9/14/2020 Week 4.

Objection to ER: $G(n, p)$
unrealistic deg seq!

What if we fix degrees? $G(n, \vec{k})$
(but make it otherwise random?)
where $\vec{k} = \{k_i\}_{i=1}^{n}$ degrees.

[How much of some observed pattern is driven by (or could be explained by) the degrees alone?]

p_{ij} = Prob that $i \leftrightarrow j$.
(ER: $p_{ij} = p \ \forall i, j$)

Let p_{ij} depends on k_i, k_j

target.
goal.

From i's perspective:

$\frac{k_j}{2m}$ prob that one edge from i goes to j

that match to j

total # slots

$$p_{ij} = k_i \left(\frac{k_j}{2m}\right) = \frac{k_i k_j}{\sum_i k_i}$$

Chung-Lu

$$p_{ij} = \frac{k_i k_j}{2m}$$

$\forall i > j \quad A_{ij} = A_{ji} = \begin{cases} 1 & \text{w.p. } p_{ij} \\ 0 & \text{else.} \end{cases}$

- simple graph: no self loops
  no multi-edges.
- Requires $\mathcal{O}(n^2)$ rand. $\left[\binom{n}{2} \text{ calls}\right]$

Generalize:

- Directed Chung Lu? $G(n, k^{in}, k^{out})$

  $p_{ij} = \frac{k_i^{out} k_j^{in}}{m}$, $\forall i, j$ (both orders)

Problem:

$$p_{ij} \leq 1$$

$$\Rightarrow \frac{k_i k_j}{2m} \leq 1$$

Let $k_{max} \geq k_i$
$\forall i$

$$\Rightarrow \frac{{k^{max}}^2}{2m} \leq 1$$

$$k^{max} \leq \sqrt{2m}$$

$$k^{max} \ll n$$

Fix

$$p_{ij} = \min\left\{1, \frac{k_i k_j}{2m}\right\}$$

New Problem: Hubs are too small!

Consistency?

$$\underline{E[k_i]} = E\left[\sum_j A_{ij}\right]$$

$$= \sum_j E[A_{ij}]$$

$$A_{ij} = \begin{cases} 1 & \text{w.p. } p_{ij} \\ 0 & \text{else} \end{cases}$$

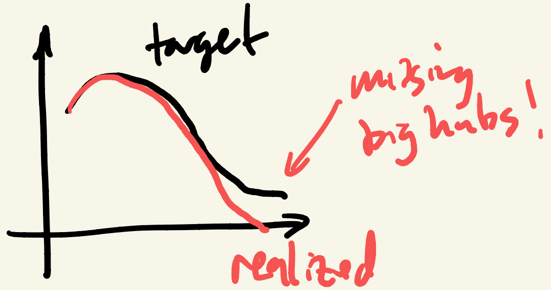
$\mathrm{Bern}(p_{ij})$

if ≤ 1 sub in $\min\{p_{ij}, 1\}$

$$= \sum_j p_{ij}$$

$$= \sum_j \frac{k_i k_j}{2m}$$

$$= \frac{k_i}{2m} \sum_j k_j = \frac{k_i}{\cancel{2m}} \cancel{2m} = k_i$$

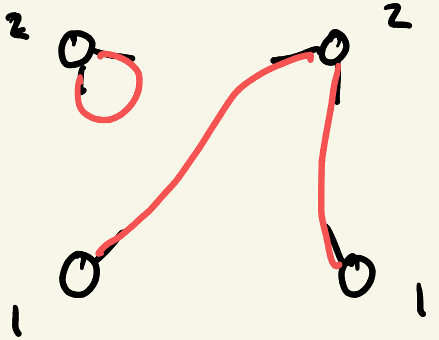


- Chung Lu gets degrees right <u>only</u> in expectation.

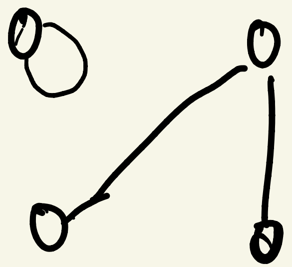
Configuration Model.

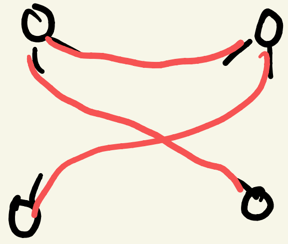
"stub matching"

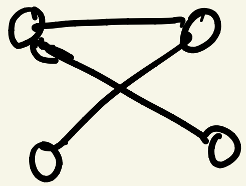
$\vec{k} = \{1, 2, 2, 1\}$



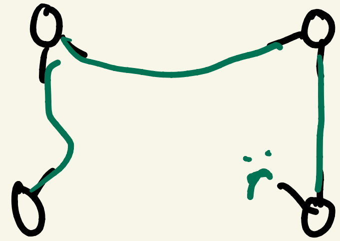
choose a pair, wire up!







$k = \{1, 2, 2, 2\}$



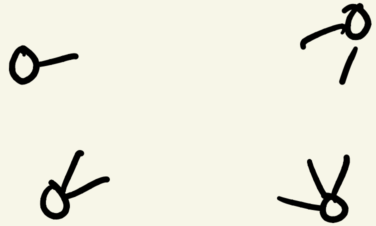
obviously even

Config Model:

$$2m = \sum_i k_i = \text{even}.$$

$k = \{1, 3, 2, 2\}$

1 2 3 4

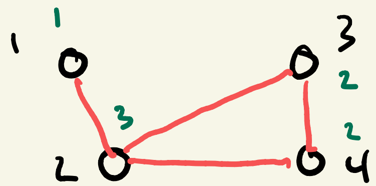


$1\ 2\ 2\ 2\ 3\ 3\ 4\ 4 = \vec{q}$

each node is in the list k_i times.

q.shuffle.

4 2 3 2 1 2 4 3

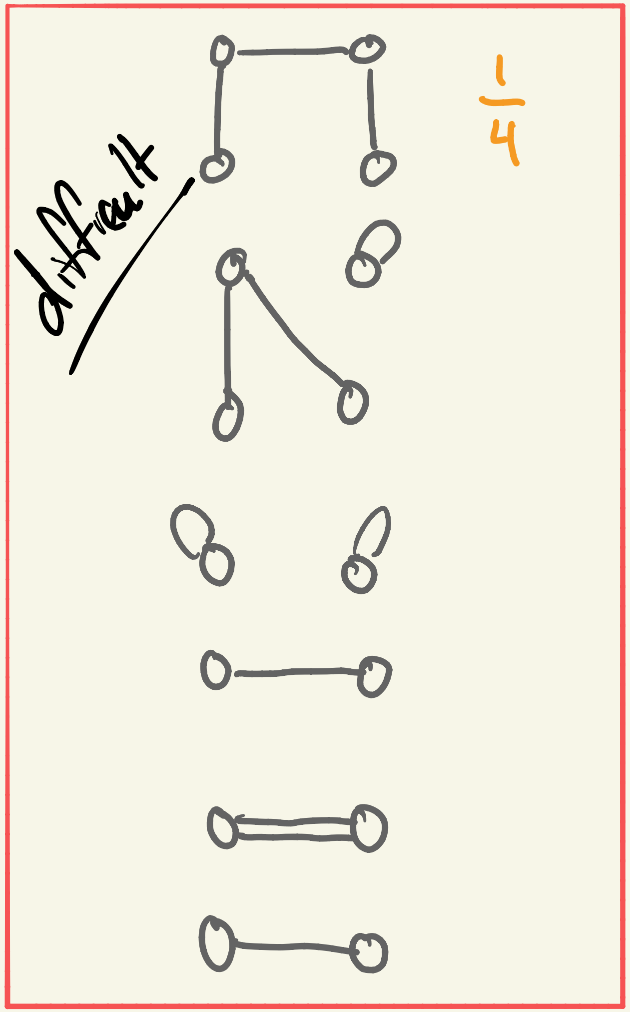


Problems! Next Time.

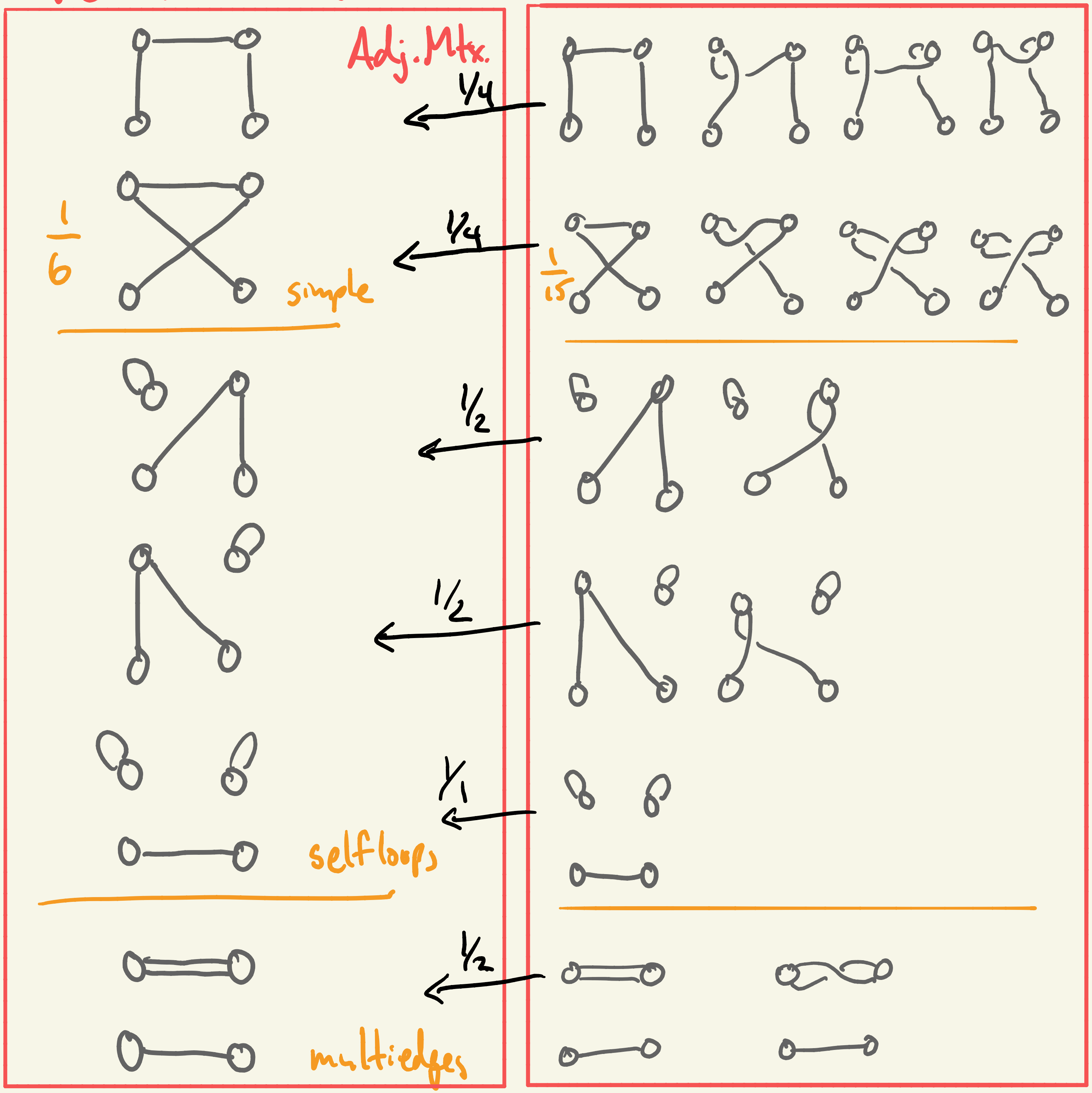
Problems:

$k = \{1, 2, 2, 1\}$

Graph Isomorphism

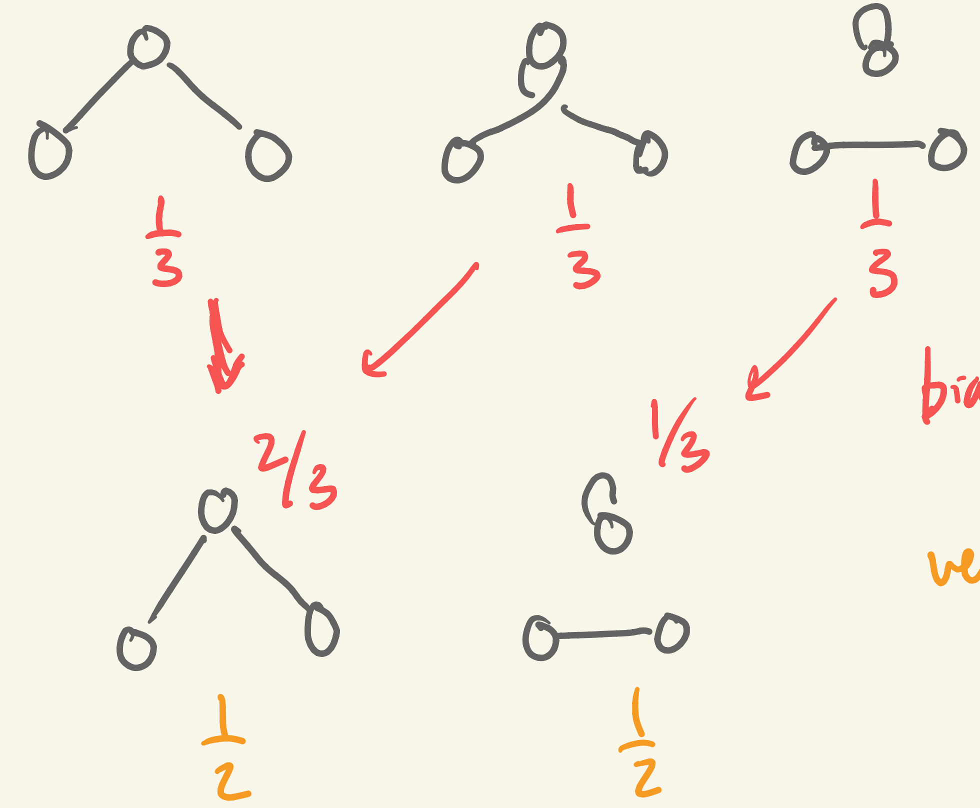


Vertex Labeled   Stub Labeled



Which space do I draw from when I stub match?

$k = \{1, 2, 1\}$



stub-labeled (stub-matching)

biased!

vertex labeled

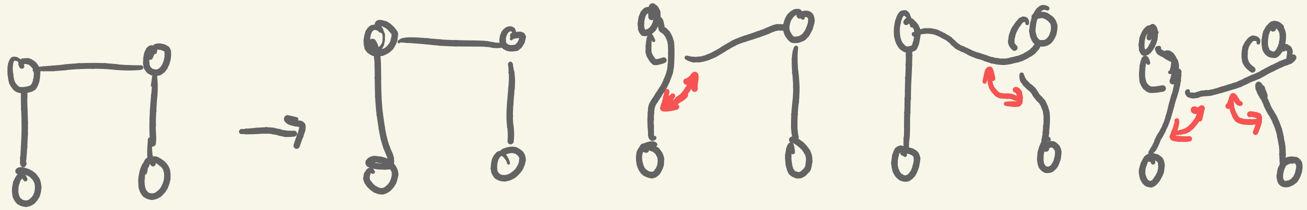
$\left\{\begin{array}{l}\text{Stub-labeled} \\ \text{Loopy} \\ \text{Multigraph}\end{array}\right.$

- Vertex-labeled multigraphs.
- Stub-labeled simple graphs.
- Stub-labeled loopy multigraphs.

sampled by stub matching

- Rejection sampling.
  ↳ inefficient
- Can we count configurations?

for a simple graph w/ deg. seq. $\vec{k}$, how many stub-labeled configurations correspond to each vertex-labeled graph?



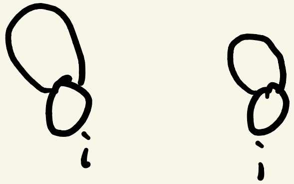
$$\boxed{q_{simple}(G) = \prod_{i=1}^{n} k_i!}$$

depends only on degs.

$1 \cdot 2! \cdot 2! \cdot 1 = 4$

There are exactly $q_{simple}(G)$ "stub-isomorphic" graphs in the stub-labeled space.

Let w_{ij} be the integer # of edges between i,j. For a single self loop, let $w_{ii} = 1$.



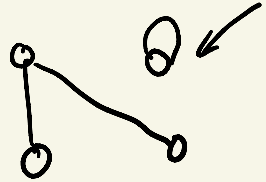
no switch possible.

$\Rightarrow$ each s.l. reduces perms by a factor of 2.

add'l factor lost.

$\prod_{i=1}^{n} w_{ii}! \, 2^{w_{ii}}$ lost

$$q_{loopy} = \frac{q_{simple}}{\prod_{i=1}^{n} w_{ii}! \, 2^{w_{ii}}}$$



$w_{ii} = 1$

$$q_{\text{multi}} = \frac{q_{\text{simple}}}{\prod_{i<j} w_{ij}!}$$

depends only on deg seq.

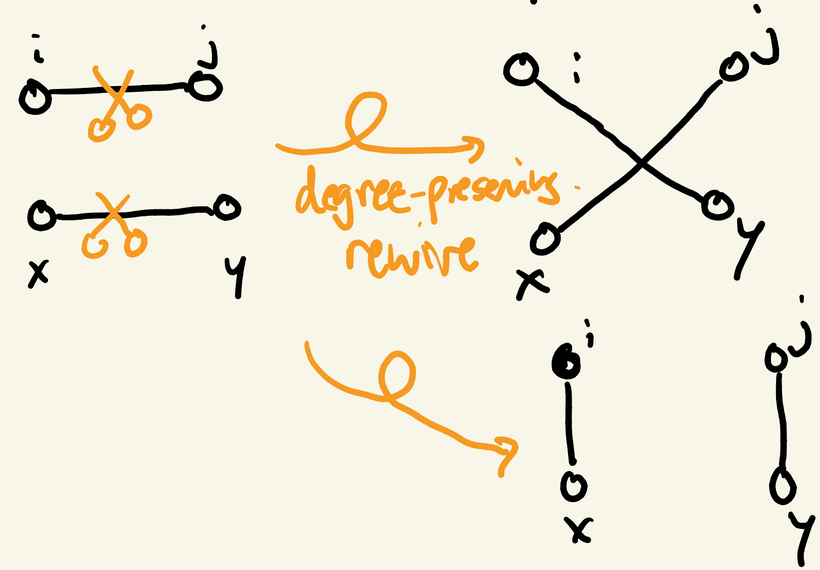
$$q_{\text{multi loopy}} = \frac{q_{\text{simple}}}{\prod_{i<j} w_{ij}! \; \prod_{i} w_{ii}! \; 2^{w_{ii}}}$$

depend on the actual config. (# SL, multiedges)

$\left.\begin{array}{l}\text{Rejection Sampling.}\\ \text{Down-sampling}\end{array}\right\}$ SLOW

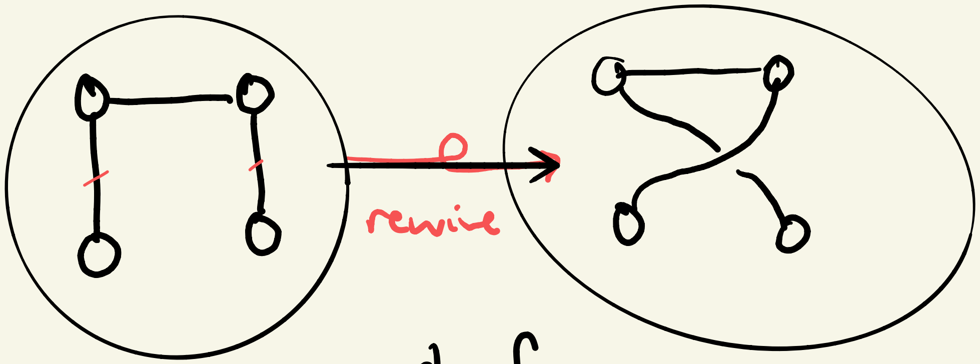
pain.

Double edge swap.



One graph $\vec{k}$ $\xrightarrow{\text{rewire}}$ another graph $\vec{k}$

1. choose a pair of distinct edge, u.a.r.
2. rewire
3. repeat 1, 2



graph of graphs.

$\boxed{\text{xzibit meme.}}$

			Markov Chain
node:	configuration	:	state
edge:	a double edge swap takes me from one graph to another	:	transition
adj.mtx:		:	transition matrix.

Markov Chain Monte Carlo?

Markov Chain → Wander around

Monte Carlo → and sample

Next: When/How/Why will this be uniform over

$\{\text{stub}, \text{vertex}\} \times \{\text{simple}, \text{loopy}, \text{multi}, \text{loopy multi}\}$

Hw Question

#2b: feel free to take $n \to \infty$

#2c: same.

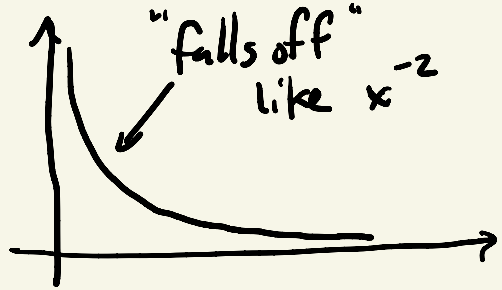
$$\frac{1}{2} \times \cancel{(x+10)} \frac{\alpha}{\cancel{x-1}} \xrightarrow[\text{like}]{\text{goes}} \frac{1}{2} \times \alpha$$

"grows as"

$\lim_{x \to \infty}$ key $x, x+10, x-1$

$$\lim_{x \to \infty} \cancel{17}\, x^{-2}$$

don't care



$$\lim_{x \to \infty} 17x^{-2} + \pi x^{-3}$$

which falls off faster?

$$\frac{1}{x^2} \quad \text{vs} \quad \frac{1}{x^3}$$

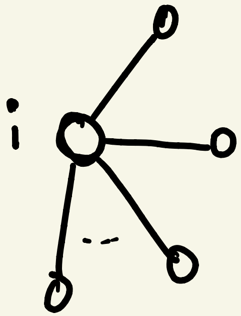
$$\frac{1}{(1000)^2} \quad \text{vs} \quad \frac{1}{(1000)^3}$$

small v.v.v. small

$$\lim_{x \to \infty} 17x^{-2} + \pi x^{-3} \xrightarrow[\text{as"}]{\text{"falls off}} x^{-2}$$

$$\langle C_m \rangle = \frac{\text{exp \# of pairs of nbrs of } i \text{ that are, themselves connected.}}{\text{exp \# of pairs of nbrs of } i}$$

$$\binom{a}{b} = \frac{a!}{b!\,(a-b)!}$$



What is the prob that two nodes are connected?

has k_i nbrs.

$\rightarrow$ has $\binom{k_i}{2}$ pairs of neighbors.

k_i choose 2

\# of ways to take k_i things and form pairs.

$$\Rightarrow \binom{k_i}{2} = \frac{k_i!}{2!\,(k_i-2)!}$$

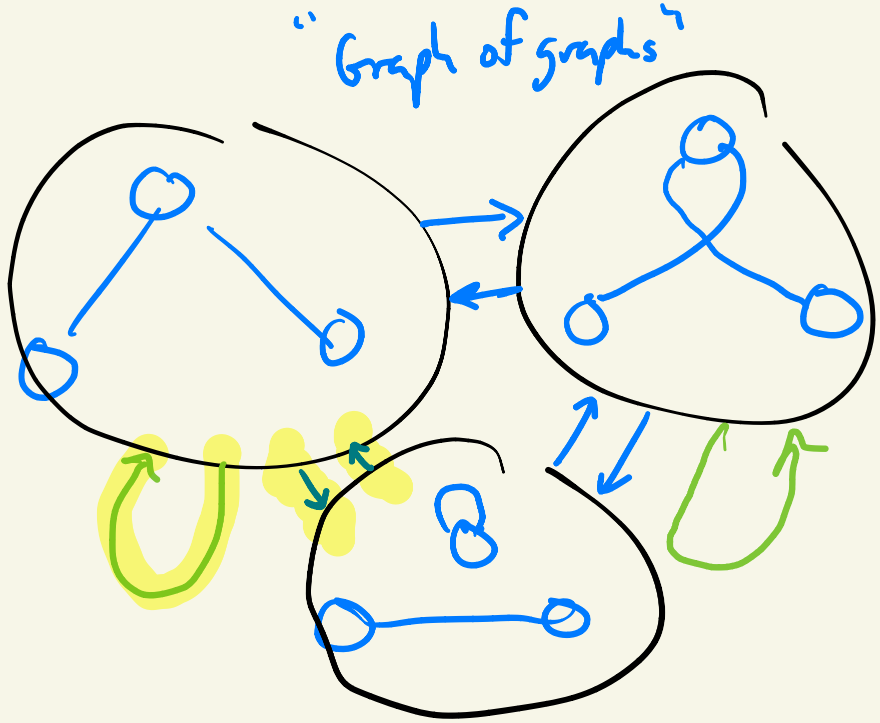
$$= \frac{k_i(k_i-1)\cancel{(k_i-2)!}}{2!\,\cancel{(k_i-2)!}}$$

$$\binom{k_i}{2} = \frac{k_i(k_i-1)}{2}$$

Goal MCMC:

Establish that if I do a Double Edge Swap. to "wander the space of $G(n,\vec{k})$" that I have equal prob of sampling each graph in that space.

$\Rightarrow$ uniform distrib.



Markov Chains!

Goal: stationary distrib. is unif.

1. transition matrix is doubly stochastic = ① GoG is degree-regular. ✓
2. irreducible = ② single component
3. aperiodic = ③ GCD cyclelength $=1$

Graph of Graphs

Fosdick CSU.
Ugander Stanford
Nishimura ASU

① If sample loopy multigraphs (stub labeled)

- stub matching
- MCMC D.E.S.

② If sample possibly loopy possibly multigraph (stub labeled)

IF in a proposed Double Edge Swap the resulting graph would have a forbidden property (loop, or multiedge)

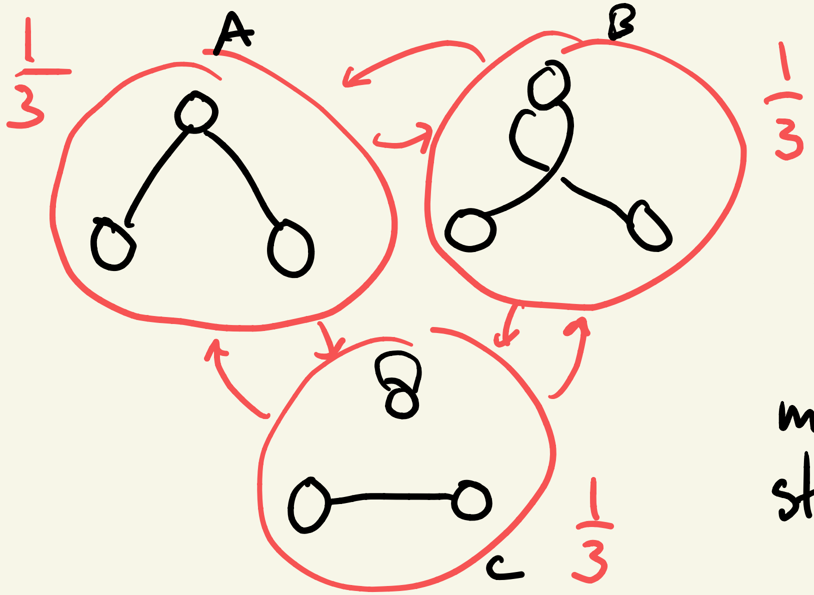
THEN: resample the current state in our graph of graphs. (see prev page in green)

③ One exception: loopy non-multi graphs.

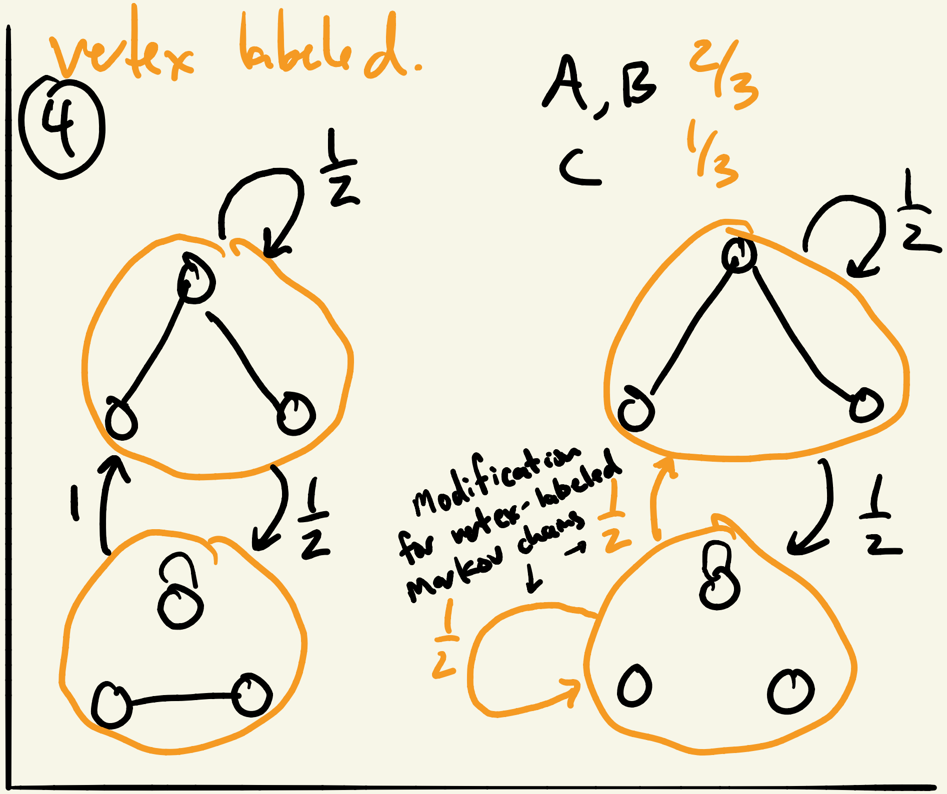
Counterexample. $\vec{k}$: $\{2, 2, 2\}$

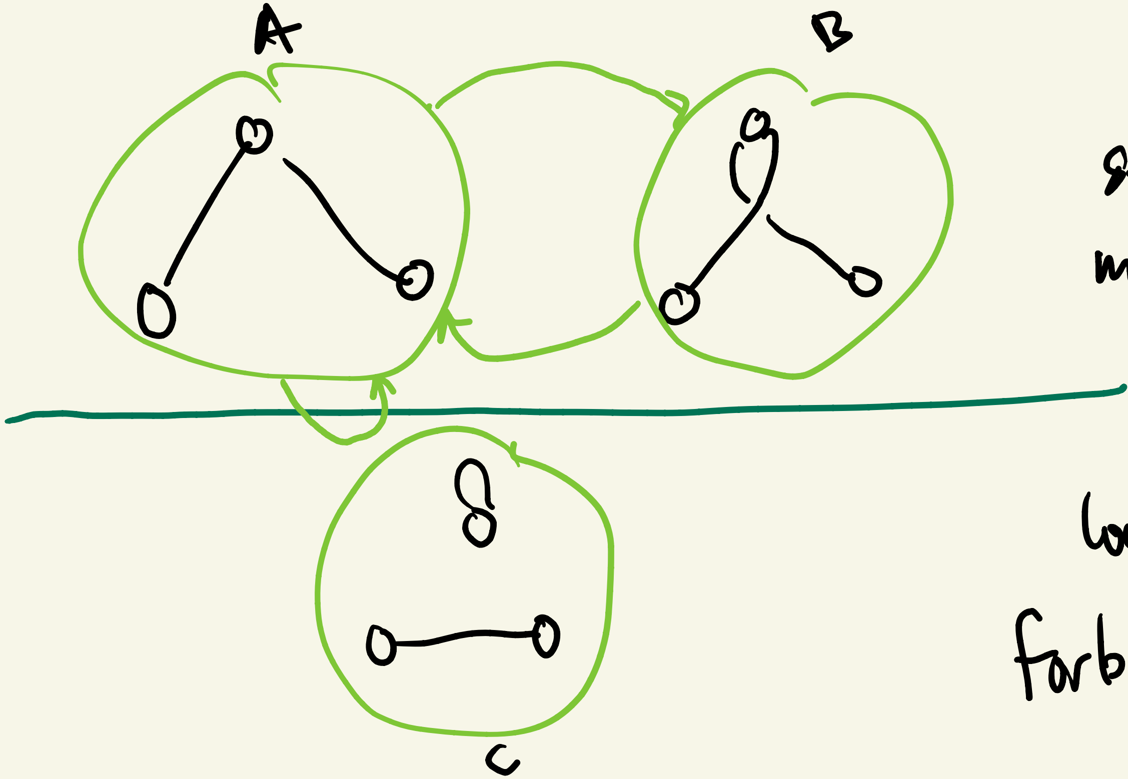
no double-edge swap!

solution: triple edge swap! Nishimura. followed up on this.



loopy multigraphs stub-labeled.





stublabeled multigraphs.

loopy! forbidden!

must resample.

A B A A

tried to leave. nope!

E.R. $G(n,p)$ (expectation)	E.R. $G(n,m)$ (exact)	density
Chung-Lu (expectation)	Config. Model (exact) stub matching	$\vec{k}$
expectation	exactly.	
"canonical"	"microcanonical"	Physicists.

$\boxed{\text{Stub matching}}$ $p_{ij} = \frac{k_i k_j}{2m}$ (4) see notes.

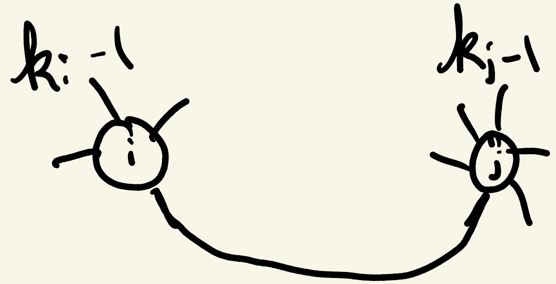
1. How many multiedges?

Under stub matching,

How many multiedges?

$$\Pr(1^{st} \text{ edge } i \leftrightarrow j) = \frac{k_i k_j}{2m}$$

$$\Pr(2^{nd} \text{ edge } i \leftrightarrow j) = \frac{(k_i - 1)(k_j - 1)}{2m}$$



$$\sum_{\substack{\text{distinct} \\ (i,j)}} \frac{k_i k_j}{2m} \frac{(k_i - 1)(k_j - 1)}{2m}$$

adjustment for self loops?

$$= \frac{1}{2} \frac{1}{(2m)^2} \sum_{ij} k_i(k_i - 1)\, k_j(k_j - 1)$$

$$= \frac{1}{2} \frac{1}{(2m)^2} \left(\sum_i k_i(k_i - 1)\right)\left(\sum_j k_j(k_j - 1)\right)$$

$$2m = \langle k \rangle n$$

$$= \frac{1}{2} \frac{1}{\langle k \rangle^2 n^2} \left(\sum_i k_i^2 - k_i\right)\left(\sum_j k_j^2 - k_j\right)$$

$$= \frac{1}{2} \frac{1}{\langle k \rangle^2} \left(\frac{1}{n}\sum_i k_i^2 - \frac{1}{n}\sum_i k_i\right)^2$$

$\langle k^2 \rangle$ $\quad$ $\langle k \rangle$

$$= \boxed{\frac{1}{2}\left[\frac{(\langle k^2 \rangle - \langle k \rangle)^2}{\langle k \rangle^2}\right]}$$

does not depend on n !!!

$\Rightarrow$ density of multiedges $\to 0$ as $n \to \infty$

How many selfloops by stub matching?

in expectation: $\dfrac{\langle k^2\rangle - \langle k\rangle}{2\langle k\rangle}$

density of self loops also goes to zero as n increases!

$$\left(p_{ii} = \frac{k_i(k_i-1)}{4m} \right)$$

How many common neighbors?

$i, j \qquad k_i, k_j$

If l is a nbr of i, j.

$\Rightarrow (i,l)$ and $(j,l) \in E$

$$n_{ij} = \sum_l \left(\frac{k_i k_l}{2m}\right)\left(\frac{k_j (k_l - 1)}{2m}\right)$$

$$= \frac{k_i k_j}{2m} \sum_l \frac{k_l(k_l-1)}{\langle k\rangle n}$$

$$= p_{ij} \frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}$$

more nbrs if you are also connected.

no dependence on n.

Giant Component?

E.R.

$c > 1$

Config Model?

$\langle k^2 \rangle - 2\langle k \rangle > 0$

For certain distributions with sufficiently heavy tails, $\langle k^2 \rangle \to \infty$ $\Rightarrow$ Always a giant component.

$\langle k^2 \rangle$ vs $\langle k \rangle^2$?

$\boxed{\text{Office Hours.}}$

$k = 1, 2, 3$

$\frac{1}{n}\left(1^2 + 2^2 + 3^2\right)$ $\left(\frac{1}{n}(1+2+3)\right)^2$

$\frac{1}{n}\sum_i k_i^2$ vs $\left(\frac{1}{n}\sum_i k_i\right)^2 \longrightarrow \left(\frac{1}{n}\sum_i t\right)^2 = (t)^2 = t^2$

Let $k_i = t$ $\frac{1}{n}\sum_i t^2 = t^2$

show!

Degree Regular $\Leftrightarrow \langle k^2 \rangle = \langle k \rangle^2$