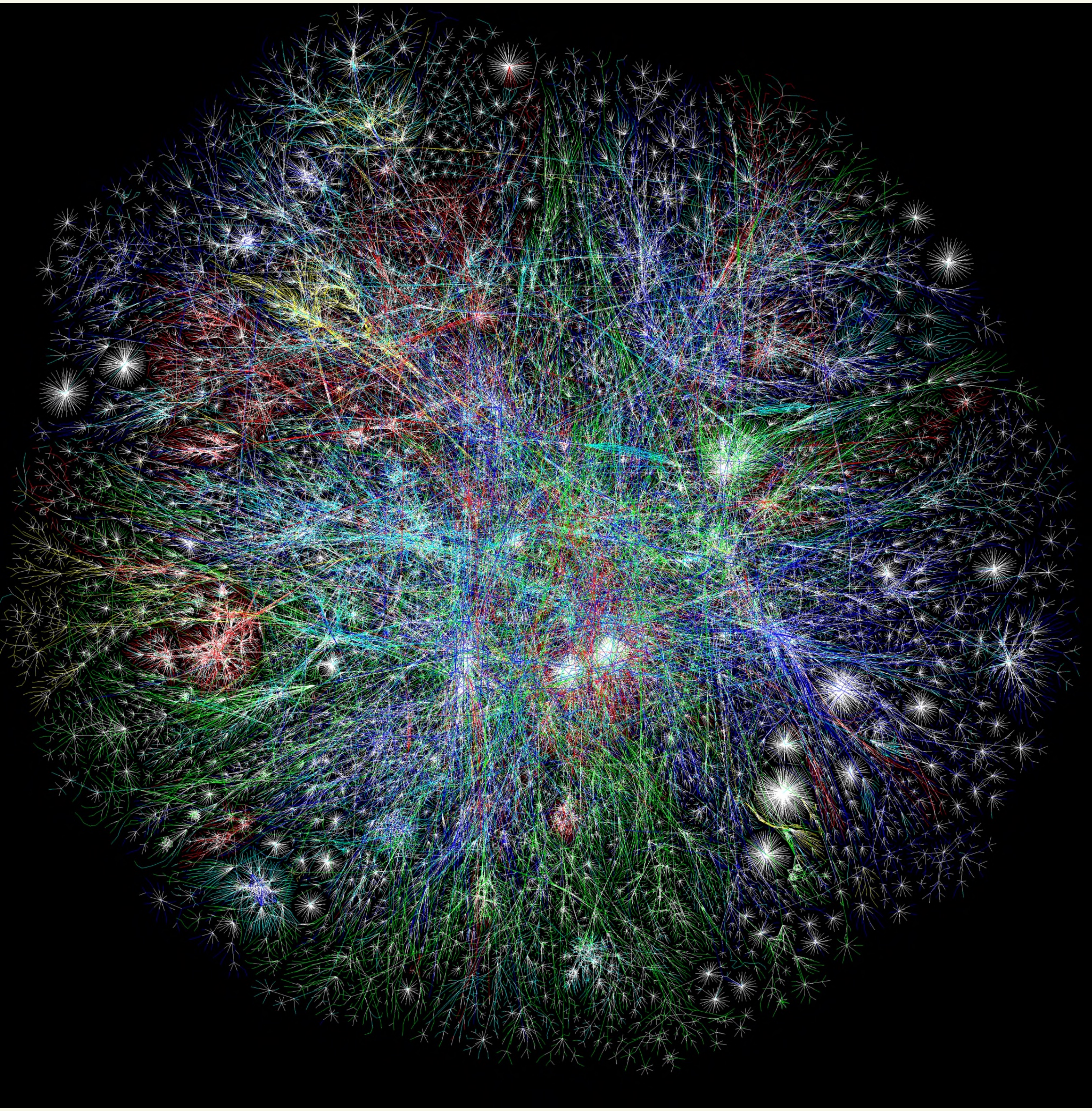
LECTURE 8

- Networks = graphs. Social (friendships, Instagram), Technological (Internet), Biological (protein interactions)

Example: Internet (vertices = IP addresses):

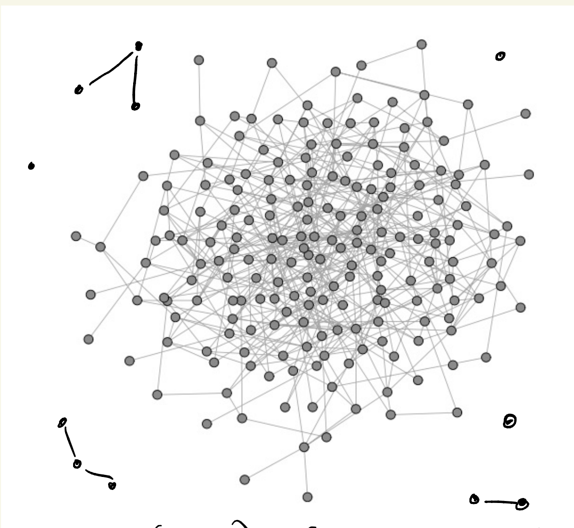


• Model: random graphs.

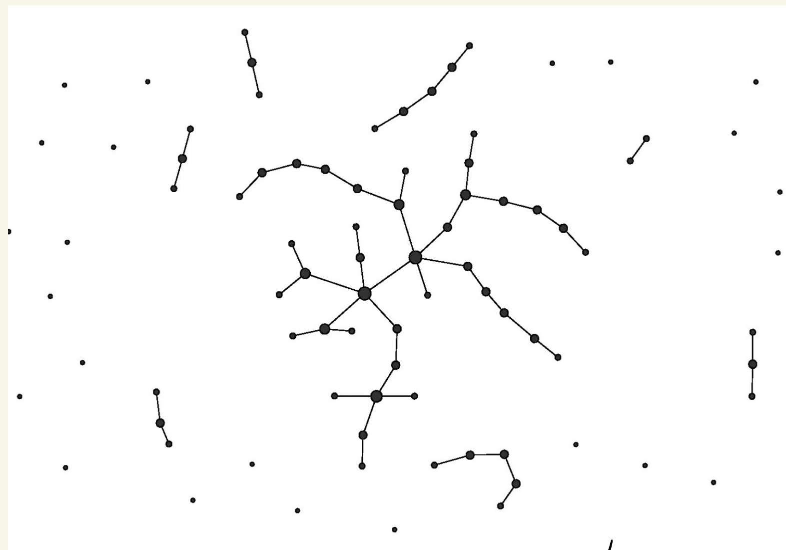
Simplest:

Def Erdös-Renyi model $G(N,p)$:

Fix a set of N vertices.
Connect each pair of vertices with an edge independently, with prob. p.



$G(N,p)$ for $N=200$, $p=\frac{1}{40}$



$G(N,p)$ for $p=\frac{1}{100}$.

• **Def** The degree of a vertex i is $\deg(i) = \#$ edges connected to i

$$\deg(i) = S_{N-1} \sim \text{Binom}(N-1, p) \Rightarrow \mathbb{E}\deg(i) = (N-1)p =: d$$

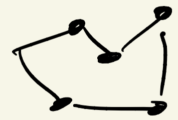
"Expected degree"

Phase transitions [Erdös-Rényi 1960] $N \to \infty$, $p = p(N)$:

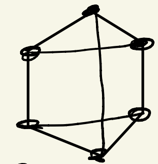
- $d = 0$: G = empty graph.
- subcritical: $d < 1-\varepsilon$: G = small connected components: isolated vertices & trees. All components have size $O(\log N)$.
- supercritical: $d > 1+\varepsilon$: G = a giant connected component (constant fraction of vertices) + isolated vertices + trees of size $O(\log N)$ as before
- connected: $d > (1+\varepsilon)\log N$: G is connected (giant component takes over).

Evolution.

Def a graph is d-regular if $\deg(i) = d \ \ \forall i$



2-regular



3-regular

We will show: $d \sim \log N$ is a phase transition for ≈ regularity of $G(N,p)$: $\begin{cases} d \gg \log N \Rightarrow \text{almost } d\text{-regular} \\ d \ll \log N \Rightarrow \text{very far from it.} \end{cases}$

THM (Regularity of dense random graphs)

$\exists$ abs. const $C > 0$ such that if $d > C \log N$ then $G(N,p)$ is almost d-regular with high prob:

$$\mathbb{P}\{\forall i: \underbrace{0.9d \le \deg(i) \le 1.1d}_{E_i}\} \ge 0.9.$$

Proof ① Fix any vertex $i \in \{1, \dots, N\}$.

$$\mathbb{P}(E_i^c) = \mathbb{P}\{|\deg(i) - d| \ge 0.1d\} \ \textcircled{\le}$$

($\deg(i)$: Binomial with mean d $\Rightarrow$ use Chernoff inequality (lec 7):)

Binom. with mean μ

$$\mathbb{P}\{|S_N - \mu| \ge \delta\mu\} \le 2\exp(-\delta^2\mu/3) \quad \forall \delta \in [0,1]$$

($\mu = d$, $\delta = 0.1$)

$$\textcircled{\le} \ 2\exp\left(-\frac{0.1^2 d}{3}\right) \le 2\exp\left(-\frac{0.01 \cdot C \log N}{3}\right) \le \frac{1}{10N}$$

if we choose C a large const.

② Union Bound

$$\mathbb{P}\left(\bigcup_{i=1}^N E_i^c\right) \le \sum_{i=1}^N \mathbb{P}(E_i^c) \le N \cdot \frac{1}{10N} = \frac{1}{10}.$$

$$\Rightarrow \mathbb{P}\left(\bigcap_{i=1}^N E_i\right) \ge 1 - \frac{1}{10} = \frac{9}{10} \text{ (de Morgan law)}.$$ QED

THM (Irregularity of sparse random graphs)

> $\exists$ abs. const $c>0$ such that if $d < c\log N$,
> $G(N,p)$ has a vertex with too large degree with high prob: ← "hub"
> $$\mathbb{P}\{\exists i : \underbrace{\deg(i) \geq 10d}_{\text{"}E_i\text{"}}\} \geq 0.9$$

Proof $\deg(i) \sim$ Binomial with mean d

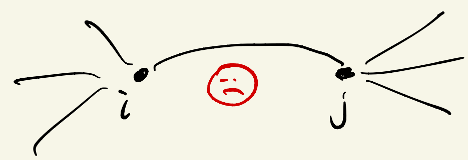
Use the "reverse Chernoff's inequality" (HW):

Binom, mean μ → $\mathbb{P}\{S_N \geq t\} \geq (\mu/t)^t \quad \forall\, t \geq \mu$ for $\mu = d$, $t = 10d$

$$\mathbb{P}(E_i) \geq e^{-d}\left(\frac{d}{10d}\right)^{10d} \geq e^{-20d} \geq e^{-20c\log N} \geq \frac{100}{N}$$ if we choose $c>0$ a small const.

- $$\mathbb{P}\Big(\bigcup_{i=1}^{N} E_i\Big) = 1 - \mathbb{P}\Big(\bigcap_{i=1}^{N} E_i^c\Big) = 1 - \prod_{i=1}^{N} \mathbb{P}(E_i^c) \quad \text{(independence)}$$
$$\geq 1 - \prod_{i=1}^{N}(1 - \mathbb{P}(E_i)) \geq 1 - \underbrace{\left(1 - \frac{100}{N}\right)^N}_{\approx\, e^{-100}} \geq 0.9 \quad \text{(DIY)}$$

ERROR! $\deg(i)$ are NOT independent

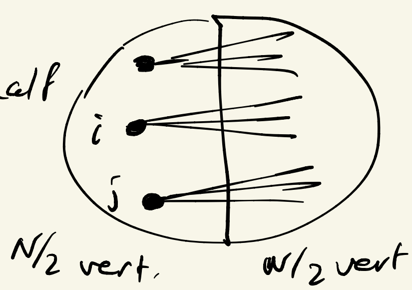


FIX: Consider a bipartite subgraph of G:

split the vertices into two halves:

Count only the edges from left half

These degrees are independent



HW

QED