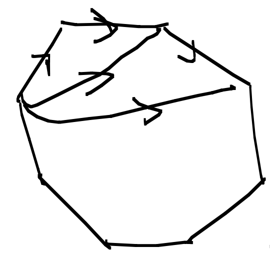
# Copying Model

Note Title 2/21/2007

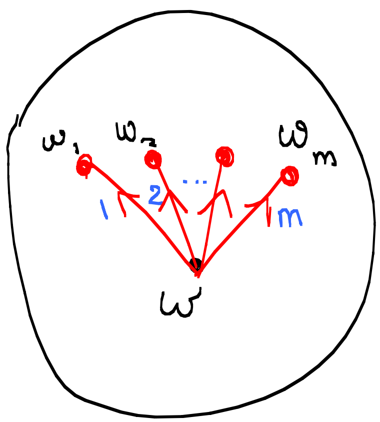
Vertices $v_1, v_2, \ldots \ldots v_t \ldots$

$D_{2m}$ =



some digraph with out-degree $m$. No loops or $\parallel$ edges.

$D_t$




$\cdot\, v_{t+1}$

$t \geq 2m$

$w$ chosen uniformly at random

$i = 1, 2, \ldots, m$

Edge $(v_{t+1}, x_i)$

$$x_i = \begin{cases} w_i : \text{Prob } 1-\alpha \\ \text{random : Prob } \alpha \\ \quad \text{from } [t] \setminus \{w_1, \ldots, w_m\} \end{cases}$$

$x_1, \ldots, x_m$ are distinct.

$N_{t,k}$ = # of vertices with indegree $k$ at time $t$.

$$E(N_{t+1,k} \mid D_t) = N_{t,k} + \frac{(1-\alpha)\left((k-1)N_{t,k-1} - kN_{t,k}\right)}{t}\left(1 + O\left(\tfrac{1}{t}\right)\right)$$

$$+ \frac{m\alpha\left(N_{t,k-1} - N_{t,k}\right)}{t}\left(1 + O\left(\tfrac{1}{t}\right)\right) + 1_{k=m}$$

accounts for making distinct choices.

$$\boxed{\begin{array}{l} N_{1,m} = 1 \\ N_{1,k} = 0,\ k \neq m \end{array}}$$

$$\overline{N}_{t+1,k} = \overline{N}_{t,k}\left(1 - \frac{k(1-\alpha) + m\alpha}{t}\right) + \overline{N}_{t,k-1}\left(\frac{(k-1)(1-\alpha) + m\alpha}{t}\right)$$

$$+ 1_{k=0} + O(1/t)$$

$$\overline{N}_{t+1,0} = \overline{N}_{t,0}\left(1 - \frac{m\alpha}{t}\right) + 1 + O(1/t)$$

$\left| \text{Assume} \leq \frac{A}{t} \right|$

<u>Lemma</u>

$$\left| \overline{N}_{t,0} - \frac{t+\alpha}{1+\alpha m} \right| \leq B \ln t, \quad B \text{ constant}$$

$B \gg A$

<u>Proof</u> True for small $t$.

<u>Upper Bound</u>

$$\left(\frac{t+\alpha}{1+\alpha m} + B\ln t\right)\left(1 - \frac{m\alpha}{t}\right) + 1 + \frac{A}{t} = \frac{t+\alpha-m\alpha}{1+\alpha m} + 1 - \frac{m\alpha^2}{t(1+\alpha m)} + (B\ln t)\left(1 - \frac{m\alpha}{t}\right) + \frac{A}{t}$$

$$\leq \frac{t+1+\alpha}{1+\alpha m} + B\ln(t+1).$$

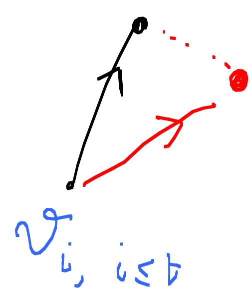
## Lower Bound

$$\left(\frac{t+\alpha}{1+\alpha m} - \beta \ln t\right)\left(1 - \frac{m\alpha}{t}\right) + 1 - \frac{A}{t}$$

$$= \frac{t+1+\alpha}{1+\alpha m} - \frac{m\alpha^2}{t(1+\alpha m)} - (\beta \ln t)\left(1 - \frac{m\alpha}{t}\right) - A/t \, .$$

$\square$

## Concentration

Changing one random choice  changes $N_{t,0}$ by $\leq 1$.

Azuma

$$P_r\left(\left|N_{t,0} - \bar{N}_{t,\theta}\right| \geq u\right) \leq \exp\left\{-\frac{2u^2}{mt}\right\}$$

$$\overline{N}_{t+1,k} = \overline{N}_{t,k}\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + \overline{N}_{t,k-1}\left(\frac{(k-1)(1-\alpha) + m\alpha}{t}\right) + \mathbb{1}_{k=0} + O(1/t).$$

Consider recurrence: (we show $\overline{N}_{t,k} \simeq a_k t$)

$$a_k\left(1 + k(1-\alpha) + m\alpha\right) = a_{k-1}\left((k-1)(1-\alpha) + m\alpha\right)$$

$$a_k = a_0 \prod_{i=1}^{k} \frac{(i-1)(1-\alpha) + m\alpha}{1 + i(1-\alpha) + m\alpha} \qquad a_0 = \frac{1}{1+\alpha m}$$

$$= a_0 \prod_{i=1}^{k}\left(1 - \frac{2-\alpha}{1+i(1-\alpha)+m\alpha}\right)$$

$$= a_0 \prod_{i=1}^{k} \left(1 - \frac{2-\alpha}{1+i(1-\alpha)+m\alpha}\right)$$

$$= a_0 \exp\left\{ - \sum_{i=1}^{k} \frac{2-\alpha}{1+i(1-\alpha)+m\alpha} - \xi_k \right\}$$

$\xi_k = O(1)$

$\xi_k \to \xi_\infty$

$$= \frac{C + O(1/k)}{k^{\frac{2-\alpha}{1-\alpha}}} \quad \text{as } k \to \infty.$$

$$\boxed{C = \text{constant} = C(\alpha, m)}$$

Now consider

$$\Delta_{k,t} = \overline{N}_{k,t} - a_k t$$

We show that this is $o(t)$. If $k \geq 1$

$$\overline{N}_{t+1,k} = \overline{N}_{t,k}\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + \overline{N}_{t,k-1}\left(\frac{(k-1)(1-\alpha)+m\alpha}{t}\right) + O(1/t)$$

or

$$\Delta_{k,t+1} + a_k(t+1) = \left(\Delta_{k,t} + a_k t\right)\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + \left(\Delta_{k-1,t} + a_{k-1} t\right)\left(\frac{(k-1)(1-\alpha)+m\alpha}{t}\right) + O(1/t)$$

or

$$\Delta_{k,t+1} = \Delta_{k,t}\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + \Delta_{k-1,t}\left(\frac{(k-1)(1-\alpha)+m\alpha}{t}\right) + O(1/t)$$

$$|\Delta_{k,t+1}| \le |\Delta_{k,t}|\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + |\Delta_{k-1,t}|\left(\frac{(k-1)(1-\alpha)+m\alpha}{t}\right) + \frac{L}{t}$$

<u>Claim</u>

$$|\Delta_{k,t}| = O(\ln t)$$

(1) $$|\Delta_{k,t}| \le \overline{N}_{t,k} + a_k t$$

$$\le \frac{mt}{k} + O\left(t / k^{\frac{2-\alpha}{1-\alpha}}\right)$$

$$\le At/k.$$

Done for $k \ge \frac{2t}{\ln t}$.

Assume $k \leq \frac{1}{2}(t+1)^{1/3}$

$$|\Delta_{k,t+1}| \leq |\Delta_{k,t}|\left(1 - \frac{k(1-\alpha)+m\alpha}{t}\right) + |\Delta_{k-1,t}|\left(\frac{(k-1)(1-\alpha)+m\alpha}{t}\right) + \frac{L}{t}$$

Assume $|\Delta_{k,t}| \leq M \ln t$ , $\forall k \leq \frac{2t}{\ln t}$

$$|\Delta_{k,t+1}| \leq M \ln t + L/t$$

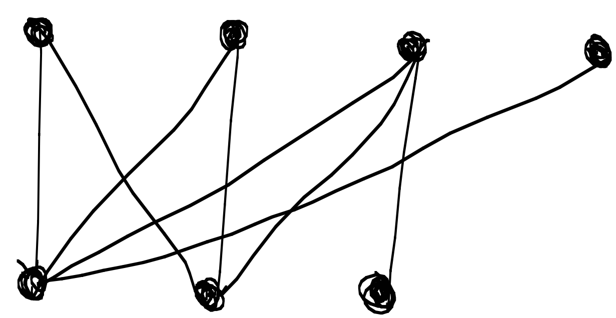
$$\leq M \ln t \qquad M \gg L.$$

## Summary

$$\overline{N}_{t,k} = \frac{C + O(1/k)}{k^{\frac{2-\alpha}{1-\alpha}}}\, t + O(\ln t).$$

# Complete Bipartite Subgraphs

Observations of web-graphs show large numbers of bipartite cliques



Whp Copy model contains many more than in $G_{n,p}$ $(p = \frac{2m}{n})$ or in PAM.

Consider $X_m$ = # copies of $K_{m,m}$.

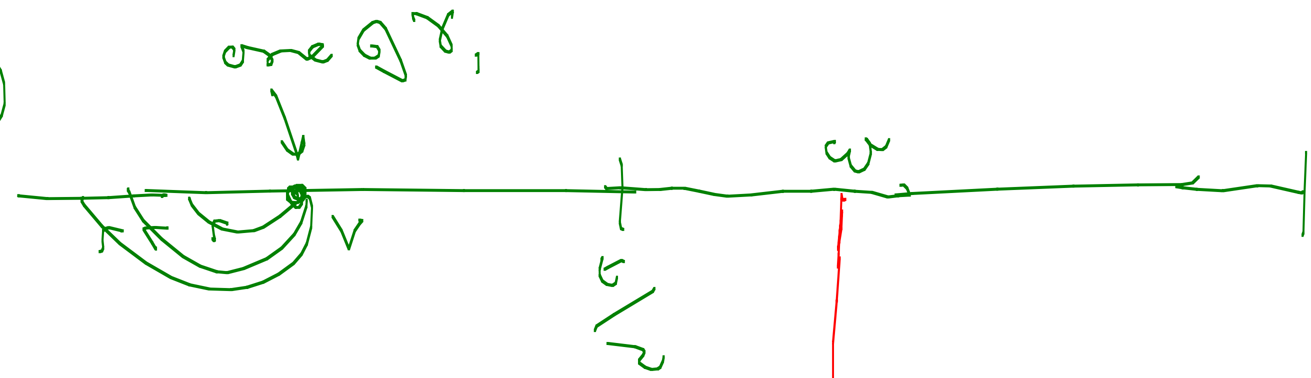
(i) Whp $\exists \geq \gamma_1 t$ vertices in $[t/4, t/2]$ with $m$ distinct out-neighbours.

$$\# \geq \mathrm{Bin}\left(\frac{t}{4}, \alpha^m \times \frac{1}{2}\right)$$

all random non-copying choices

distinct.

(ii)




$\text{Prob} \geq \frac{1}{t} \times (1-\alpha)^m$

of having same out neighbours as $v$.

So # of $w$ having same out neighbours is

$\geq Bin\left(t/2, \frac{1}{t}(1-\alpha)^m\right)$

$P_r(\geq m-1) \geq \gamma_2$.

So $E(\# v \text{ producing } K_{m,m}) \geq \gamma_1 \gamma_2 t$.

Use Azuma to show existence whp.