Random Graphs I

Y. Kohayakawa (São Paulo)

Chorin, 31 July 2006

Outline of Lecture I

1. **Probabilistic preliminaries:** basics, binomial distribution
2. **Models of random graphs:** the models, monotonicity, equivalence
3. **Jumbledness and expansion:** edge-distribution, expansion
4. **Threshold phenomena:** Thresholds, giant component

Probabilistic preliminaries

- Focus on *discrete probability spaces*: $(\Omega, \mathbb{P})$
  - $|\Omega| < \infty$
  - $\mathbb{P}\colon \Omega \to [0,1]$
  - $\sum_{\omega\in\Omega} \mathbb{P}(\omega) = 1$
- *Random variable (r.v.)*: $X\colon \Omega \to \mathbb{R}$

Expectation and linearity

- *Expectation*:

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) = \sum_{x} x\mathbb{P}(X = x) \tag{1}$$

- *Linearity*:

$$\mathbb{E}(\sum_{i} a_i X_i) = \sum_{i} a_i \mathbb{E}(X_i) \tag{2}$$

Variance and standard deviation

- *Variance*:

$$\sigma^2(X) = \mathsf{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \qquad (3)$$

- *Standard deviation*:

$$\sigma(X) = \sqrt{\mathsf{Var}(X)} \qquad (4)$$

Indicator random variables

- $\triangleright$ $X_E = [\text{event } E \text{ holds}]$
- $\triangleright$ $X = \sum_{E \in \mathcal{E}} X_E$ [$=$ number of $E \in \mathcal{E}$ that hold]
- $\triangleright$ $\mathbb{E}(X) = \sum_{E \in \mathcal{E}} \mathbb{E}(X_E) = \sum_{E \in \mathcal{E}} \mathbb{P}(E \text{ holds})$
- $\triangleright$ $\mathrm{Var}(X) = \sum_{(E,E')} \mathrm{Cov}(X_E, X_{E'})$
- $\triangleright$ $\mathrm{Cov}(X, X') = \mathbb{E}(XX') - \mathbb{E}(X)\mathbb{E}(X')$ [$= 0$ if X and X' independent]

Markov's and Chebyshev's inequality

▷ Markov: if $X \geq 0$, then for all $t > 0$ we have

$$\mathbb{P}(X \geq t) \leq \frac{1}{t}\mathbb{E}(X). \tag{5}$$

∘ **Consequence:** if X is integer-valued, taking $t = 1$ gives

$$\mathbb{P}(X > 0) = \mathbb{P}(X \geq 1) \leq \mathbb{E}(X). \tag{6}$$

Often, just estimate $\mathbb{E}(X)$ and show that $\mathbb{E}(X) = o(1)$.

Markov's and Chebyshev's inequality

- Chebyshev: for all $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{1}{t^2}\mathrm{Var}(X). \tag{7}$$

Proof*.* Apply Markov to $Y = (X - \mathbb{E}(X))^2$. □

- Taking $t = \mathbb{E}(X)$, we have

$$\mathbb{P}(X = 0) \leq \mathbb{P}(|X - \mathbb{E}(X)| \geq \mathbb{E}(X)) \leq \frac{\mathrm{Var}(X)}{\mathbb{E}(X)^2}. \tag{8}$$

Markov's and Chebyshev's inequality

▷ Cauchy–Schwarz: May obtain small improvement applying CS:

$$\mathbb{P}(X = 0) \leq \frac{\mathrm{Var}(X)}{\mathbb{E}(X)^2 + \mathrm{Var}(X)} = \frac{\mathrm{Var}(X)}{\mathbb{E}(X^2)}. \qquad (9)$$

For non-negative integer-valued r.vs:

$$\mathbb{P}(X \geq 1) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}. \qquad (10)$$

Proof. Exercise 1. □

Basic concentration

If $\mathrm{Var}(X) \ll \mathbb{E}(X)^2$, then X is concentrated around its expectation: for any fixed $\varepsilon > 0$,

$$\mathbb{P}[|X - \mathbb{E}(X)| \geq \varepsilon \mathbb{E}(X)] \leq \frac{\mathrm{Var}(X)}{\varepsilon^2 \mathbb{E}(X)^2} = o(1). \tag{11}$$

Therefore, have $\mathbb{P}[X = (1 \pm \varepsilon)\mathbb{E}(X)]$ with probability $1 - o(1)$.

Binomial distribution

$X \sim \mathrm{Bi}(n, p)$: $X = X_1 + \cdots + X_n$, with each $X_i \sim \mathrm{Be}(p)$

- ▷ $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- ▷ $\mathbb{E}(X) = np$
- ▷ $\mathbb{E}_r(X) = \mathbb{E}[(X)_r] = \mathbb{E}[X(X-1)\dots(X-r+1)] = (n)_r p^r$. This gives $\mathrm{Var}(X) = np(1-p)$.
- ▷ X concentrated around $\mathbb{E}(X)$ if $np \to \infty$

Poisson distribution

$X \sim \mathrm{Po}(\lambda)$: integer-valued, mean $\lambda > 0$, with

$$\mathbb{P}(X = k) = \frac{1}{k!}e^{-\lambda}\lambda^k \tag{12}$$

- $\mathbb{E}_r(X) = \mathbb{E}[(X)_r] = \lambda^r$
- $\mathrm{Bi}(n, p) \xrightarrow{d} \mathrm{Po}(\lambda)$ if $np \to \lambda$ as $n \to \infty$

Hypergeometric distribution

$X \sim \text{Hyp}(n, b, d)$: $X = |D \cap B|$ when $D \in \binom{[n]}{d}$ uniformly at random, and $B \subset [n]$ with $|B| = b$ is fixed

- ▷ $\mathbb{P}(X = k) = \binom{b}{k}\binom{n-b}{d-k}\binom{n}{d}^{-1} = \binom{d}{k}\binom{n-d}{b-k}\binom{n}{b}^{-1}$
- ▷ $\mathbb{E}(X) = bd/n$

Exponential bounds for the binomial

Suppose $X \sim \mathrm{Bi}(n, p)$.

Theorem 1. *We have*

$$\mathbb{P}(X \geq k) \leq \binom{n}{k} p^k \leq \left(\frac{enp}{k}\right)^k. \qquad (13)$$

Proof. Exercise 2. □

▷ If $k = \lambda np$, bound is $(e/\lambda)^{\lambda np} = e^{-c_\lambda np}$, where $c_\lambda = \lambda(\log \lambda - 1)$.

Exponential bounds for the binomial

Suppose $X \sim \mathrm{Bi}(n, p)$.

Theorem 2. *Let* $\mu = \mathbb{E}(X) = np$ *and* $t \geq 0$*. Then*

$$\mathbb{P}(X \geq \mu + t) \leq \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\} \tag{14}$$

and

$$\mathbb{P}(X \leq \mu - t) \leq \exp\left\{-\frac{t^2}{2\mu}\right\}. \tag{15}$$

Exponential bounds for the binomial

Suppose $X \sim \mathrm{Bi}(n, p)$; $\mu = np$.

Theorem 3. *If* $\varepsilon \leq 3/2$, *then*

$$\mathbb{P}\left(|X - \mu| \geq \varepsilon\mu\right) \leq 2\exp\left\{-\frac{1}{3}\varepsilon^2\mu\right\}. \tag{16}$$

Exponential bounds for the hypergeometric

Suppose $X \sim \text{Hyp}(n, b, d)$.

Theorem 4. *We have*

$$\mathbb{P}(X \geq k) \leq \binom{d}{k} \left(\frac{b}{n}\right)^k \leq \left(\frac{ebd}{kn}\right)^k. \quad (17)$$

Proof*.* Exercise 3. □

▷ If $k = \lambda bd/n$, then the bound is $(e/\lambda)^{\lambda bd/n} = e^{-c_\lambda bd/n}$, where $c_\lambda = \lambda(\log \lambda - 1)$.

Exponential bounds for the hypergeometric

Suppose $X \sim \text{Hyp}(n, b, d)$.

Theorem 5. *Let* $\mu = \mathbb{E}(X) = bd/n$ *and* $t \geq 0$*. Then*

$$\mathbb{P}(X \geq \mu + t) \leq \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\} \tag{18}$$

and

$$\mathbb{P}(X \leq \mu - t) \leq \exp\left\{-\frac{t^2}{2\mu}\right\}. \tag{19}$$

Exponential bounds for the hypergeometric

Suppose $X \sim \mathrm{Hyp}(n, b, d)$, $\mu = bd/n$.

Theorem 6. *If* $\varepsilon \leq 3/2$, *then*

$$\mathbb{P}\left(|X - \mu| \geq \varepsilon\mu\right) \leq 2\exp\left\{-\frac{1}{3}\varepsilon^2\mu\right\}. \tag{20}$$

Models of random graphs

- $\triangleright$ $G(n, p)$: each element of $\binom{[n]}{2}$ is present with probability p, independently of all others
- $\triangleright$ $G(n, M)$: uniform space on $\binom{\binom{[n]}{2}}{M}$
- $\triangleright$ $\mathbf{G} = (G_t)_{t=0}^{N}$: random processes $G_0 \subset G_1 \subset \cdots \subset G_N$ ($N = \binom{n}{2}$), with each G_i on $[n]$, say, and G_i obtained from G_{i-1} by the addition of a new random edge. Space has cardinality $N!$.

Always interested in $n \to \infty$. Use the terms 'almost surely', 'almost every', 'almost always', etc to mean 'with probability $\to 1$ as $n \to \infty$'.

Monotonicity theorems

Definition 7 (Graph property)**.** *A graph property is a family of graphs closed under isomorphism.*

Definition 8 (Increasing and decreasing properties)**.** *A graph property is decreasing if the removal of an edge does not destroy the property. A graph property is increasing if the addition of an edge does not destroy the property (vertices are not added).*

▷ Examples: being planar, being connected

Monotonicity theorems

Theorem 9. *Suppose $0 \leq p \leq p' \leq 1$. If $\mathcal{P}$ is an increasing graph property, then $\mathbb{P}(G(n,p) \in \mathcal{P}) \leq \mathbb{P}(G(n,p') \in \mathcal{P})$.*

Proof. Exercise 4. □

▷ '2-round exposure trick': $G(n,p') = G(n,p) \cup G(n,p'')$ (union of two independent r.gs), with $1 - p' = (1-p)(1-p'')$

Monotonicity theorems

Theorem 10. *Suppose* $0 \leq M \leq M' \leq N = \binom{n}{2}$. *If* $\mathcal{P}$ *is an increasing graph property, then* $\mathbb{P}(G(n, M) \in \mathcal{P}) \leq \mathbb{P}(G(n, M') \in \mathcal{P})$.

Proof. Exercise 5. □

Equivalence theorems

Theorem 11. *Suppose* $\mathcal{P}$ *is an increasing property, let* $M = M(n) \to \infty$, *and suppose* $\delta > 0$ *is a constant with* $(1+\delta)M/N = (1+\delta)M/\binom{n}{2} \leq 1$. *Set* $p = p(n) = M/N$.

(i) *If* $\mathbb{P}(G(n,p) \in \mathcal{P}) \to 1$, *then* $\mathbb{P}(G(n,M) \in \mathcal{P}) \to 1$.

(ii) *If* $\mathbb{P}(G(n,p) \in \mathcal{P}) \to 0$, *then* $\mathbb{P}(G(n,M) \in \mathcal{P}) \to 0$.

(iii) *If* $\mathbb{P}(G(n,M) \in \mathcal{P}) \to 1$, *then* $\mathbb{P}(G(n,(1+\delta)p) \in \mathcal{P}) \to 1$.

(iv) *If* $\mathbb{P}(G(n,M) \in \mathcal{P}) \to 0$, *then* $\mathbb{P}(G(n,(1-\delta)p) \in \mathcal{P}) \to 0$.

Proof*.* Exercise 6. □

Jumbledness

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

Definition 12 ((p,η)-uniform)**.** *Let p and $\eta > 0$ be given. We say that G is (p,η)-uniform if, for all U, $W \subset V$, with $U \cap W = \emptyset$ and $|U|$, $|W| \geq \eta n$, we have*

$$\Big| e(U, W) - p|U||W| \Big| \leq \eta p |U||W|, \tag{21}$$

where $e(U, W)$ denotes the number of edges with one endvertex in U and the other in W.

Jumbledness

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

Definition 13 ((p, α)-bijumbled). *Let p and $\alpha > 0$ be given. We say that G is (p, α)-bijumbled if, for all U, $W \subset V$, with $U \cap W = \emptyset$ and $1 \leq |U| \leq |W| \leq pn|U|$, we have*

$$\Big| e(U, W) - p|U||W| \Big| \leq \alpha\sqrt{|U||W|}. \tag{22}$$

Particular interest: $\alpha = O(\sqrt{np})$. We often set $d = np$ (and call this the 'average degree', which is, of course, not quite right).

Jumbledness

Theorem 14. *Let* $G = G^n = (V, E)$ *be a* (p, α)*-bijumbled graph. Then, for all* $U \subset V$*, we have*

$$\left| e(G[U]) - p\binom{|U|}{2} \right| \leq \alpha |U|. \tag{23}$$

Proof. Exercise 7. □

Jumbledness

Theorem 15. *For every* $\eta > 0$ *there is* C *such that if* $d = pn \geq C$*, then* $G(n,p)$ *is a.s.* (p,η)*-uniform.*

Proof. Exercise 8. □

Theorem 16. *For every* $0 < p = p(n) < 1$*, the random graph* $G(n,p)$ *is a.s.* $(p, e^{3/2}\sqrt{d})$*-bijumbled, where* $d = np$*.*

Proof. Exercise 9. □

Exercise 10: why do we have the condition $1 \leq |U| \leq |W| \leq pn|U|$ in Definition 13?

Jumbledness

Corollary 17. *Suppose* $pn \geq C \log n$ *for some constant* $C > 3$. *Then a.e.* $G(n,p)$ *satisfies* (22) *for every pair of disjoint sets* $U, W \subset V(G(n,p))$ *with* $\alpha = e^{3/2}\sqrt{d}$.

Proof (Sketch). Theorem 16 tells us that $G(n,p)$ is a.s. $(p, e^{3/2}\sqrt{d})$-bi-jumbled. Now let U and W be such that $|W| > d|U|$. Then $e^{3/2}\sqrt{d|U||W|} > e^{3/2}d|U|$. In particular, $p|U||W| - e^{3/2}\sqrt{d|U||W|} \leq p|U|n - e^{3/2}d|U| < 0 \leq e(U,W)$.

As $d = np = C \log n$ and $C > 3$, we have that $\Delta(G(n,p)) \leq 2d$ almost surely. Therefore $e(U,W) \leq 2d|U| \leq e^{3/2}d|U| \leq p|U||W| + e^{3/2}\sqrt{d|U||W|}$. $\square$

Expansion results

Definition 18 ((b, f)-expansion). *Let $B = (U, W; E)$ be a bipartite graph with vertex classes U and W and edge set E. Let positive reals b and f be given. We say that B is $(b, f; U)$-expanding if, for every $X \subset U$ with $|X| \le b$, we have $|\Gamma(X)| \ge f|X|$. If B is both $(b, f; U)$-expanding and $(b, f; W)$-expanding, let us say that B is (b, f)-expanding.*

As usual, $\Gamma(X)$ is the neighbourhood of X, that is, the set of all vertices adjacent to some $x \in X$.

Expansion results

Let $G = G^n = (V, E)$ be $(p, A\sqrt{d})$-bijumbled, where $d = np$. Suppose U and $W \subset V$ are disjoint; let $|W| = \alpha n$. Suppose

$$d_W(u) = |\Gamma(u) \cap W| \geq \rho p|W| \tag{24}$$

for all $u \in U$.

Theorem 19. *For any* $\eta > 0$ *and any* $0 < f \leq (\eta\alpha\rho/A)^2 d$*, the bipartite graph* $G[U, W]$ *is* $((1-\eta)\rho|W|/f, f; U)$*-expanding.*

Expansion results

Proof. By contradiction: let f be as in the statement. Let $X \subset U$ be such that $|X| \le (1-\eta)\rho|W|/f$. Let $Y = \Gamma(X) \cap W$ and suppose $|Y| < f|X|$.

By the $(p, A\sqrt{d})$-bijumbledness condition on G, we have

$$e(X,Y) \le p|X||Y| + A\sqrt{d|X||Y|} < p|X|(1-\eta)\rho|W| + A\sqrt{d|X||Y|}, \tag{25}$$

and, from (24), we deduce that

$$e(X,Y) = e(X,W) \ge \rho p|W||X|. \tag{26}$$

Combining (25) and (26), we have $(\eta\rho p|W||X|)^2 < A^2 d|X||Y|$. Therefore

$$|Y| > \frac{(\eta\rho p|W||X|)^2}{A^2 d|X|} \ge \left(\frac{\eta\rho\alpha}{A}\right)^2 d|X| \ge f|X|. \tag{27}$$

As we supposed that $|Y| < f|X|$, we have a contradiction. □

Long paths in expanding bipartite graphs

The following lemma is known as the bipartite version of *Posá's lemma*.

Lemma 20. *Let $b \geq 1$ be an integer. If the bipartite graph B is $(b, 2)$-expanding, then B contains a path P^{4b} on $4b$ vertices.*

Proof. Later we shall see a proof of Posá's original lemma. □

The Friedman–Pippenger lemma

Suppose $G = (V, E)$ is (b, f)*-expanding*: *every* $X \subset V$ *with* $|X| \leq b$ *is such that* $|\Gamma_G(X)| \geq f|X|$.

Theorem 21 (Friedman and Pippenger 1987)**.** *Any* $(2n-2, d+1)$*-expander contains every tree* $T = T^n$ *with maximum degree* $\Delta(T) \leq d$*.*

Proof*.* Exercise 11^{++}. □

Open problem 12: give an efficient algorithm for finding the tree guaranteed in Theorem 21.

Random graphs are fault tolerant

Write $G \rightarrow_\eta \mathcal{J}$ if every $H \subset G$ with $|E(H)| \geq \eta |E(G)|$ contains a copy of every $J \in \mathcal{J}$ as a subgraph.

Theorem 22. *For any $\eta > 0$ and any Δ, there is C such that a.e. $G = G(n, p)$ with $p = C/n$ satisfies*

$$G(n, p) \rightarrow_\eta \mathcal{T}, \tag{28}$$

where $\mathcal{T}$ is the family of all trees $T = T^t$ with $t \leq n/C$ and $\Delta(T) \leq \Delta$.

Proof. Exercise 13^+. □

▷ There exist linear fault-tolerant graphs for trees. Exercise 14^{++}: how about for even cycles?

Threshold functions

Consider $G(n,p)$ [similar for $G(n,M)$]. Let $\mathcal{P}$ be an increasing graph property.

Definition 23 (Threshold)**.** *The function $p_0 = p_0(n)$ is a threshold function for $\mathcal{P}$ if*

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \text{ has } \mathcal{P}) = \begin{cases} 0 & \text{if } p \ll p_0 \\ 1 & \text{if } p \gg p_0. \end{cases} \tag{29}$$

▷ 0-*statement*, 1-*statement*

Sharp threshold functions

Let $\mathcal{P}$ be an increasing graph property.

Definition 24 (Sharp and coarse thresholds). *The function $p_0 = p_0(n)$ is a sharp threshold function for $\mathcal{P}$ if, for every $\varepsilon > 0$, we have*

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \text{ has } \mathcal{P}) = \begin{cases} 0 & \text{if } p \le (1-\varepsilon)p_0 \\ 1 & \text{if } p \ge (1+\varepsilon)p_0. \end{cases} \tag{30}$$

Coarse threshold: not sharp

Threshold functions, examples

- $K^4 \subset G(n,p)$: $p_0 = p_0(n) = n^{-2/3}$ [Exercise 15]. This threshold is coarse [Exercise 16].
- $G(n,p) \to (K^3)^{\mathsf{v}}_2$: $p_0 = n^{-2/3}$ [Exercise 17^{++}].
- $G(n,p) \to (K^3)^{\mathsf{e}}_2$: $p_0 = n^{-1/2}$ [Exercise 18^{++}; > 2 colours: Exercise 19^{++}; Open problem 20: conjectured to be sharp for all $k \geq 2$; very tough for $k = 2$]
- $G(n,p) \to_{1/2+\eta} K^3$: $p_0 = n^{-1/2}$ [Exercise 21^{++}; Open problem 22: conjectured to be sharp]

The Bollobás–Thomason theorem

Theorem 25. *Let $\mathcal{P}$ be an increasing property. Then $\mathcal{P}$ admits a threshold.*

Proof (Sketch). Consider p_ε so that $G(n, p_\varepsilon)$ has $\mathcal{P}$ with probability $\geq \varepsilon$. Let $G = G_1 \cup \cdots \cup G_t$, where each G_i is an independent copy of $G(n, p_\varepsilon)$. Then $G = G(n, p')$ with $p' \leq tp$. Suppose $t = t(\varepsilon)$ is such that $(1-\varepsilon)^t \leq \varepsilon$. Then $G(n, p')$ has $\mathcal{P}$ with probability at least $1 - \varepsilon$. This implies the theorem [Exercise 23]. □

A (very) sharp threshold

Theorem 26. *Let*

$$p = \frac{1}{n}\left(\log n + c_n\right). \tag{31}$$

Then

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \text{ is connected}) = \begin{cases} 0 & \text{if } \lim_n c_n = -\infty, \\ e^{-e^{-c}} & \text{if } \lim_n c_n = c \in \mathbb{R}, \\ 1 & \text{if } \lim_n c_n = \infty. \end{cases} \tag{32}$$

Proof*.* Exercise 24^{++}. □

The Friedgut theorem for sharp thresholds

Theorem 27. *Let $\mathcal{P}$ be an increasing graph property with a coarse threshold. Then there exist real constants $0 < c < C$ and $\beta > 0$, a rational ρ, and a sequence $p = p(n)$ satisfying*

$$cn^{-1/\rho} < p(n) < Cn^{-1/\rho}, \tag{33}$$

such that $\beta < \mathbb{P}[G(n, p) \in \mathcal{P}] < 1 - \beta$ for infinitely many n.

The Friedgut theorem for sharp thresholds

Given a graph M and a disjoint set of n vertices, let M^* be a labelled copy of M placed uniformly at random on one of the $n!/(n-|V(M)|)!$ possible ways.

Theorem 28. *Furthermore, there exist α and $\xi > 0$ and a balanced graph M with density ρ for which the following holds: For every graph property $\mathcal{G}$ such that $G(n,p) \in \mathcal{G}$ a.s., there are infinitely many values of n for which there exists a graph G on n vertices for which the following holds:*

(i) $G \in \mathcal{G}$,

(ii) $G \notin \mathcal{P}$,

(iii) $\mathbb{P}(G \cup M^* \in \mathcal{P}) > 2\alpha$,

(iv) $\mathbb{P}(G \cup G(n, \xi p) \in \mathcal{P}) < \alpha$.

The emergence of the giant component

For a graph G, write let $L_k(G)$ for the number of vertices in the kth largest component.

Theorem 29. *Let $\varepsilon > 0$ be fixed. For almost every random graph process $\mathbf{G} = (G_t)_{t=0}^N$, the following holds:*

(i) we have $L_1(G_t) = o(n)$ for all $t \leq (1/2 - \varepsilon)n$,

(ii) we have $L_1(G_t) \geq cn$ and $L_2(G_t) = o(n)$ for all $t \geq (1/2 + \varepsilon)n$, where $c = c(\varepsilon)$ is a constant that depends only on ε.

Thus, at around time $t = n/2$, our evolving graph G_t suffers a sudden change in structure: the so called *giant component* emerges.