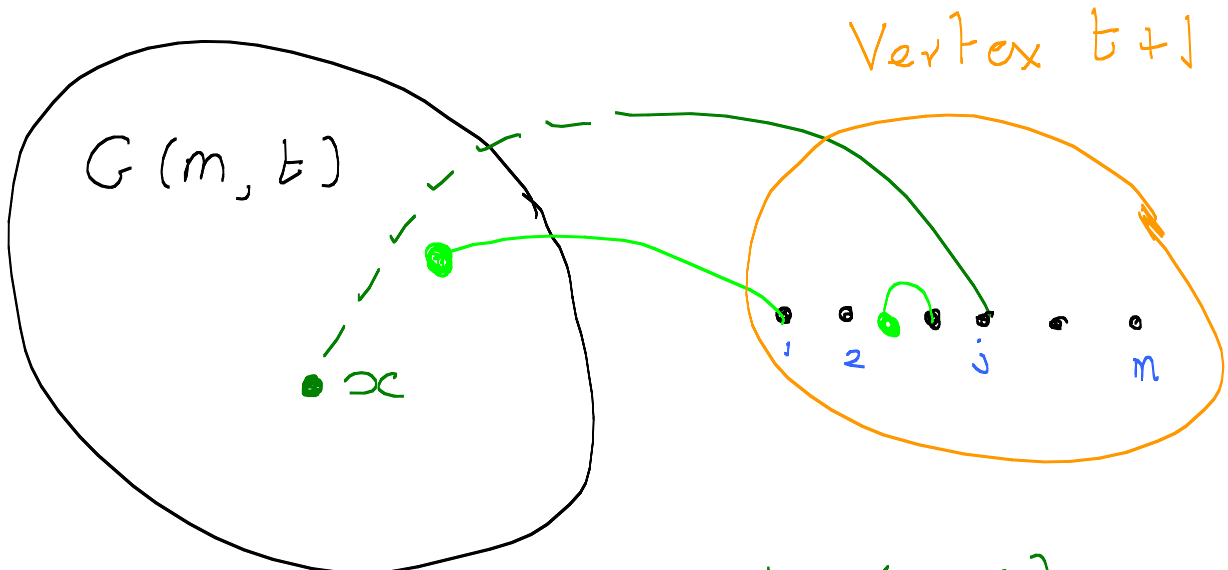
Coupling "Web Graph" and $G_{n,p}$

Note Title 1/14/2004

Bollobás & Riordan

Model



$$P_r(x = v) = \frac{\deg(v,t)}{2mt + 2j - 1}$$

$$\left[\quad P_r(\text{loop}) = \frac{1}{2mt + 2j - 1} \quad \right]$$

The paper studies following problem: If all but αt vertices of $G(m,t)$ are removed, how big is the largest component in what is left?

A question of robustness under "attack".

Theorem 1

There are absolute constants m_0, c such that whp $G(m,t)$ has the following properties if $m \geq m_0$:

(a) Every induced subgraph on $\geq \frac{100 \log m}{m} t$ vertices contains a component of order $\geq \frac{2 \log m}{m} t$.

(b) $G(m,t)$ contains an independent set of size $\geq \frac{c \log m}{m} t$.

Theorem 2

Let $\eta < \frac{1}{2}$ be fixed. Then there exist constants $A, c > 0$ such that for each fixed m we can generate random graphs G_1, G_2 on same vertex set such that

(a) $G_1 = G(m, t)$

(b) $G_2 \sim G_{t,p}$ $\qquad p = \frac{\eta m}{t}$

(c) Whp $e(G_2 \setminus G_1) \leq Ae^{-cm}t$

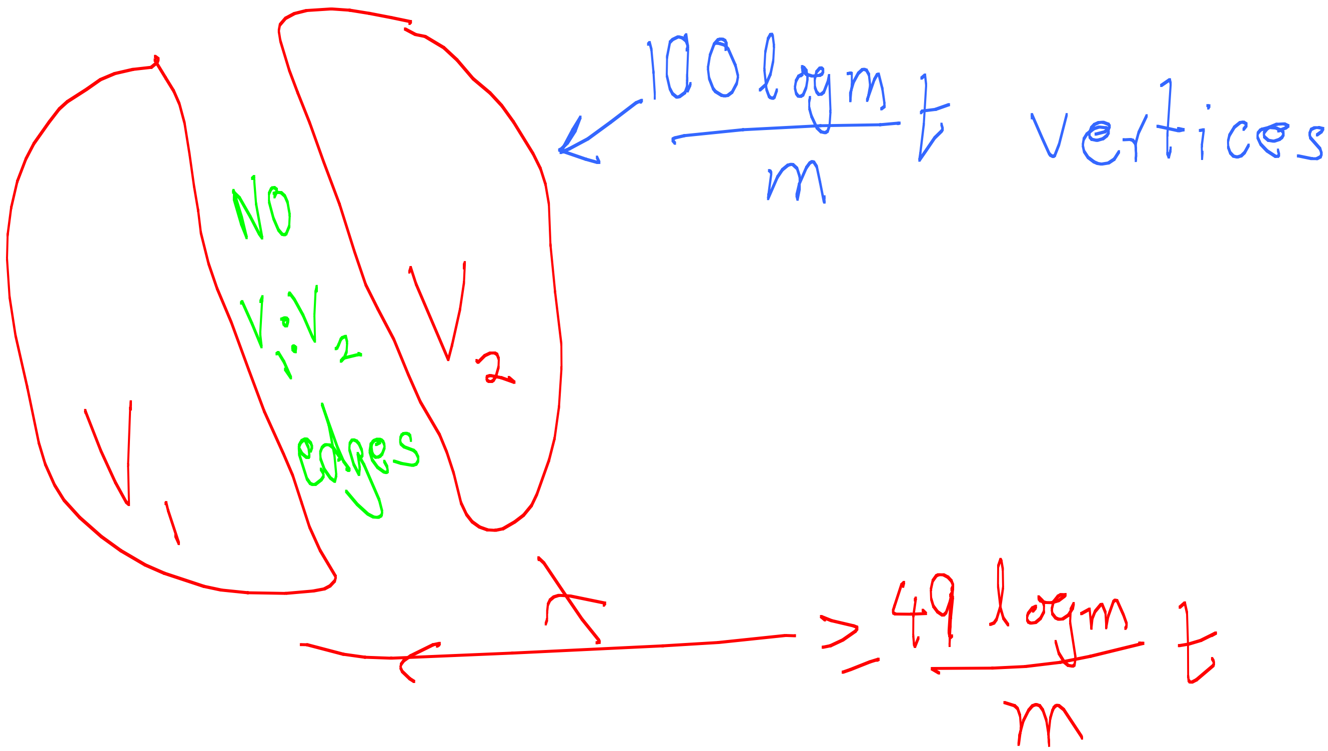
Half of Theorem 1 follows:

$\eta = \frac{6}{13}$.

Suppose in G_1

$\frac{100 \log m}{m} t$ vertices

NO COMPONENT $\geq \frac{2 \log m}{m} t$



Assume $e(G_2 \setminus G_1) \leq x = Ae^{-cm} t$:

So in G_2, $\exists\, V_1, V_2$ with $\leq x$ edges between them

$$\Pr\left(G_2 \text{ contains } V_1, V_2 \text{ such that } e(V_1 : V_2) \le x \right)$$

$$\le \binom{t}{\frac{49 \log m}{m} t}^2 \times$$

$$\Pr\left(B\left(\frac{(49)^2 (\log m)^2}{m^2} t^2, \frac{\eta m}{t} \right) \le A e^{-cmt} \right)$$

$$E(B) = \frac{(49)^2 \eta (\log m)^2}{m} t \qquad \le \frac{1}{100} E(B)$$

$$\Pr(\cdots) \le \exp\left\{ -\frac{1}{10} \cdot \frac{(49)^2 \eta (\log m)^2}{m} t \right\}$$

So $P_1(V_1, V_2 \text{ exist})$

$$\leq \left(\frac{em}{49 \log m}\right)^{\frac{98 \log m}{m} t} \exp\left\{-\frac{(49)^2 \eta (\log m)^2}{10 m} t\right\}$$

Take logs: Compare (for large m)

$$\frac{98 (\log m)^2}{m} t \quad v \quad \frac{(49)^2 \eta}{10} \frac{(\log m)^3}{m} t$$

RHS wins $\frac{49\eta}{10} > 2$

Theorem 3

Let $\epsilon > 0$ be given. $\exists$ constant C such that we can construct G_1, G_2 on same vertex set as before and such that

$$G_2 \sim G_{t,p}, \qquad p = \frac{Cm}{t}$$

$$G_2 \supseteq G_1 \setminus B \quad \text{(as a graph)}$$

B is a set of vertices such that

$$|\{i \in B : i \geq \epsilon t\}| \leq \epsilon \frac{t}{m}$$

$$\left(S \circ \quad |B| \leq \left(1 + \frac{1}{m}\right) \epsilon t \quad \right)$$

Other half of Theorem 1

$\epsilon = \frac{1}{2}$

Fact: $\alpha(G_{n,\frac{d}{n}}) \gtrsim \frac{2\log d}{d} n$ whp

whp $G_2 \setminus \{i \geq \frac{1}{2}t\}^*$ contains an independent set of size

$$\gtrsim \frac{2\log(\frac{1}{2}cm)}{\frac{1}{2}cm} \frac{t}{2}$$

* B may be correlated with G_2

So

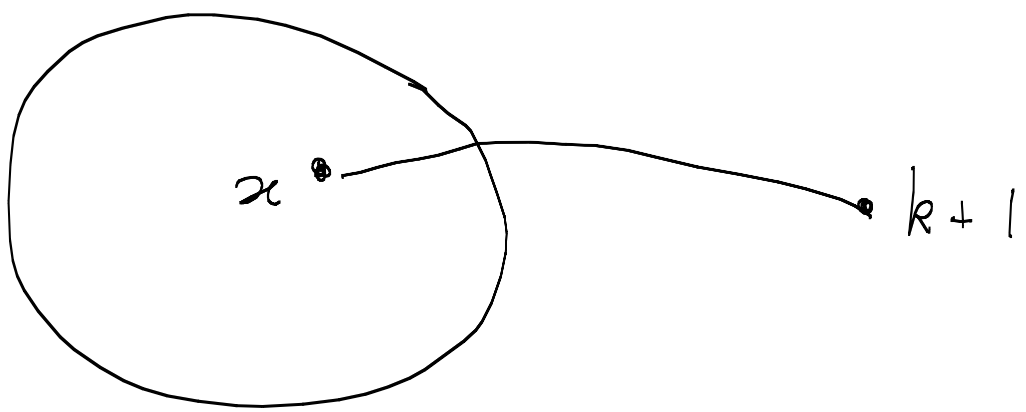
G_1 contains an independent set of size

$$\geq \sim \left(\frac{2 \log m}{m} - \frac{1}{2m} \right) t$$

Proof of Theorem 2

Assume $k \geq k_0 = \log t$.

Early construction $k < k_0$ not coupled.



G_1

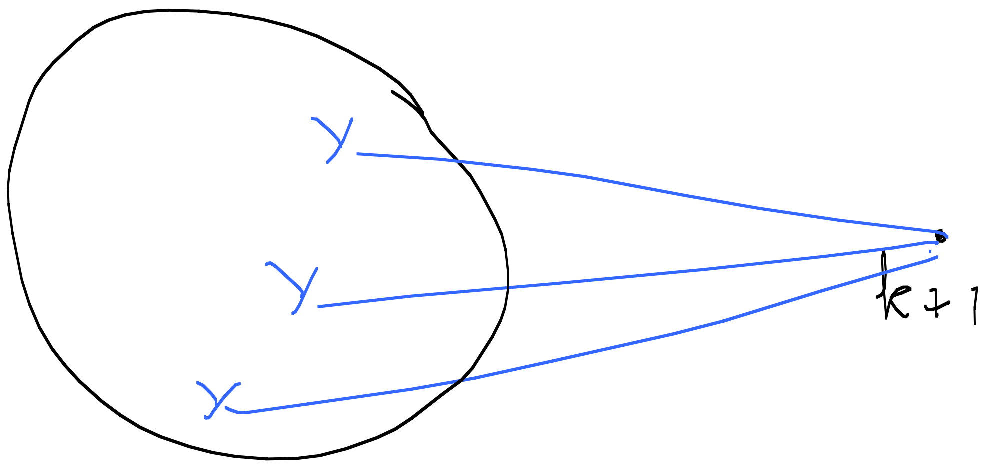
$$\Pr(x \text{ chosen}) \geq \frac{m}{2km+m} = \frac{1}{2k+1}$$

Choose $s_i \in [k] \cup \{\emptyset\}$, $i = 1, 2, \dots, m$

$$\Pr(s_i = j) = \frac{1}{2k+1}$$

$$X = \# i : s_i \neq \emptyset = B\left(m, \frac{k}{2k+1}\right)$$

Now finish addition of $k+1$ by adding $m-X$ more neighbours.



G_2

$$Y = B\left(k, \frac{2m}{t}\right)$$

$= \#$nbrs of $k+1$ in $[k]$ in $G_{t,2m/t}$

If $Y \le X$ choose random Y subset of $\{S_i \neq \emptyset\}$

If $Y > X$ add $Y-X$ random nbrs.

(i) $P(Y \leq X) \geq 1 - e^{-cm}$

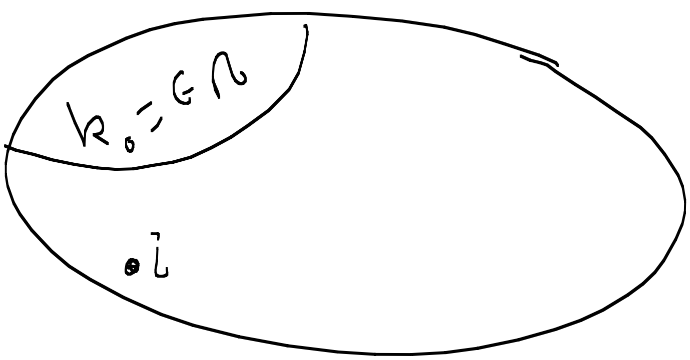
(ii) Whp $\sum (Y-X)^+ \leq t e^{-c'm}$

Chernoff

Proof of Theorem 3

G is given

A is some large constant.



$k+1$ $\quad G_1$

i is bad if degree $\geq Am$

i stays bad

whp # bad $\leq \dfrac{\epsilon t}{m}$

$$t \sum_{k \geq Am} \left[\frac{C_1}{k^3} + \frac{C_2}{k^4} + t^{-\theta} \right] < \frac{\epsilon t}{m}$$

• $k+1 : G_A$

Choices for $k+1$ are $t_1, t_2, \dots, t_m$

For $j \in V_k$

$$\Pr(t_q = j \mid t_1, \dots, t_{i-1}) \le \frac{Am+m}{2mk} \le \frac{A}{\epsilon t}$$

$$\Pr(t_i \notin V_k \mid t_1, \dots, t_{i-1}) \ge \frac{\epsilon t m}{2 t m} \ge \epsilon/2$$

• $k+1 : G_{\underline{A}}$

So

$$\Pr\left(t_i = j_r \,\middle|\, t_1, t_2 \ldots t_{i-1} \text{ and } t_i \notin \{j_1, \ldots, j_{r-1}\}\right)$$

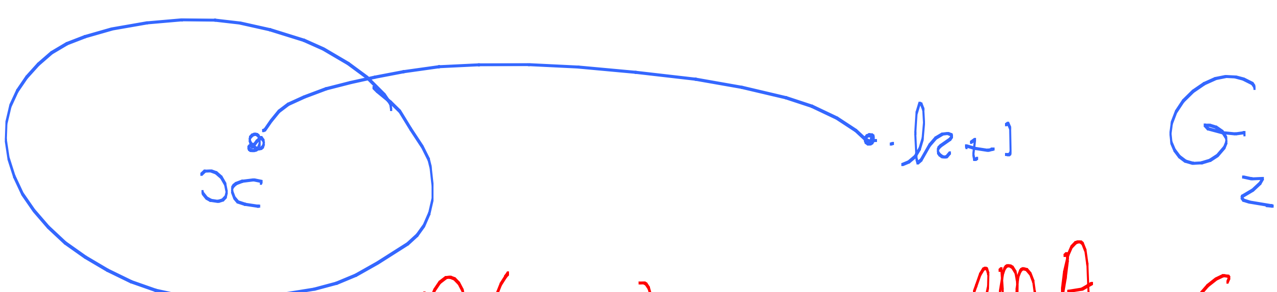
$$\leq \frac{\Pr(t_i = j_r \mid t_1, t_2, \ldots, t_{i-1})}{\Pr(t \notin V_k \mid t_1, t_2, \ldots, t_{i-1})}$$

$$\leq p = \frac{2A}{G^2 t}$$

Conclusion: can can construct random subset T_i, $|T_i| = B(k,p)$ such $t_i \in T_i$ whenever $t_i \in V_k$.

The T_i can be constructed independently.

The neighbours of $k+1$ in G_2 are $T_1 \cup T_2 \cup \dots \cup T_m$



$$Pr(\text{edge}) \leq mp = \frac{2mA}{\epsilon^2 t} = \frac{Cm}{t}$$

At end of construction only edges of G_1 not in G_2 are incident with

$$\underbrace{\{1, 2, \dots, k_0\}}_{\epsilon t} \cup \underbrace{\underline{BAD}}_{\le \frac{6}{m} t}$$