q^d) = \bar{A} \,\dot\cup\, \bar{B}$ s.t.

- every vertex in $\bar{A}$ is within distance 2 of A
- every vertex in $\bar{B}$ is within distance 2 of B



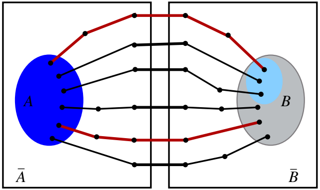
Edge-isoperimetry in Q^d [Harper 64; Lindsey 64; Bernstein 67; Hart 76]

$$|E(X, X^c)| \geq |X|\,(d - \log_2 |X|), \qquad \forall X \subseteq V(Q^d) \quad \text{with} \quad |X| \leq 2^{d-1}$$

Contribution of S_3 to $N(S)$ – extending and connecting

(2) Extending the partition $V(L_1') = A \dot\cup B$ to a partition $V(Q^d) = \bar{A} \dot\cup \bar{B}$ s.t.

- every vertex in $\bar{A}$ is within distance 2 of A
- every vertex in $\bar{B}$ is within distance 2 of B



(3) Sprinkle with $q_2 = \frac{\delta_2}{d}$

Lemma

whp $\exists$ at least $\frac{|A|}{\text{poly}(d)}$ vertex-disjoint A-B-paths of length at most 5 in $Q^d_{q_2}$

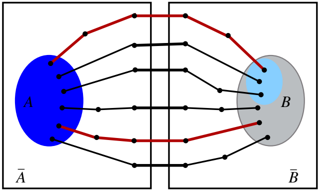
Contribution of S_3 to $N(S)$

L_1' $=$ split into a family $\mathcal{C}$ of 'pieces', each of order poly(d)

S_2 $=$ vertices in pieces $C \in \mathcal{C}$ with $C \cap S \neq \emptyset$ and $S - C \neq \emptyset$

S_3 $=$ vertices in pieces $C \in \mathcal{C}$ with $C \subseteq S$

$= A$



Each A-B-path in $Q_{q_2}^d$ contributes at least one edge to $N(S)$,

unless it goes to S_2

Thus $$|N(S)| \geq \frac{|S_3|}{\text{poly}(d)} - d|S_2|$$

Expansion properties and consequences

$L_1 \;=\;$ largest component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem [ERDE-K.-KRIVELEVICH 22]

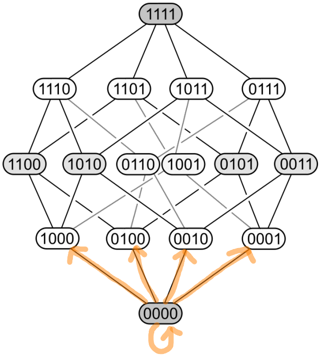
whp L_1

- is $c\, d^{-5}$–expander
- contains a $c'\, d^{-2}\, (\log d)^{-1}$–expander on $\geq 0.99\, |L_1|$ vertices
- has diameter $O\left(d^3\right)$
- has Cheeger constant $\Omega\left(d^{-5}\right)$

Mixing time of lazy random walk on Q^d

In each step,

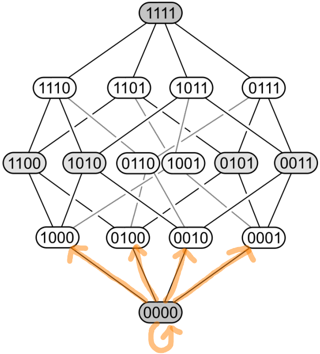
- it remains at the current position with prob $\frac{1}{2}$
- it moves to a uniformly chosen random neighbour with prob $\frac{1}{2}$



Mixing time of lazy random walk on Q^d

In each step,

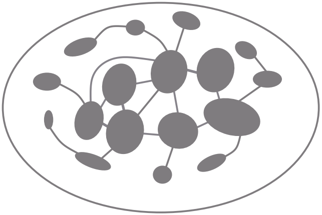
- it remains at the current position with prob $\frac{1}{2}$
- it moves to a uniformly chosen random neighbour with prob $\frac{1}{2}$



Mixing time: $O(d \log d)$

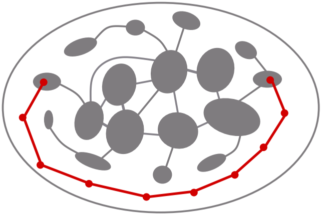
Mixing time of lazy random walk on the giant

$L_1 \;=\;$ giant component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$



Mixing time of lazy random walk on the giant

$L_1 \;=\;$ giant component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$



$*$ whp L_1 contains bare paths of length $\Omega(d)$

$\Longrightarrow$ (worst-case) mixing time : $\Omega(d^2)$

Mixing time of lazy random walk

Given a graph G,

$t_{\mathrm{mix}}(G) \;=\;$ (worst-case) mixing time of a lazy random walk on a graph G

$\Phi(G) \;=\;$ Cheeger constant of G (= bottleneck ratio)

$\pi_{\min}(G) \;=\; \min\{ \frac{d_G(x)}{2|E(G)|} \;:\; x \in V(G)\}$

[LAWLER–SOKAL 88; JERRUM–SINCLAIR 89; LEVIN–PERES–WILMER 07]

$$t_{\mathrm{mix}}(G) \;\leq\; \frac{2}{\Phi(G)^2} \log\left(\frac{4}{\pi_{\min}(G)}\right)$$

Mixing time of lazy random walk on the giant

Given a graph G,

$t_{\text{mix}}(G) \;=\;$ (worst-case) mixing time of a lazy random walk on a graph G

$\Phi(G) \;=\;$ Cheeger constant of G (= bottleneck ratio)

$\pi_{\min}(G) \;=\; \min\{\frac{d_G(x)}{2|E(G)|} \;:\; x \in V(G)\}$

[Lawler–Sokal 88; Jerrum–Sinclair 89; Levin–Peres–Wilmer 07]

$$t_{\text{mix}}(G) \;\le\; \frac{2}{\Phi(G)^2} \log\left(\frac{4}{\pi_{\min}(G)}\right)$$

$L_1 \;=\;$ giant component of Q_p^d when $p \;=\; \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

[Erde-K.-Krivelevich 22]

whp

$$\Phi(L_1) \;=\; \Omega\left(d^{-5}\right) \quad \text{and} \quad \pi_{\min}(L_1) \;=\; \Omega\left(2^{-d}\right)$$

$$t_{\text{mix}}(L_1) \;=\; O\left(d^{11}\right)$$

Summary and open problems

L_1 $=$ largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem [ERDE-K.-KRIVELEVICH 22]

whp L_1

- is $c\,d^{-5}$–expander
- contains a $c'\,d^{-2}\,(\log d)^{-1}$–expander on $\geq 0.99\,|L_1|$ vertices
- has diameter $O\left(d^3\right)$

whp the mixing time of the lazy simple random walk on L_1 is $O\left(d^{11}\right)$.

Summary and open problems

L_1 = largest component of Q_p^d when $p = \frac{1+\varepsilon}{d}$ for $\varepsilon > 0$

Theorem [ERDE-K.-KRIVELEVICH 22]

whp L_1

- is $c\, d^{-5}$–expander
- contains a $c'\, d^{-2} (\log d)^{-1}$–expander on $\geq 0.99\, |L_1|$ vertices
- has diameter $O\left(d^3\right)$

whp the mixing time of the lazy simple random walk on L_1 is $O\left(d^{11}\right)$.

$Q1.$ What is the correct order of the diameter of L_1 ?

$Q2.$ What is the mixing time of the lazy random walk on L_1 ?

Outline of the minicourse – Day 4

I. Prelude

II. Basic probabilistic tools

III. Erdős-Rényi random graphs

IV. Higher-dimensional analogues

V. Random subgraphs of the hypercube

VI. Topological aspects of random graphs

Part VI.

Topological aspects of random graphs

- Typical genus of the Erdős-Rényi random graph
- Random graphs on surfaces with constant genus
- Benjamini-Schramm local weak limits
- Random graphs on surfaces with non-constant genus

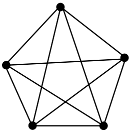
Guiding questions/themes

(1) What is a typical genus of the Erdős-Rényi random graph?

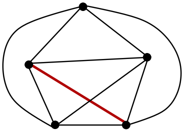
$*$ genus of a graph G

$=$ minimum number of handles that must be attached to sphere in order to embed G without any crossing edges

$*$ the case when genus is 0 corresponds to planar graphs



K_5



genus of K_5 = 1

Guiding questions/themes

(1) What is a typical genus of the Erdős-Rényi random graph?

(2) How does a topological constraint such as

- being planar
- being embeddable on the orientable surface with given genus

affect the global and local structure of a random graph, e.g.,

- component structures
- local weak limits

A uniform random graph

$G(n,m) \;\in_R\; \mathcal{G}(n,m)$

$$\begin{aligned} \mathcal{G}(n,m) &= \text{set of all vertex-labelled simple graphs} \\ &\quad \text{on vertex set } [\,n\,] := \{1,\ldots,n\} \text{ with } m = m(n) \text{ edges} \\ G(n,m) &= \text{chosen uniformly at random from } \mathcal{G}(n,m) \end{aligned}$$

$*$ $G(n,m)$ and $G(n,p)$ are 'essentially equivalent' when $m \sim \binom{n}{2}\, p$

Planarity of $G(n, m)$

$m = d \cdot \frac{n}{2}$

Theorem [Erdős-Rényi 1959–60]

- If $d < 1$, whp
  - each component is either a tree or unicyclic component
  - $G(n, m)$ is planar
- If $d > 1$, whp
  - largest component contains $\geq$ two cycles ('complex')
  - $G(n, m)$ is not planar

Genus g of $G(n,m)$

$m \gg n$ (superlinear)

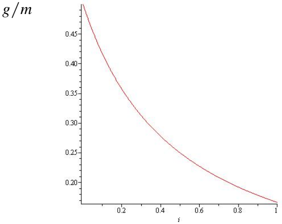
$g = g(G(n,m))$ denote the genus of $G(n,m)$

Theorem [Rödl-Thomas 1995]

- If $n^{1+\frac{1}{j+1}} \ll m \ll n^{1+\frac{1}{j}}$ for $j \in \mathbb{N}$, then whp,

$$g = (1+o(1)) \frac{j}{2(j+2)} m$$

- If $m = \Theta(n^2)$, then whp $g = (1+o(1)) \frac{1}{6} m$.



Genus g of supercritical $G(n,m)$

$m = d \cdot \frac{n}{2}$ for $d > 1$

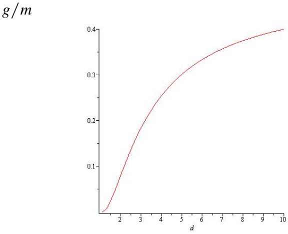
$g = g(G(n,m))$ denote the genus of $G(n,m)$

Theorem [Dowden–K.–Krivelevich 2019]

whp

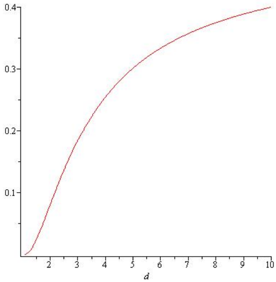
$$g = (1+o(1))\,\mu(d)\,m$$

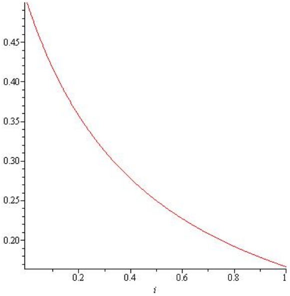
where $\mu(d) \to 0$ as $d \to 1$ and $\mu(d) \to \frac{1}{2}$ as $d \to \infty$



Summary – genus g of $G(n, m)$

$$\frac{g}{m} \sim \begin{cases} 0 & \text{if } m - n/2 \ll n \\ \mu(d) & \text{if } 2m/n \to \ d > 1 \quad \left(\lim_{d\to 1} \mu(d) = 0,\ \lim_{d\to\infty} \mu(d) = \frac{1}{2}\right) \\ \frac{j}{2(j+2)} & \text{if } n^{1+\frac{1}{j+1}} \ll m \ll n^{1+\frac{1}{j}} \ \ \textit{for} \ \ j \in \mathbb{N} \\ \frac{1}{6} & \text{if } m = \Theta\left(n^2\right) \end{cases}$$





Random graphs on surfaces

$g \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$

$\mathbb{S}_g$ = the orientable surface of genus g

$S_g(n,m) \in_R \mathcal{S}_g(n,m)$

$\mathcal{S}_g(n,m)$ = set of all vertex-labelled simple graphs on $[n]$ with $m = m(n)$ edges that are embeddable on $\mathbb{S}_g$

$S_g(n,m)$ = chosen uniformly at random from $\mathcal{S}_g(n,m)$

Random graphs on surfaces

$g \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$

$\mathbb{S}_g$ = the orientable surface of genus g

$S_g(n, m) \in_R \mathcal{S}_g(n, m)$

$\mathcal{S}_g(n, m)$ = set of all vertex-labelled simple graphs on $[n]$ with $m = m(n)$ edges that are embeddable on $\mathbb{S}_g$

$S_g(n, m)$ = chosen uniformly at random from $\mathcal{S}_g(n, m)$

Note

- $\mathcal{S}_0(n, m) \subset \ldots \subset \mathcal{S}_g(n, m) \subset \mathcal{S}_{g+1}(n, m) \subset \ldots \subset \mathcal{G}(n, m)$

Random graphs on surfaces

$g \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$

$\mathbb{S}_g$ = the orientable surface of genus g

$S_g(n,m) \in_R \mathcal{S}_g(n,m)$

$\mathcal{S}_g(n,m)$ = set of all vertex-labelled simple graphs on $[\,n\,]$ with $m = m(n)$ edges that are embeddable on $\mathbb{S}_g$

$S_g(n,m)$ = chosen uniformly at random from $\mathcal{S}_g(n,m)$

Note

- $\mathcal{S}_0(n,m) \subset \ldots \subset \mathcal{S}_g(n,m) \subset \mathcal{S}_{g+1}(n,m) \subset \ldots \subset \mathcal{G}(n,m)$
- If $1 \leq m < \frac{n}{2}$, then $\dfrac{|\mathcal{S}_0(n,m)|}{|\mathcal{G}(n,m)|} \xrightarrow{n\to\infty} 1$
- If $m > 3n - 6 + 6g$, then $\mathcal{S}_g(n,m) = \emptyset$

Random graphs on surfaces – vertex model

$S_g(n) \in_R \mathcal{S}_g(n)$

$\mathcal{S}_g(n)$ = set of all vertex-labelled simple graphs on $[n]$ that are embeddable on $\mathbb{S}_g$

$S_g(n)$ = chosen uniformly at random from $\mathcal{S}_g(n)$

Random graphs on surfaces – vertex model

$S_g(n) \in_R \mathcal{S}_g(n)$

$\mathcal{S}_g(n)$ = set of all vertex-labelled simple graphs on $[\,n\,]$ that are embeddable on $\mathbb{S}_g$

$S_g(n)$ = chosen uniformly at random from $\mathcal{S}_g(n)$

Theorem [McDiarmid-Steger-Welsh 05 ($g = 0$); McDiarmid 08]

whp the order of largest component L_1 in $S_g(n)$ is $n - O(1)$

Theorem [Giménez–Noy 09 ($g = 0$); Chapuy–Fusy–Giménez–Mohar–Noy 11]

$$|\mathcal{S}_g(n)| \quad \sim \quad \alpha_g\, n^{\frac{5}{2}g - \frac{7}{2}}\, \gamma^n\, n!$$

where $\alpha_g > 0$ and $\gamma \approx 27.23$ is the exponential growth rate of planar graphs

random graphs on surfaces

$S_g(n) \in_R \mathcal{S}_g(n)$

Theorem [Giménez–Noy 09 ($g = 0$); Chapuy–Fusy–Giménez–Mohar–Noy 11]

- $X_n = \#$ edges in $S_g(n)$

$$\frac{X_n - \mathbb{E}(X_n)}{\sigma(X_n)} \xrightarrow{\text{D}} N(0,1)$$

where $\mathbb{E}(X_n) \approx 2.21\, n$ and $\sigma^2(X_n) \approx 0.43\, n$ (same as planar case)

'Dense' random graphs on surfaces

$S_g(n) \in_R \mathcal{S}_g(n)$

Theorem [Giménez–Noy 09 ($g = 0$); Chapuy–Fusy–Giménez–Mohar–Noy 11]

- $X_n = \#$ edges in $S_g(n)$

$$\frac{X_n - \mathbb{E}(X_n)}{\sigma(X_n)} \xrightarrow{\mathrm{D}} N(0,1)$$

where $\mathbb{E}(X_n) \approx 2.21\, n$ and $\sigma^2(X_n) \approx 0.43\, n$ (same as planar case)

- For $m = d \cdot \frac{n}{2}$ with $d \in (2,6)$,

$$|\mathcal{S}_g(n,m)| \sim \alpha_g(d)\, n^{\frac{5}{2}g-4}\, \gamma(d)^n\, n!$$

where $\alpha_g(d) > 0$ and $\gamma(d)$ is same as planar case

$*$ whp $|L_1| = n - o(1)$

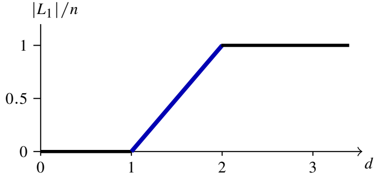
'Sparse' random graphs on surfaces

$L_1 \;=\;$ largest component in $S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$

$m \;=\; d\cdot\frac{n}{2}$ for $d \;\in\; (1,2)$

Theorem [K.–ŁUCZAK 2012 ($g=0$); K.–MOSSHAMMER–SPRÜSSEL 2020]

whp $\qquad |L_1| \;=\; (1+o(1))\,(d-1)n$



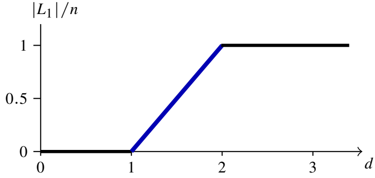
'Sparse' random graphs on surfaces

$L_1 \; = \;$ largest component in $S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$

$m \; = \; d \cdot \frac{n}{2}$ for $d \;\in\; (1,2)$

Theorem [K.–Łuczak 2012 ($g=0$); K.–Mosshammer–Sprüssel 2020]

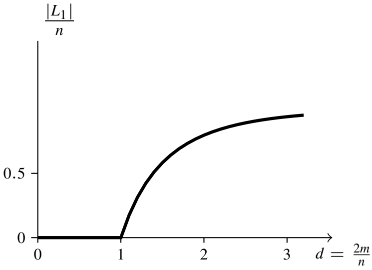
whp $\qquad |L_1| \;=\; (1+o(1))\,(d-1)n$

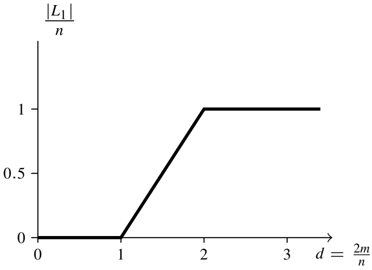


$Q1.$ What is the limit distribution of $|L_1|$?

$Q2.$ Joint distribution of the number of vertices and edges in L_1 ?

ER random graph vs random graphs on surfaces

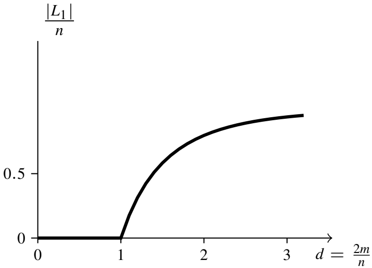




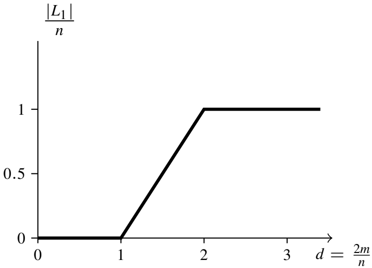
Uniform random graph $G(n, m)$

Random graph on a surface $S_g(n, m)$

ER random graph vs random graphs on surfaces



Uniform random graph $G(n, m)$



Random graph on a surface $S_g(n, m)$

$*$ fragment $R = G(n, m) \setminus L_1$ is subcritical (i.e., $2m_R/n_R < 1$)

$*$ fragment $R = S_g(n, m) \setminus L_1$ is critical (i.e., $2m_R/n_R \to 1$)

Part VI.

Topological aspects of random graphs

- ✓ Typical genus of the Erdős-Rényi random graph
- ✓ Random graphs on surfaces with constant genus
- Benjamini-Schramm local weak limits
- Random graphs on surfaces with non-constant genus

Local structure of Erdős-Rényi random graph

$G = G(n,m) \in_R \mathcal{G}(n,m)$

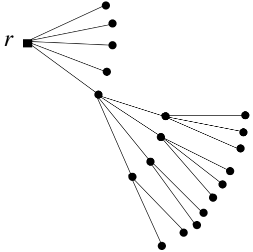
$m = d \cdot \frac{n}{2}$ for $d > 0$

Local structure of Erdős-Rényi random graph

$G \;=\; G(n,m) \;\in_R\; \mathcal{G}(n,m)$

$m \;=\; d \cdot \frac{n}{2}$ for $d \;>\; 0$

$r \;\in_R\; V(G) \;=\;$ a vertex chosen uniformly at random from $V(G)$



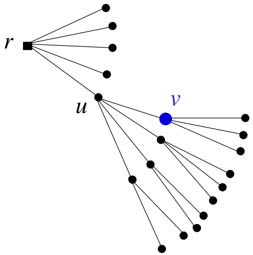
$d^{+}(r) \;\sim\; \mathrm{Po}\,(d)$

Local structure of Erdős-Rényi random graph

$G \;=\; G(n,m) \;\in_R\; \mathcal{G}(n,m)$

$m \;=\; d \cdot \frac{n}{2}$ for $d \;>\; 0$

$r \;\in_R\; V(G) \;=\;$ a vertex chosen uniformly at random from $V(G)$



$d^{+}(r) \;\sim\; \mathrm{Po}\,(d)$

$d^{+}(u) \;\sim\; \mathrm{Po}\,(d)$

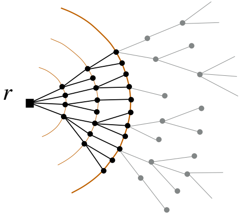
$d^{+}(v) \;\sim\; \mathrm{Po}\,(d)$

Benjamini–Schramm local weak limit

[BENJAMINI–SCHRAMM 2001; ALDOUS–STEELE 2004]

- Given a rooted graph (H, r) and $\ell \in \mathbb{N} := \{1, 2, \ldots\}$, let

$$B_\ell\,(H, r) \;:=\; H\left[\{v \in V(H) : d_H(v, r) \le \ell\}\right]$$

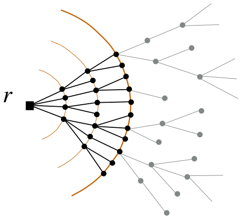


Benjamini–Schramm local weak limit

[Benjamini–Schramm 2001; Aldous–Steele 2004]

- Given a rooted graph (H,r) and $\ell \in \mathbb{N} := \{1,2,\ldots\}$, let

$$B_\ell\,(H,r) \;:=\; H\left[\{v \in V(H) : d_H(v,r) \le \ell\}\right]$$



- Two rooted graphs (H,r) and (H',r') are isomorphic,

$$(H,r) \quad \cong \quad (H',r')$$

if $\exists$ isomorphism ϕ from H onto H with $\phi(r) = r'$

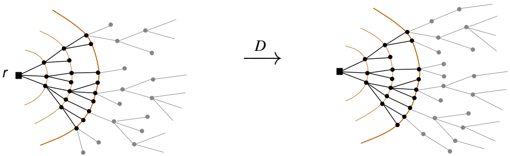
Benjamini–Schramm local weak limit — cont'd

Given a sequence $\big((G_n, r_n)\big)_n$ of random connected rooted graphs, a random connected rooted graph (G_0, r_0) is the local weak limit of (G_n, r_n)

$$(G_n, r_n) \xrightarrow{D} (G_0, r_0)$$

if for each fixed rooted graph (H, r_H) and $\ell \in \mathbb{N}$,

$$\mathbb{P}\Big[B_\ell\,(G_n, r_n) \;\cong\; (H, r_H)\Big] \xrightarrow{n\to\infty} \mathbb{P}\Big[B_\ell\,(G_0, r_0) \;\cong\; (H, r_H)\Big]$$



Benjamini–Schramm local weak limit — cont'd

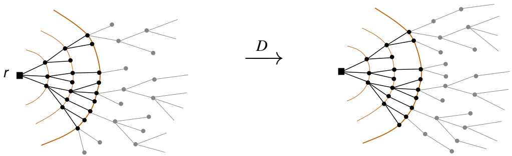
Given a sequence $\big((G_n, r_n)\big)_n$ of random connected rooted graphs,

a random connected rooted graph (G_0, r_0) is the local weak limit of (G_n, r_n)

$$(G_n, r_n) \xrightarrow{D} (G_0, r_0)$$

if for each fixed rooted graph (H, r_H) and $\ell \in \mathbb{N}$,

$$\mathbb{P}\Big[B_\ell\,(G_n, r_n) \cong (H, r_H)\Big] \xrightarrow{n\to\infty} \mathbb{P}\Big[B_\ell\,(G_0, r_0) \cong (H, r_H)\Big]$$



* For not necessarily connected (G_n, r_n), its local weak limit?

$\Longrightarrow$ define it as the local weak limit of the component of G_n containing r_n

Erdős-Rényi random graph vs GW tree

$G \;=\; G(n,m) \;\;\in_R\;\; \mathcal{G}(n,m)$

$r \;\in_R\; V(G)$

$2m/n \;\xrightarrow{n\to\infty}\; d \in [0,\infty)$

$\mathrm{GWT}(d) \;=\;$ Galton-Watson tree with offspring distribution $\mathrm{Po}(d)$

Theorem [Dembo–Montanari 2010; van der Hofstad 2022+]

$$(G,r) \quad \xrightarrow{D} \quad \mathrm{GWT}(d)$$

Local weak limit of a random tree

$T \;=\; T(n) \;\in_R\; \mathcal{T}(n)$

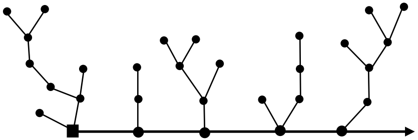
$\;=\;$ a tree chosen uniformly at random from the class of all trees on $[n]$

$r \;\in_R\; V(T)$

Theorem [GRIMMETT 1980/1981]

$$(T, r) \quad \xrightarrow{D} \quad T_\infty$$

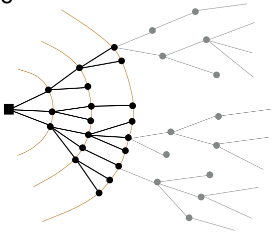
Skeleton tree $\;T_\infty$



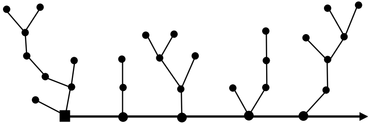
$=$ a rooted tree obtained from an infinite path by replacing each vertex of the path by an independent Galton-Watson tree $\mathrm{GWT}\,(1)$

Local weak limits

$\mathrm{GWT}\,(d)$ Galton-Watson tree



T_∞ Skeleton tree



Local weak limit of a random graph on a surface

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$

$r \;\in_R\; V(S)$ a vertex chosen uniformly at random from $V(S)$

$2m/n \;\xrightarrow{n\to\infty}\; d \in [1,2]$

Local weak limit of a random graph on a surface

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$

$r \;\in_R\; V(S)$ a vertex chosen uniformly at random from $V(S)$

$2m/n \xrightarrow{n\to\infty} d \in [1,2]$

Theorem [K.–MISSETHAN 2022+]

$$(S,r) \quad \xrightarrow{D} \quad (2-d)\ \mathrm{GWT}(1) \;+\; (d-1)\, T_\infty$$

meaning that for each rooted graph (H, r_H) and $\ell \in \mathbb{N}$, we have

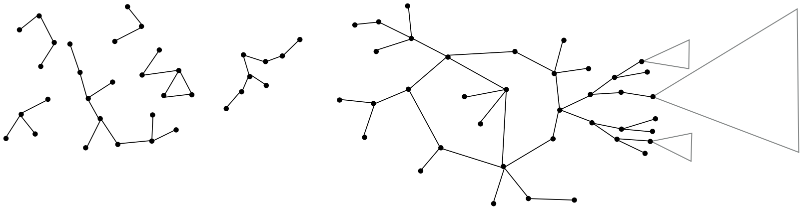
$$\mathbb{P}\Big[B_\ell(S,r) \cong (H,r_H)\Big] \quad \xrightarrow{n\to\infty}$$
$$(2-d)\,\mathbb{P}\Big[B_\ell(\mathrm{GWT}(1)) \cong (H,r_H)\Big] \;+\; (d-1)\,\mathbb{P}\Big[B_\ell(T_\infty) \cong (H,r_H)\Big]$$

Global structure of a random graph on a surface

$S = S_g(n,m) \in_R \mathcal{S}_g(n,m)$ and $2m/n \to d \in (1,2)$

L_1 largest component of S

$R = S \setminus L_1$ fragment of S



Global structure of the fragment

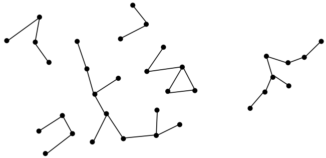
$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S

$R \;=\; S \setminus L_1$ fragment of S

The fragment R ‘behaves similarly’ like

a critical ER random graph $G(\bar{n},\bar{m})$ with $\bar{n} = (2-d)\,n$ and $2\bar{m}/\bar{n} \;\to\; 1$



Global structure of the fragment

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

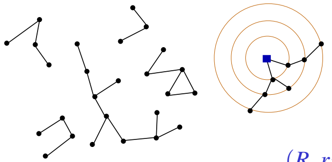
L_1 largest component of S

$R \;=\; S \setminus L_1$ fragment of S

$r_R \;\in_R\; V(R)$

The fragment R 'behaves similarly' like

a critical ER random graph $G(\bar{n},\bar{m})$ with $\bar{n} = (2-d)\,n$ and $2\bar{m}/\bar{n} \;\to\; 1$



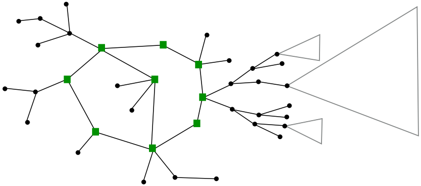
$(R, r_R) \xrightarrow{D} \mathrm{GWT}\,(1)$

Internal structure of the largest component

$S \ = \ S_g(n,m) \ \in_R \ \mathcal{S}_g(n,m)$ and $2m/n \ \to \ d \in (1,2)$

L_1 largest component of S_g

C 2-core $=$ max subgraph of L_1 with min deg $\geq$ two



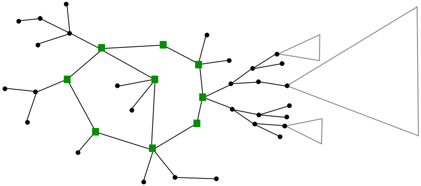
Internal structure of the largest component

$S \; = \; S_g(n,m) \; \in_R \; \mathcal{S}_g(n,m)$ and $2m/n \; \to \; d \in (1,2)$

L_1 largest component of S_g

C 2-core $=$ max subgraph of L_1 with min deg $\geq$ two

- $|L_1| \; = \; (1+o(1))\,(d-1)\,n$ and $|C| \; = \; o(n)$
- $L \; = \; C \; + $ each vertex in $V(C)$ replaced by a rooted tree



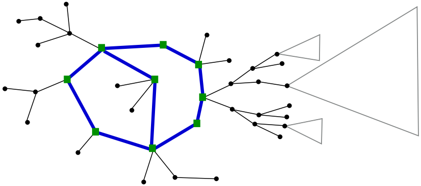
Internal structure of the largest component

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S_g

C 2-core $=$ max subgraph of L_1 with min deg $\geq$ two

- $|L_1| \;=\; (1+o(1))\,(d-1)\,n$ and $|C| \;=\; o(n)$
- $L \;=\; C \;+$ each vertex in $V(C)$ replaced by a rooted tree



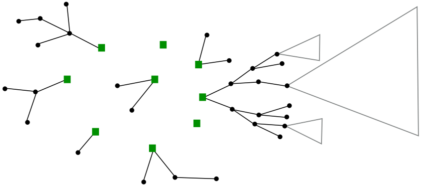
Internal structure of the largest component

$S = S_g(n,m) \in_R \mathcal{S}_g(n,m)$ and $2m/n \to d \in (1,2)$

L_1 largest component of S_g

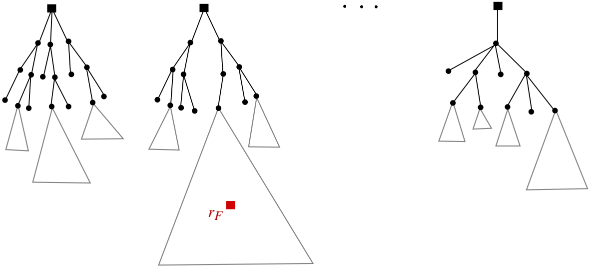
C 2-core $=$ max subgraph of L_1 with min deg $\geq$ two

- $|L_1| = (1+o(1))\,(d-1)\,n$ and $|C| = o(n)$
- $L = C$ + each vertex in $V(C)$ replaced by a rooted tree



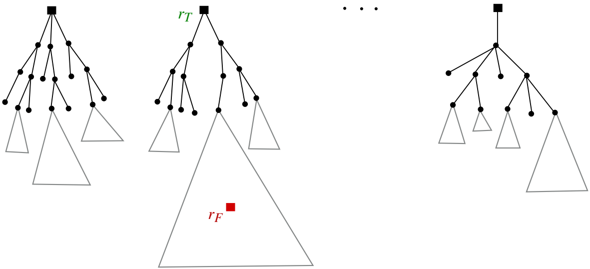
Local weak limit of a random forest

- $F = F(n,t) \in_R \mathcal{F}(n,t)$ a forest on vertex set $[n]$ with t tree components
- $r_F \in_R V(F)$ a vertex chosen uniformly at random from $V(F)$



Local weak limit of a random forest

- $F = F(n,t) \in_R \mathcal{F}(n,t)$ a forest on vertex set $[n]$ with t tree components
- $r_F \in_R V(F)$ a vertex chosen uniformly at random from $V(F)$
- r_T the root of the tree component T in F that contains r_F



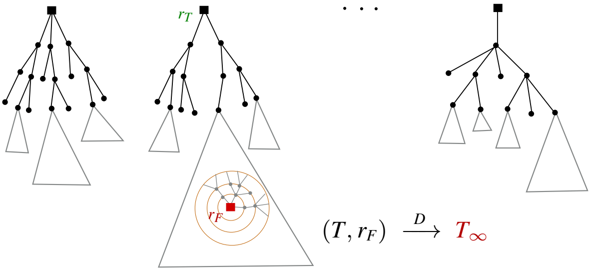
Local weak limit of a random forest

- $F = F(n,t) \in_R \mathcal{F}(n,t)$ a forest on vertex set $[n]$ with t tree components
- $r_F \in_R V(F)$ a vertex chosen uniformly at random from $V(F)$
- r_T the root of the tree component T in F that contains r_F

Lemma [K.–MISSETHAN 2022+]

If $t = t(n) = o(n)$, then whp $d_F(r_F, r_T) = \omega(1)$ and

$$(F, r_F) \xrightarrow{D} T_\infty$$



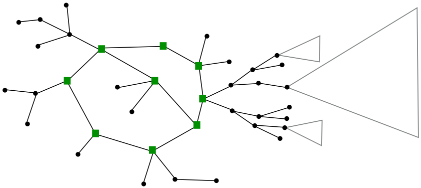
Internal structure of the largest component

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S

$R \;=\; S \setminus L_1$ fragment of S

- $|L_1| \;=\; (1+o(1))\,(d-1)\,n$ and $|C| \;=\; o(n)$
- $L_1 \;=\; C \;+\;$ each vertex in $V(C)$ replaced by a rooted tree



Internal structure of the largest component

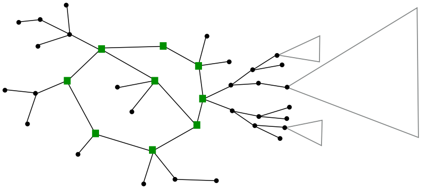
$S \ = \ S_g(n,m) \ \in_R \ \mathcal{S}_g(n,m)$ and $2m/n \ \to \ d \in (1,2)$

L_1 largest component of S

$R \ = \ S \setminus L_1$ fragment of S

$r_{L_1} \ \in_R \ V(L_1)$

- $|L_1| \ = \ (1+o(1))\,(d-1)\,n$ and $|C| \ = \ o(n)$
- $L_1 \ = \ C$ + each vertex in $V(C)$ replaced by a rooted tree



Internal structure of the largest component

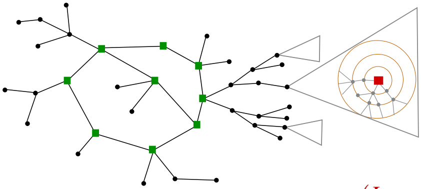
$S = S_g(n,m) \in_R \mathcal{S}_g(n,m)$ and $2m/n \to d \in (1,2)$

L_1 largest component of S

$R = S \setminus L_1$ fragment of S

$r_{L_1} \in_R V(L_1)$

- $|L_1| = (1+o(1))\,(d-1)\,n$ and $|C| = o(n)$
- $L_1 = C$ + each vertex in $V(C)$ replaced by a rooted tree



$(L_1, r_{L_1}) \xrightarrow{D} T_\infty$

Finer view of local weak limits

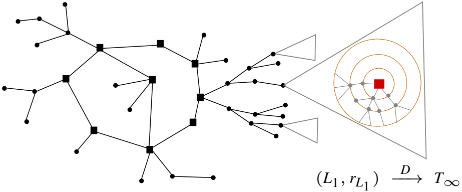
$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S and $|L_1| \;\sim\; (d-1)\,n$

$r_{L_1} \;\in_R\; V(L_1)$

Theorem [K.–MISSETHAN 2022+]

$$(L_1, r_{L_1}) \xrightarrow{D} T_\infty$$



Finer view of local weak limits

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S and $|L_1| \;\sim\; (d-1)\,n$

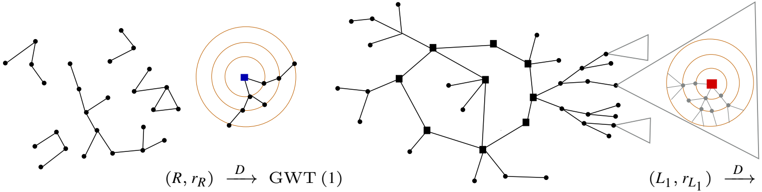
$R = S \setminus L_1 \;\sim$ crtitical ER random graph and $|R| \;\sim\; (2-d)\,n$

$r_R \;\in_R\; V(R)$, $r_{L_1} \;\in_R\; V(L_1)$

Theorem [K.–MISSETHAN 2022+]

$$(R, r_R) \xrightarrow{D} \mathrm{GWT}(1)$$

$$(L_1, r_{L_1}) \xrightarrow{D} T_\infty$$



$(R, r_R) \xrightarrow{D} \mathrm{GWT}(1)$ $(L_1, r_{L_1}) \xrightarrow{D} T_\infty$

Finer view of local weak limits

$S \;=\; S_g(n,m) \;\in_R\; \mathcal{S}_g(n,m)$ and $2m/n \;\to\; d \in (1,2)$

L_1 largest component of S and $|L_1| \;\sim\; (d-1)\,n$

$R = S \setminus L_1 \;\sim\;$ crtitical ER random graph and $|R| \;\sim\; (2-d)\,n$

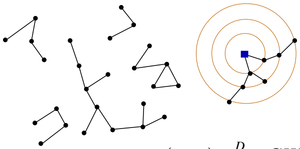
$r_R \;\in_R\; V(R)$, $r_{L_1} \;\in_R\; V(L_1)$ and $r_S \;\in_R\; V(S)$

Theorem [K.–Missethan 2022+]

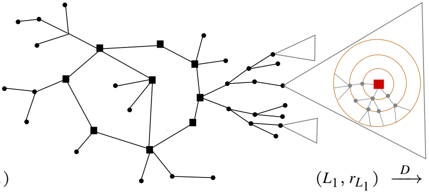
$$(R, r_R) \xrightarrow{D} \mathrm{GWT}(1)$$

$$(L_1, r_{L_1}) \xrightarrow{D} T_\infty$$

$$(S, r_S) \xrightarrow{D} (2-d)\,\mathrm{GWT}(1) \;+\; (d-1)\,T_\infty$$



$(R, r_R) \xrightarrow{D} \mathrm{GWT}(1)$



$(L_1, r_{L_1}) \xrightarrow{D} T_\infty$

Part V.

Topological aspects of random graphs

- ✓ Typical genus of the Erdős-Rényi random graph
- ✓ Random graphs on surfaces with constant genus
- ✓ Benjamini-Schramm local weak limits
- Random graphs on surfaces with non-constant genus

ER random graph vs random graphs on surfaces

IF whp the genus of $G(n,m)$ is $T = T(n,m)$,

THEN $\forall\, g \geq T$

$$\frac{|\mathcal{S}_g(n,m)|}{|\mathcal{G}(n,m)|} \geq \frac{|\mathcal{S}_T(n,m)|}{|\mathcal{G}(n,m)|} \xrightarrow{n\to\infty} 1.$$

ER random graph vs random graphs on surfaces

IF whp the genus of $G(n,m)$ is $T = T(n,m)$,

THEN $\forall\, g \geq T$

$$\frac{|\mathcal{S}_g(n,m)|}{|\mathcal{G}(n,m)|} \geq \frac{|\mathcal{S}_T(n,m)|}{|\mathcal{G}(n,m)|} \xrightarrow{n\to\infty} 1.$$

In other words, for $\forall\, g \geq T$,

$S_g(n,m)$ is indistinguishable from $G(n,m)$ under viewpoint of whp-properties

If for every property $\mathcal{A}$

whp $G(n,m)$ satisfies $\mathcal{A}$ iff whp $S_g(n,m)$ satisfies $\mathcal{A}$

then we say $G(n,m)$ and $S_g(n,m)$ are contiguous.

Contiguity threshold and more on the giant

$m = d \cdot \frac{n}{2}$ for $1 < d < 2$

$T = \nu(d) \cdot n \;=\;$ contiguity threshold

Contiguity threshold and more on the giant

$m = d \cdot \frac{n}{2}$ for $1 < d < 2$

$T = \nu(d) \cdot n =$ contiguity threshold

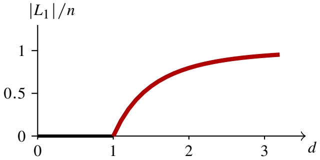
$L_1 =$ largest component in $S_g(n,m) \in_R \mathcal{S}_g(n,m)$

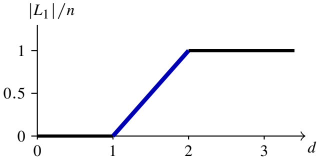
Theorem [Dowden–K.–Mosshammer–Sprüssel 2022+]

whp

$$|L_1| = (1+o(1))\, \rho\, n \qquad \text{if} \quad g \gg T$$

$$|L_1| = (1+o(1))\, (d-1)\, n \qquad \text{if} \quad g \ll T$$





Proof sketch – asymptotic enumeration

$$
\begin{aligned}
|\mathcal{S}_g(n,m)| &= \#\text{ graphs on } [n] \text{ with } m \text{ edges and genus} \leq g \\
&= \sum_{k,\ell} \binom{n}{k} C_g(k,k+\ell)\, U(n-k, m-k-\ell)
\end{aligned}
$$

where

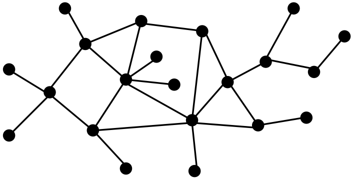
$C_g(k,k+\ell) \quad = \quad$ # complex part on $[k]$ with $k+\ell$ edges

$U(n-k, m-k-\ell) \quad = \quad$ # graphs consisting of
trees or unicyclic components
on $[n-k]$ with $m-k-\ell$ edges

✓ Asymptotic behaviour of $U(n-k, m-k-\ell)$ is well understood

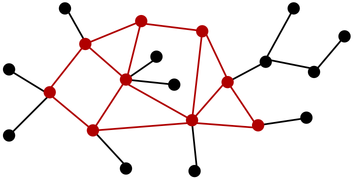
Proof sketch – core-kernel approach

- Complex part G



Proof sketch – core-kernel approach

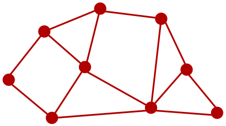
- Complex part G



2-Core $=$ max. subgraph of G with min. degree ≥ 2

Proof sketch – core-kernel approach

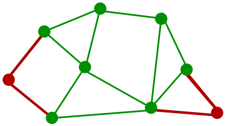
- Complex part G



2-Core $=$ max. subgraph of G with min. degree ≥ 2

Proof sketch – core-kernel approach

- Complex part G

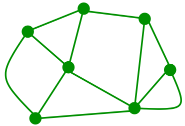


2-Core $=$ max. subgraph of G with min. degree ≥ 2

Kernel $=$ replace each bare-path in 2-core by an edge

Proof sketch – core-kernel approach

- Complex part G

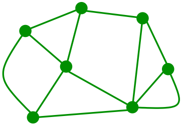


2-Core $=$ max. subgraph of G with min. degree ≥ 2

Kernel $=$ replace each bare-path in 2-core by an edge

Proof sketch – core-kernel approach

- Complex part G



2-Core $=$ max. subgraph of G with min. degree ≥ 2

Kernel $=$ replace each bare-path in 2-core by an edge

$*$ g is genus of G iff g is genus of kernel of G

Proof sketch – construction of complex part

- Construct complex part G

$C_g(k, k+\ell) \;=\; \#$ complex part on $[k]$ with $k+\ell$ edges

Proof sketch – construction of complex part

- Construct complex part G by
  - ▷ choosing kernel of G from the set of possible candidates

$$C_g(k, k+\ell) \;=\; \#\text{ complex part on } [k] \text{ with } k+\ell \text{ edges}$$

$$= \sum_{i,j} K_g(2\ell - j)$$

Proof sketch – construction of complex part

- Construct complex part G by
  - ▷ choosing kernel of G from the set of possible candidates
  - ▷ putting on its edges vertices of degree 2 to obtain 2-core

$$C_g(k, k+\ell) = \#\text{ complex part on } [k] \text{ with } k+\ell \text{ edges}$$

$$= \sum_{i,j} K_g(2\ell - j) \frac{(k)_i}{(2\ell - j)!} \binom{i - a\ \ell - 1}{3\ell - j - 1}$$

Proof sketch – construction of complex part

- Construct complex part G by
  - ▷ choosing kernel of G from the set of possible candidates
  - ▷ putting on its edges vertices of degree 2 to obtain 2-core
  - ▷ adding a forest rooted at vertices of 2-core

$$C_g(k, k+\ell) = \#\text{ complex part on } [k] \text{ with } k+\ell \text{ edges}$$

$$= \sum_{i,j} K_g(2\ell - j) \frac{(k)_i}{(2\ell - j)!} \binom{i - a\,\ell - 1}{3\ell - j - 1} i\, k^{k-i-1}$$

Proof sketch – construction of complex part

- Construct complex part G by
  - ▷ choosing kernel of G from the set of possible candidates
  - ▷ putting on its edges vertices of degree 2 to obtain 2-core
  - ▷ adding a forest rooted at vertices of 2-core

$$C_g(k, k+\ell) = \#\text{ complex part on } [k] \text{ with } k+\ell \text{ edges}$$

$$= \sum_{i,j} K_g(2\ell - j) \, \frac{(k)_i}{(2\ell - j)!} \binom{i - a\,\ell - 1}{3\ell - j - 1} \, i\, k^{k-i-1}$$

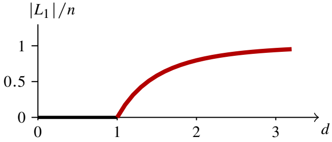
∗ Asymptotic behaviour of $C_g(k, k+\ell)$?

$\Longrightarrow$ combinatorial variants of Laplace method ✓

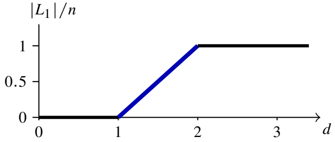
Summary and open problems

Global properties of $S_g(n,m)$ when $m = d \cdot \frac{n}{2}$ for $d > 1$

- contiguity threshold $T = \nu(d) \cdot n$
- largest component L_1



$|L_1| \sim \rho\, n$ if $g \gg T$



$|L_1| \sim (d-1)\, n$ if $g \ll T$

$Q1.$ Order of largest component when $g = \Theta(T)$?

$Q2.$ Length of longest cycle when $g \ll T$ or $g = \Theta(T)$?

$*$ when $g \gg T$, it follows from $G(n,m)$ [AJTAI-KOMLÓS-SZEMERÉDI 1981]

Summary of the minicourse

I. Prelude

II. Basic probabilistic tools

III. Erdős-Rényi random graphs

IV. Higher-dimensional analogues

V. Random subgraphs of the hypercube

VI. Topological aspects of random graphs

* Slides available at

  https://www.math.tugraz.at/ ~ kang/talks/Kang-RandNET2022.pdf