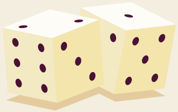
Complex networks & sparsity
Part II: Random models

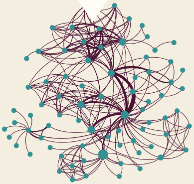


Felix Reidl
Blair D. Sullivan

DOCCOURSE '18

Graph classes vs. networks

Network instances



Mathematical Theory

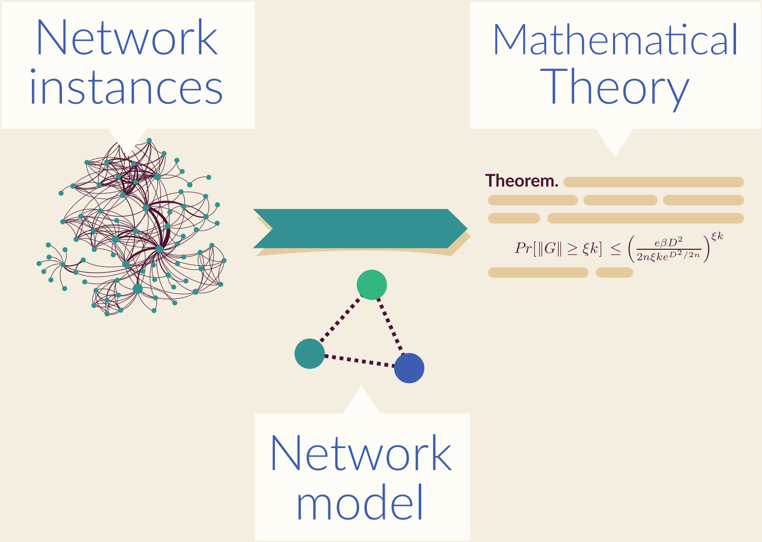
Theorem.

$$Pr[\|G\| \geq \xi k] \leq \left(\frac{e\beta D^2}{2n\xi k e^{D^2/2n}}\right)^{\xi k}$$

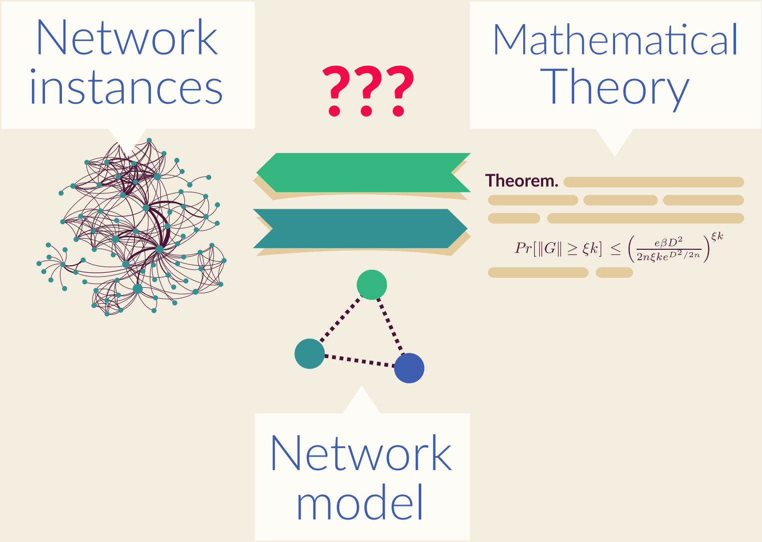
???

...

Graph classes vs. networks



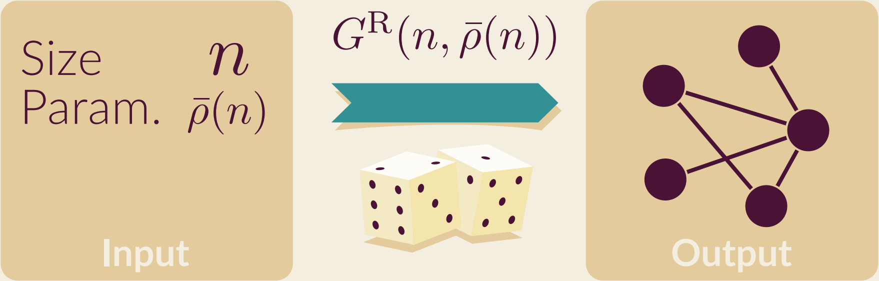
Graph classes vs. networks



Network models

A *network model* is a random process that generates *n*-vertex graphs*.

Besides the size *n* of the graph, a network model usually takes a vector of *parameters* $\bar{\rho}(n)$.



* In mathematical terms: a network model is a probability distribution over all *n*-vertex graphs.

Asymptotic bounded expansion

We know how to define bounded exp. for graph classes, but what about random graphs/models?

Def. A graph model $G^{\mathrm{R}}(n, \bar{\rho}(n))$ has bounded expansion *asymptotically almost surely* (a.a.s) if there exists a function f such that, for all r,

$$\lim_{n\to\infty} P\big[\widetilde{\nabla}_r\big(G^{\mathrm{R}}(n, \bar{\rho}(n))\big) < f(r)\big] = 1.$$

It has bounded expansion *with high probability* (w.h.p) if for every $c \geqslant 1$ there exists f s.t.

$$P\big[\widetilde{\nabla}_r\big(G^{\mathrm{R}}(n, \bar{\rho}(n))\big) < f(r)\big] \geqslant 1 - O(n^{-c}).$$

Asymptotic nowhere dense

Def. A graph model $G^{\mathrm{R}}(n, \bar{\rho}(n))$ is nowhere dense *asymptotically almost surely* (a.a.s) if there exists a function f such that, for all r,

$$\lim_{n\to\infty} P\big[\tilde{\omega}_r\big(G^{\mathrm{R}}(n, \bar{\rho}(n))\big) < f(r)\big] = 1.$$

It is nowhere dense *with high probability* (w.h.p) if for every $c \geqslant 1$ there exists f s.t.

$$P\big[\tilde{\omega}_r\big(G^{\mathrm{R}}(n, \bar{\rho}(n))\big) < f(r)\big] \geqslant 1 - O(n^{-c}).$$

$$\tilde{\omega}_r(G) = \omega(G \,\tilde{\nabla}\, r)$$

Asymptotic denseness

Def. A graph model $G^{\mathrm{R}}(n, \bar{\rho}(n))$ is *a.a.s. somewhere dense* if there exists r such that for all functions f it holds that

$$\lim_{n \to \infty} P\big[\tilde{\omega}_r\big(G^{\mathrm{R}}(n, \bar{\rho}(n))\big) > f(r)\big] = 1.$$

$$\tilde{\omega}_r(G) = \omega(G \,\tilde{\nabla}\, r)$$

Asymptotic denseness

Def. A graph model $G^{\mathrm{R}}(n, \bar{\rho}(n))$ is *a.a.s. somewhere dense* if there exists r such that for all functions f it holds that

$$\lim_{n\to\infty} P\big[\tilde{\omega}_r\big(G^{\mathrm{R}}(n,\bar{\rho}(n))\big) > f(r)\big] = 1.$$

Is is *not a.a.s. nowhere dense* if it only holds that

$$\lim_{n\to\infty} P\big[\tilde{\omega}_r\big(G^{\mathrm{R}}(n,\bar{\rho}(n))\big) > f(r)\big] > 0.$$

$$\tilde{\omega}_r(G) = \omega(G \,\tilde{\nabla}\, r)$$

(uniformly) random graphs

The sparse Erdős–Rényi model

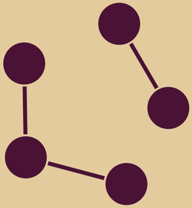
The mother of all random graphs!

Size n
Density μ

Input

$G(n, \frac{\mu}{n})$

$\frac{\mu}{n}$



Output

The Erdős–Rényi model

Erdős–Rényi graphs famously exhibit *phase transitions*, meaning that some property suddenly changes when we vary the parameter around a critical point.

There is a huge body of work on properties of Erdős–Rényi graphs!

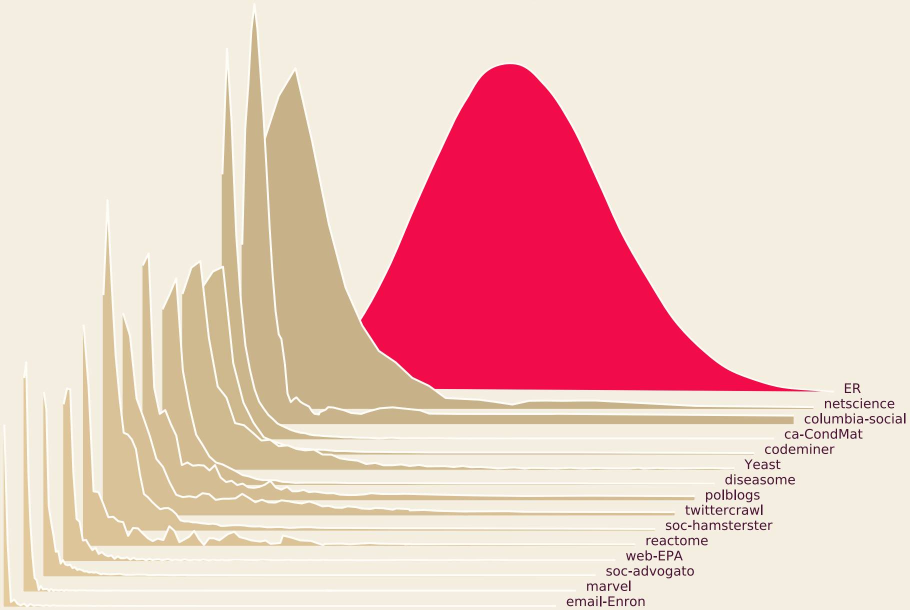
The Erdős–Rényi model

Erdős–Rényi graphs famously exhibit *phase transitions*, meaning that some property suddenly changes when we vary the parameter around a critical point.

Jarik & Patrice:

Sparse Erdős–Rényi graphs have asymptotically almost surely bounded expansion.

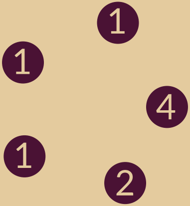
The Erdős–Rényi model



…but I want my powerlaw…

The Chung–Lu model

Generate graphs with a given degree distribution (in expectation).

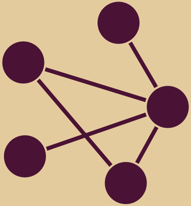


Degree distribution
$D_n = (\omega_1, \ldots, \omega_n)$

Input

$$G^{\mathrm{CL}}(D_n)$$

$$\frac{\omega_s \omega_t}{\sum_i \omega_i} = \frac{\omega_s \omega_t}{\mu n}$$



Output

Phase transition of Chung–Lu

Thm. Let $(D_n)_{n \in \mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CL}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Phase transition of Chung–Lu

The phase transition exists because of how generalized harmonic numbers behave:

Lemma 4. *For all integers* $0 < \delta \leqslant \Delta$ *the bound* $R_\gamma \leqslant \sum_{k=\delta}^{\Delta} \frac{1}{k^\gamma} \leqslant \frac{1}{\delta^\gamma} + R_\gamma$ *holds where*

$$R_\gamma = \begin{cases} \frac{1}{\gamma-1}(\delta^{1-\gamma} - \Delta^{1-\gamma}) & \textit{for } \gamma > 1, \\ \ln \Delta - \ln \delta & \textit{for } \gamma = 1, \textit{ and} \\ \frac{1}{1-}(\Delta^{1-\gamma} - \delta^{1-\gamma}) & \textit{for } 0 < \gamma < 1. \end{cases}$$

Lemma 5. *For* $\gamma > 0$, $r \geqslant 1$ *and integers* $r^{2r} < \delta \leqslant \Delta$ *the bound*

$$R'_\gamma \leqslant \sum_{k=\delta}^{\Delta} \frac{\ln^r k}{k^\gamma} \leqslant \frac{\ln^r \delta}{\delta^\gamma} + \zeta R'_\gamma$$

holds where ζ *is a constant and*

$$R'_\gamma = \begin{cases} \frac{1}{\gamma-1}(\delta^{1-\gamma} \ln^r \delta - \Delta^{1-\gamma} \ln^r \Delta) & \textit{for } \gamma > 1, \\ \frac{1}{\zeta(r+1)}(\ln^{r+1} \Delta - \ln^{r+1} \delta) & \textit{for } \gamma = 1, \textit{ and} \\ \frac{1}{(1-\gamma)^{r+1}}(\Delta^{1-\gamma} \ln^r \Delta - \delta^{1-\gamma} \ln^r \delta) & \textit{for } 0 < \gamma < 1. \end{cases}$$

Phase transition of Chung–Lu

The phase transition exists because of how generalized harmonic numbers behave:

Lemma 4. *For all integers* $0 < \delta \leqslant \Delta$ *the bound* $R_\gamma \leqslant \sum_{k=\delta}^{\Delta} \frac{1}{k^\gamma} \leqslant \frac{1}{\delta^\gamma} + R_\gamma$ *holds where*

$$R_\gamma = \begin{cases} \frac{1}{\gamma-1}(\delta^{1-\gamma} - \Delta^{1-\gamma}) & \text{for } \gamma > 1, \\ \ln \Delta - \ln \delta & \text{for } \gamma = 1, \text{ and} \\ \frac{1}{1-\gamma}(\Delta^{1-\gamma} - \delta^{1-\gamma}) & \text{for } 0 < \gamma < 1. \end{cases}$$

Lemma 5. *For* $\gamma > 0$, $r \geqslant 1$ *and integers* $r^{2r} < \delta \leqslant \Delta$ *the bound*

$$R'_\gamma \leqslant \sum_{k=\delta}^{\Delta} \frac{\ln^r k}{k^\gamma} \leqslant \frac{\ln^r \delta}{\delta^\gamma} + \zeta R'_\gamma$$

holds where ζ *is a constant and*

$$R'_\gamma = \begin{cases} \frac{1}{\gamma-1}(\delta^{1-\gamma} \ln^r \delta - \Delta^{1-\gamma} \ln^r \Delta) & \text{for } \gamma > 1, \\ \frac{1}{\zeta(r+1)}(\ln^{r+1} \Delta - \ln^{r+1} \delta) & \text{for } \gamma = 1, \text{ and} \\ \frac{1}{(1-\gamma)^{r+1}}(\Delta^{1-\gamma} \ln^r \Delta - \delta^{1-\gamma} \ln^r \delta) & \text{for } 0 < \gamma < 1. \end{cases}$$

Phase transition of Chung–Lu

Thm. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CL}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Degree distributions

Def. An *n-vertex degree distribution* D_n is a rand. variable with probability mass function f_n s.t.

1. $f_n(d) = 0$ for $d \notin [1, n-1]$,
2. $nf_n(d) \in \mathbb{N}_0$ for all $d \in \mathbb{N}$.

An n-vertex graph G *matches* D_n if the number of vertices of degree d in G is $nf_n(d)$.

If G matches D_n, then the expected value of D_n is precisely the average degree of G:

$$\bar{d}(G) = \frac{1}{n}\sum_{i=1}^{n} |\{v \mid d(v) = i\}| = E[D_n]$$

Degree distribution sequence

In order to talk about degree distr's of classes, we need a notion of *asymptotic* degree distr's.

Def. We call a sequence $(D_n)_{n\in\mathbb{N}}$ of *n*-vertex degree distr's a *degree distribution sequence*. The limit of (D_n) is a random variable D with

$$(D_n) \xrightarrow{d} D.$$

We say that (D_n) is *sparse* if

$$E[D] < \infty \quad \text{and} \quad (E[D_n]) \to E[D].$$

Imagine the limit of a degree distr sequence as its overall 'shape'.

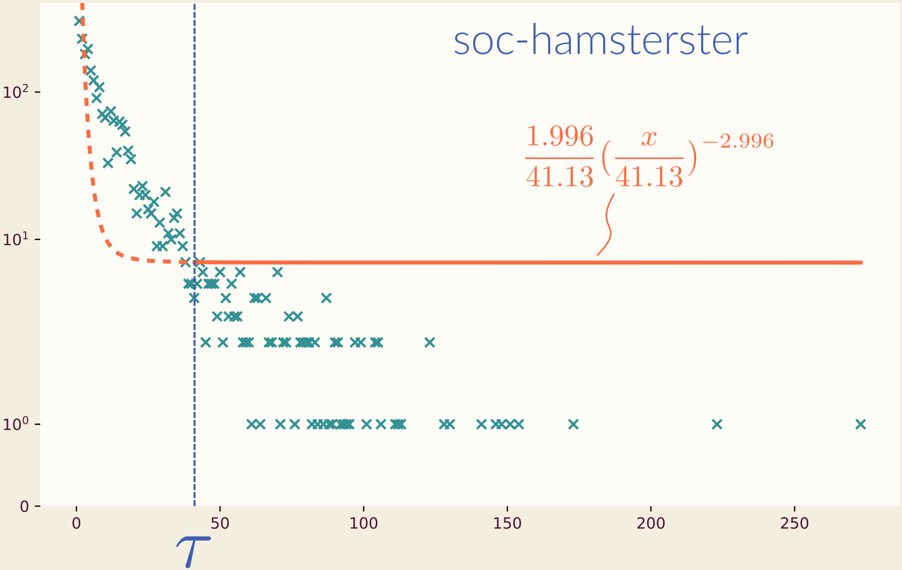
Tail-bounds

Def. A DDS $(D_n)_{n\in\mathbb{N}}$ with limit D has h as an *upper* (*lower*) *tail-bound* if there exists a constant $\tau \geqslant 0$ such that, for all $d \geqslant \tau$ and large enough n, it holds that

$$P[D_n \geqslant d] = O\Big(\frac{1}{h(d)}\Big). \quad \Big(\Omega\Big(\frac{1}{h(d)}\Big)\Big)$$

If a DDS has a lower and an upper tail bound, we simply call this function the *tail*.

Tail-bounds



soc-hamsterster
$\frac{1.996}{41.13}\left(\frac{x}{41.13}\right)^{-2.996}$
10^2
10^1
10^0
0
0
50
100
150
200
250
τ

Thm. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CL}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Supercubic regime: 1st ingredient

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-3-\epsilon})$. Then the probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most

$$c_r \frac{\omega_s \omega_t}{\mu n}$$

for some constant c_r.

Supercubic regime: 1st ingredient

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-3-\epsilon})$. Then the probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most

$$c_r \frac{\omega_s \omega_t}{\mu n}$$

for some constant c_r.

Probability of a dense r-shallow minor
$\approx$ probability of a dense subgraph!

Supercubic regime: 1st ingredient

Probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most $c_r \frac{\omega_s \omega_t}{\mu n}$.

$$\mathrm{P}\left[\, P_{st} \subseteq G \mid d_s, d_t \,\right] = \binom{n}{r-1} \frac{d_s d_t \prod_i d_i^2}{\mu^r n^r} \, \mathrm{P}[\mathbb{D}^r = (d_1, \ldots, d_{r-1})]$$

Supercubic regime: 1st ingredient

Probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most $c_r \frac{\omega_s \omega_t}{\mu n}$.

$$\mathrm{P}\left[\, P_{st} \subseteq G \mid d_s, d_t \,\right] = \binom{n}{r-1} \frac{d_s d_t \prod_i d_i^2}{\mu^r n^r} \mathrm{P}[\mathbb{D}^r = (d_1, \dots, d_{r-1})]$$

$$\begin{aligned}
\prod_i d_i^2 \cdot \mathrm{P}\Big[\bigwedge_i \hat{D}_{i,n} = d_i\Big] &\leqslant \sum_{d_1,\dots,d_{r-1}} \prod_i d_i^2 \cdot \mathrm{P}[\hat{D}_{i,n} = d_i] \\
&= \prod_{d_1,\dots,d_{r-1}} \sum_i d_i^2 \cdot \mathrm{P}[\hat{D}_{i,n} = d_i] = \prod_{d_1,\dots,d_{r-1}} \mathrm{E}[\hat{D}_{i,n}^2] \\
&= \Theta(\mathrm{E}[D_n^2]^{r-1}).
\end{aligned}$$

Supercubic regime: 1st ingredient

Probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most $c_r \frac{\omega_s \omega_t}{\mu n}$.

$$\mathrm{P}\big[\, P_{st} \subseteq G \mid d_s, d_t \,\big] = \binom{n}{r-1} \frac{d_s d_t \prod_i d_i^2}{\mu^r n^r} \, \mathrm{P}[\mathbb{D}^r = (d_1, \ldots, d_{r-1})]$$

$$\begin{aligned}
\prod_i d_i^2 \cdot \mathrm{P}\big[\bigwedge_i \hat{D}_{i,n} = d_i\big] &\leqslant \sum_{d_1,\ldots,d_{r-1}} \prod_i d_i^2 \cdot \mathrm{P}[\hat{D}_{i,n} = d_i] \\
&= \prod_{d_1,\ldots,d_{r-1}} \sum_i d_i^2 \cdot \mathrm{P}[\hat{D}_{i,n} = d_i] = \prod_{d_1,\ldots,d_{r-1}} \mathrm{E}[\hat{D}_{i,n}^2] \\
&= \Theta(\mathrm{E}[D_n^2]^{r-1}).
\end{aligned}$$

$$\mathrm{P}[P_{st} \subseteq G \mid d_s, d_t \,] \leqslant \binom{n}{r-1} \frac{d_s d_t}{\mu^r n^r} \Theta(\mathrm{E}[D_n^2]^{r-1}) = \frac{d_s d_t}{\mu n} O(\mathrm{E}[D_n^2]^{r-1})$$

Supercubic regime: 1st ingredient

Probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most $c_r \frac{\omega_s \omega_t}{\mu n}$.

$$\mathrm{P}\left[\, P_{st} \subseteq G \mid d_s, d_t \,\right] = \binom{n}{r-1} \frac{d_s d_t \prod_i d_i^2}{\mu^r n^r} \, \mathrm{P}[\mathbb{D}^r = (d_1, \ldots, d_{r-1})]$$

$$E[D_n^2] = \sum_{d=1}^{\Delta} \frac{d^2}{d^{3+\varepsilon}} = \sum_{d=1}^{\Delta} \frac{1}{d^{1+\varepsilon}} = O(1)$$

$$\mathrm{P}[P_{st} \subseteq G \mid d_s, d_t \,] \leqslant \binom{n}{r-1} \frac{d_s d_t}{\mu^r n^r} \Theta(\mathrm{E}[D_n^2]^{r-1}) = \frac{d_s d_t}{\mu n} O(\mathrm{E}[D_n^2]^{r-1})$$

Supercubic regime: 2nd ingredient

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-3-\epsilon})$. Then for $\xi \geqslant \Theta(1)$, $c \geqslant 2e$ and $n \geqslant 4\xi$ it holds that

$$\mathrm{P}\left[\,\exists H \subseteq G^{CL}(D_n)) : |H| \leqslant n/c \textit{ and } \nabla_0(H) \geqslant \xi\,\right] \leqslant \frac{1}{n^{\xi}}.$$

Supercubic regime: 2nd ingredient

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-3-\epsilon})$. Then for $\xi \geqslant \Theta(1)$, $c \geqslant 2e$ and $n \geqslant 4\xi$ holds that

$$\mathrm{P}\big[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \textit{ and } \nabla_0(H) \geqslant \xi\,\big] \leqslant \frac{1}{n^{\xi}}.$$

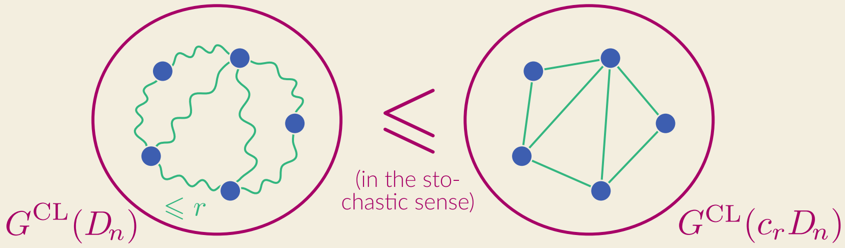
Probability of dense subgraphs is low.

Supercubic regime: 2nd ingredient

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-3-\epsilon})$. Then for $\xi \geqslant \Theta(1)$, $c \geqslant 2e$ and $n \geqslant 4\xi$ holds that

$$\mathrm{P}\left[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \text{ and } \nabla_0(H) \geqslant \xi\,\right] \leqslant \frac{1}{n^{\xi}}.$$

Probability of dense subgraphs is low.



Supercubic regime: 2nd ingredient

$$\mathrm{P}\big[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \;\; \mathit{and} \;\; \nabla_0(H) \geqslant \xi\,\big] \leqslant \frac{1}{n^{\xi}}$$

$$\sum_{k=2\xi}^{n/c} \binom{n}{k} \mathrm{P}\big[\, \|G^{\mathrm{CL}}(D_n)[X_k])\| \geqslant k\xi\,\big]$$

Union bound: number of dense subgraphs

Supercubic regime: 2nd ingredient

$$\mathrm{P}\left[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \;\; and \;\; \nabla_0(H) \geqslant \xi\,\right] \leqslant \frac{1}{n^{\xi}}$$

$$\sum_{k=2\xi}^{n/c} \binom{n}{k} \mathrm{P}\left[\,\|G^{\mathrm{CL}}(D_n)[X_k])\| \geqslant k\xi\,\right]$$

$$\leqslant \sum_{k=2\xi}^{n/c} \binom{n}{k} \sum_{d=k}^{\Delta k} \left(\frac{ecd^2}{2n\xi k e^{d^2/2n}}\right)^{\xi k} \mathrm{P}[\mathbb{D} = d]$$

Group subgraphs by sum of weights $\mathbb{D}$

Probability of dense subgraph, given $\mathbb{D}$

Supercubic regime: 2nd ingredient

$$\mathrm{P}\big[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \ \textit{and} \ \nabla_0(H) \geqslant \xi\,\big] \leqslant \frac{1}{n^{\xi}}$$

$$\sum_{k=2\xi}^{n/c} \binom{n}{k} \mathrm{P}\big[\,\|G^{\mathrm{CL}}(D_n)[X_k])\| \geqslant k\xi\,\big]$$

$$\leqslant \sum_{k=2\xi}^{n/c} \binom{n}{k} \sum_{d=k}^{\Delta k} \left(\frac{ecd^2}{2n\xi k e^{d^2/2n}}\right)^{\xi k} \mathrm{P}[\mathbb{D} = d]$$

$$\vdots$$

(a lot of work)

$$\leqslant \sum_{k=2\xi}^{n/c} \left(\frac{k}{n}\right)^{\xi k} \frac{n^k}{k^k}$$

Supercubic regime: 2nd ingredient

$$\mathrm{P}\left[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \;\; and \;\; \nabla_0(H) \geqslant \xi\,\right] \leqslant \frac{1}{n^{\xi}}$$

$$\sum_{k=2\xi}^{n/c} \binom{n}{k} \mathrm{P}\left[\,\|G^{\mathrm{CL}}(D_n)[X_k])\| \geqslant k\xi\,\right]$$

$$\leqslant \sum_{k=2\xi}^{n/c} \binom{n}{k} \sum_{d=k}^{\Delta k} \left(\frac{ecd^2}{2n\xi k e^{d^2/2n}}\right)^{\xi k} \mathrm{P}[\mathbb{D} = d]$$

$$\vdots$$

$$\leqslant \sum_{k=2\xi}^{n/c} \left(\frac{k}{n}\right)^{\xi k} \frac{n^k}{k^k} \leqslant 2\left(\frac{2\xi}{n}\right)^{2\xi^2} \left(\frac{n}{2\xi}\right)^{2\xi} = 2\left(\frac{2\xi}{n}\right)^{2(\xi^2-\xi)}$$

Sum is supergeometric:
bound by twice the first term.

Supercubic regime: 2nd ingredient

$$\mathrm{P}\big[\,\exists H \subseteq G^{\mathrm{CL}}(D_n)) : |H| \leqslant n/c \ \mathit{and}\ \nabla_0(H) \geqslant \xi\,\big] \leqslant \frac{1}{n^{\xi}}$$

$$\sum_{k=2\xi}^{n/c} \binom{n}{k} \mathrm{P}\big[\,\|G^{\mathrm{CL}}(D_n)[X_k])\| \geqslant k\xi\,\big]$$

$$\leqslant \sum_{k=2\xi}^{n/c} \binom{n}{k} \sum_{d=k}^{\Delta k} \left(\frac{ecd^2}{2n\xi k e^{d^2/2n}}\right)^{\xi k} \mathrm{P}[\mathbb{D} = d]$$

$$\vdots$$

For $n \geqslant 4\xi$ and $\xi \geqslant 2.5$

$$\leqslant \sum_{k=2\xi}^{n/c} \left(\frac{k}{n}\right)^{\xi k} \frac{n^k}{k^k} \leqslant 2\left(\frac{2\xi}{n}\right)^{2\xi^2} \left(\frac{n}{2\xi}\right)^{2\xi} = 2\left(\frac{2\xi}{n}\right)^{2(\xi^2-\xi)} \leqslant \frac{1}{n^{\xi}}$$

Thm. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CL}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Cubic regime: ingredients

Lem. Let $(D_n)_{n \in \mathbb{N}}$ be a sparse DDS with a tail in $\Theta(d^{-3})$. Then the probability that $G^{\mathrm{CL}}(D_n)$ contains an s-t-path of length r is at most

$$\frac{\omega_s \omega_t}{\mu n} \cdot \Theta(\log n)$$

Lem. Let $(D_n)_{n \in \mathbb{N}}$ be a sparse DDS with upper tail-bound $O(d^{-2-\epsilon})$. Then for $\xi \geqslant \Theta(1)$ and large enough n it holds that

$$\mathrm{P}\Big[\omega(G^{\mathrm{CL}}(D_n)) \geqslant 4\sqrt{\xi(\alpha+1)/(\alpha-1)}\Big] \leqslant \frac{1}{n^{\xi}}.$$

Cubic regime: dense minor

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$

with high probability.

Cubic regime: dense minor

High-degree vertices must be important!

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$

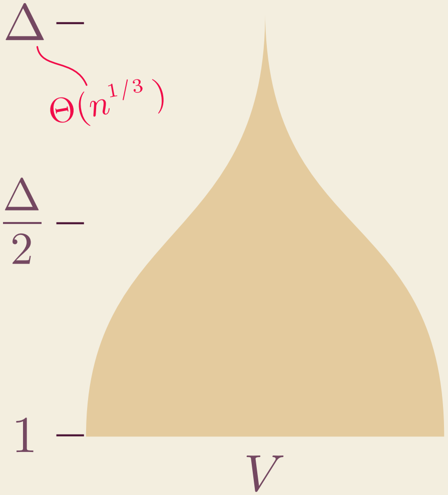
with high probability.

Goal: construct dense 1-shallow minor

Small minor suffices!

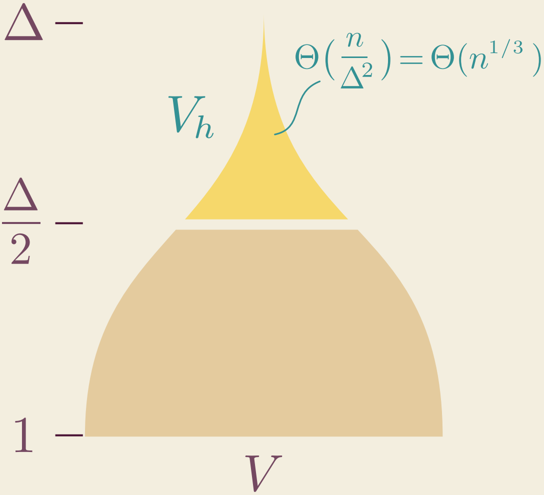
Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



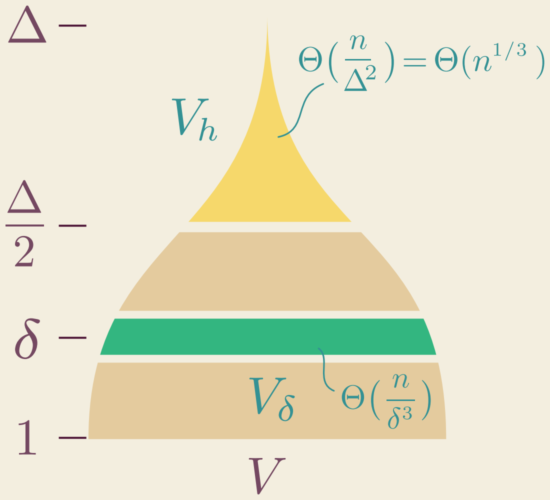
Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



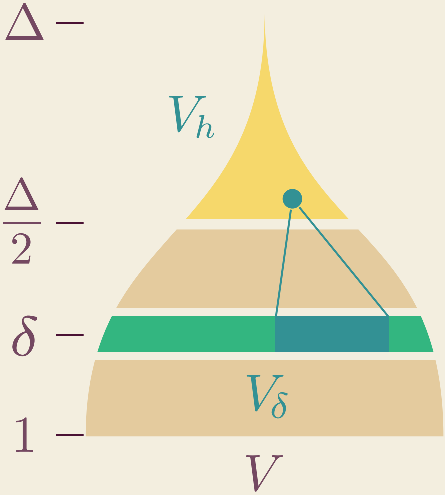
Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



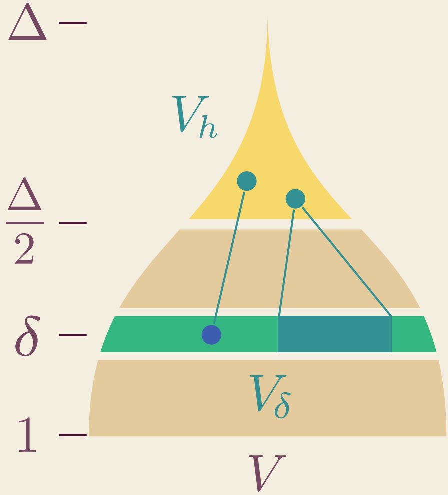
$E[|N(x) \cap V_\delta|] \geqslant \Theta(\frac{\Delta}{\delta^2})$

for $x \in V_h$

Vertices in V_h have many neighbours in V_δ.

Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



$E[|N(x) \cap V_\delta|] \geqslant \Theta(\frac{\Delta}{\delta^2})$

for $x \in V_h$

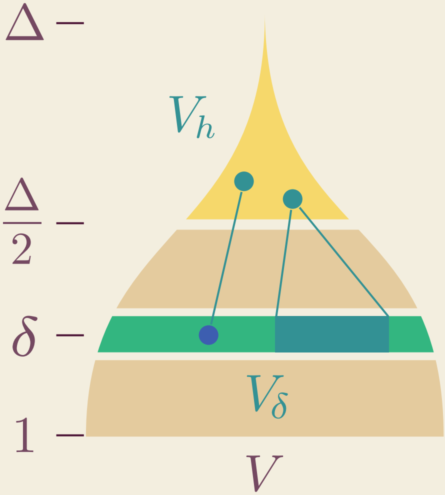
Vertices in V_h have many neighbours in V_δ.

$E[|N(y) \cap V_h|] \leqslant \Theta(\frac{\delta}{\Delta})$

for $y \in V_\delta$

Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



$E[|N(x) \cap V_\delta|] \geqslant \Theta(\frac{\Delta}{\delta^2})$

for $x \in V_h$

Vertices in V_h have many neighbours in V_δ.

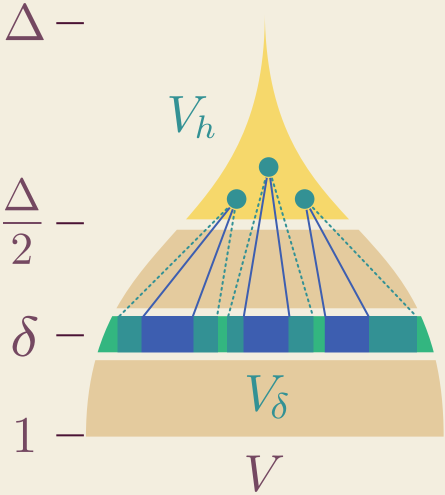
Vertices in V_δ have very few neighbours in V_h.

$E[|N(y) \cap V_h|] \leqslant \Theta(\frac{1}{\log n})$

for $y \in V_\delta$ and $\delta \leqslant \frac{\Delta}{\log n}$

Cubic regime: dense minor

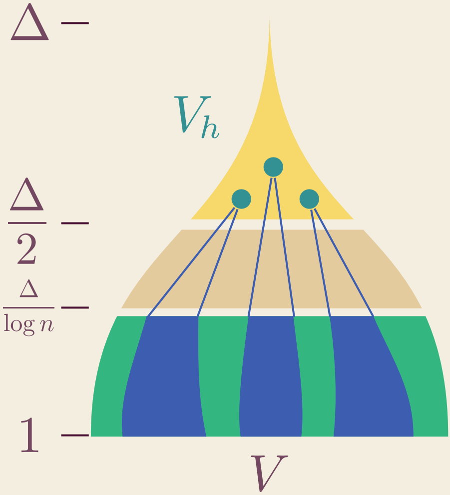
$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



With high probability, half of $N(x) \cap V_\delta$ has no other neighbours than x in V_h for *all* $x \in V_h$.

Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



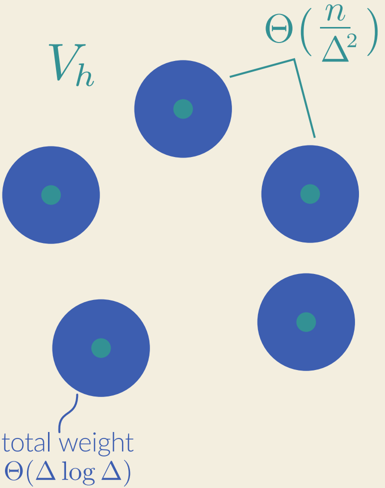
With high probability, half of $N(x) \cap V_\delta$ has no other neighbours than x in V_h for *all* $x \in V_h$.

Therefore every $x \in V_h$ has an *exclusive neighbourhood* $S_x \subseteq V_{\leqslant \frac{\Delta}{\log n}}$ whose *total weight* is at least

$$\Theta(\Delta \log \Delta)$$

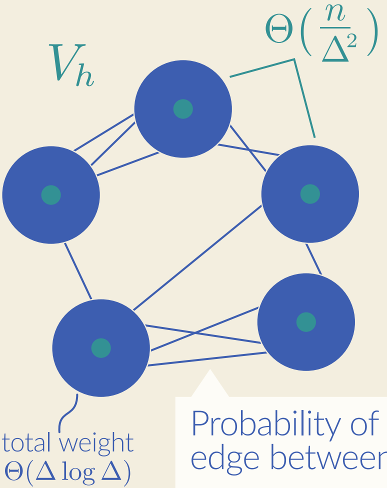
Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



Cubic regime: dense minor

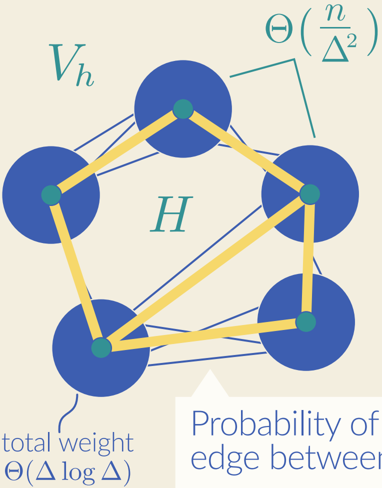
$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



Probability of at least one edge between sets: $\Theta\left(\frac{\Delta^2 \log^2 \Delta}{n}\right)$

Cubic regime: dense minor

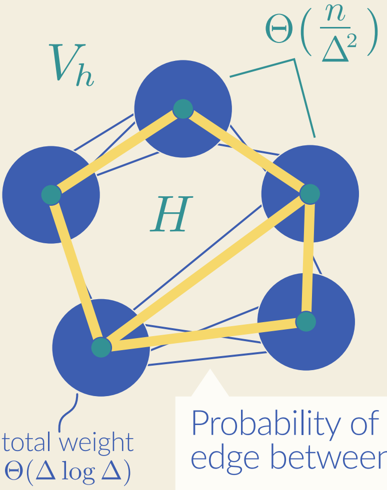
$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



Probability of at least one edge between sets: $\Theta\left(\frac{\Delta^2 \log^2 \Delta}{n}\right)$

Cubic regime: dense minor

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$



With high probability:

$$\frac{\|H\|}{|H|} = \Theta\Big(\frac{\Delta^2}{n} \cdot \frac{n^2}{\Delta^4} \cdot \frac{\Delta^2 \log^2 \Delta}{n}\Big)$$
$$= \Theta\Big(\frac{n}{\Delta^2} \cdot \frac{\Delta^2 \log^2 \Delta}{n}\Big)$$
$$= \Theta\big(\log^2 \Delta\big) = \Theta\big(\log^2 n\big)$$

Probability of at least one edge between sets: $\Theta\left(\frac{\Delta^2 \log^2 \Delta}{n}\right)$

Cubic regime: dense minor

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega(\log^2 n)$$

with high probability.

Thm. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CL}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Subcubic regime: dense minor

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3-\varepsilon})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega\left(n^{\frac{\varepsilon}{2(1-\varepsilon/3)}}\right)$$

with high probability.

Subcubic regime: dense minor

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3-\varepsilon})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega\Big(n^{\frac{\varepsilon}{2(1-\varepsilon/3)}}\Big)$$

with high probability.

Thm. (Dvořák, Jiang) Let $\ell \in \mathbb{N}$ and $\varepsilon > 0$. There exists $n_{\ell,\varepsilon}$ and c_ε such that every graph G with $n > n_{\ell,\varepsilon}$ vertices and at least $n^{1+\varepsilon}$ edges contains a c_ε-subdivision of K_ℓ.

Jiang T. **Compact topological minors in graphs.** Journal of Graph Theory. 2011 Jun;67(2):139-52.

Dvořák Z. **Asymptotical structure of combinatorial objects.** PhD thesis.

Subcubic regime: dense minor

Lem. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse DDS with lower tail-bound $\Omega(d^{-3-\varepsilon})$. Then

$$\nabla_1(G^{\mathrm{CL}}(D_n)) = \Omega\left(n^{\frac{\varepsilon}{2(1-\varepsilon/3)}}\right)$$

with high probability.

Thm. (Dvořák, Jiang) Let $\ell \in \mathbb{N}$ and $\varepsilon > 0$. There exists $n_{\ell,\varepsilon}$ and c_ε such that every graph G with $n > n_{\ell,\varepsilon}$ vertices and at least $n^{1+\varepsilon}$ edges contains a c_ε-subdivision of K_ℓ.

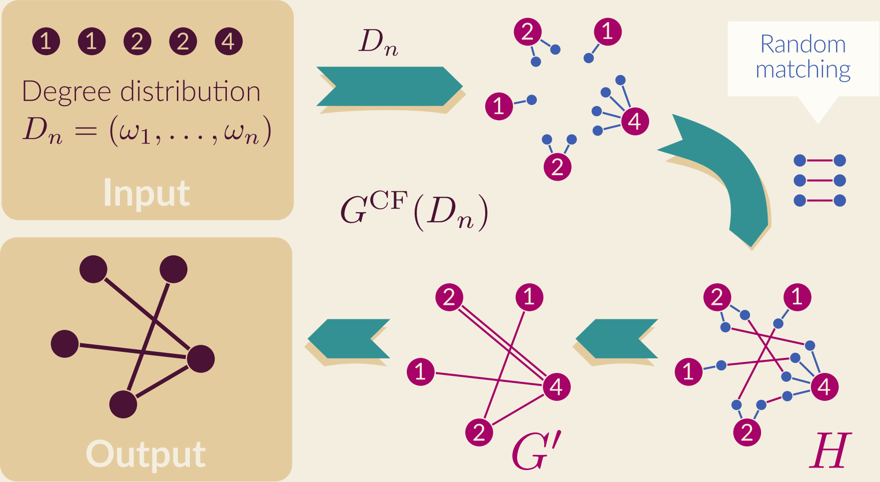
Somewhere dense!

Jiang T. **Compact topological minors in graphs.** Journal of Graph Theory. 2011 Jun;67(2):139-52.

Dvořák Z. **Asymptotical structure of combinatorial objects.** PhD thesis.

Configuration model

Generate multi-graphs with a given degree distribution.



Configuration model: sparseness

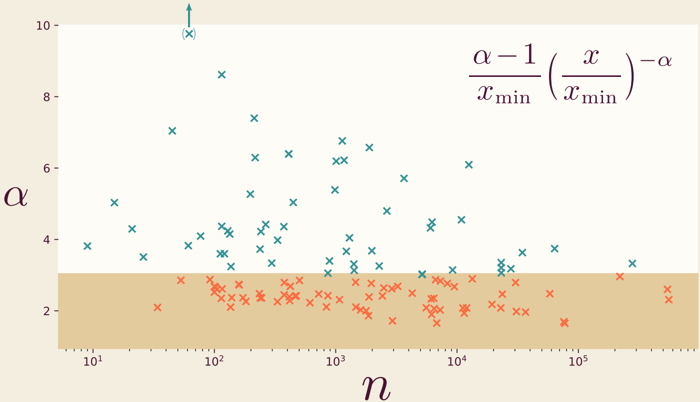
Thm. Let $(D_n)_{n\in\mathbb{N}}$ be a sparse degree distribution sequence. Then $G^{\mathrm{CF}}(D_n)$

1. has **bounded expansion** if (D_n) has an upper tail-bound of $O(\frac{1}{d^{3+\varepsilon}})$,
2. is **nowhere dense** with unbounded expansion if (D_n) has a tail in $\Theta(\frac{1}{d^3})$, and
3. is **somewhere dense** if (D_n) has a lower tail-bound of $\Omega(\frac{1}{d^{3-\varepsilon}})$.

Proofs very similar to Chung–Lu!

Real tail-bounds!

We estimated the tail of 129 networks using the `powerlaw` package.



Supercubic:
56
networks

Subcubic:
73
networks

Alstott J, Bullmore E, Plenz D. **powerlaw: a Python package for analysis of heavy-tailed distributions.** PloS one. 2014 Jan 29;9(1):e85777.

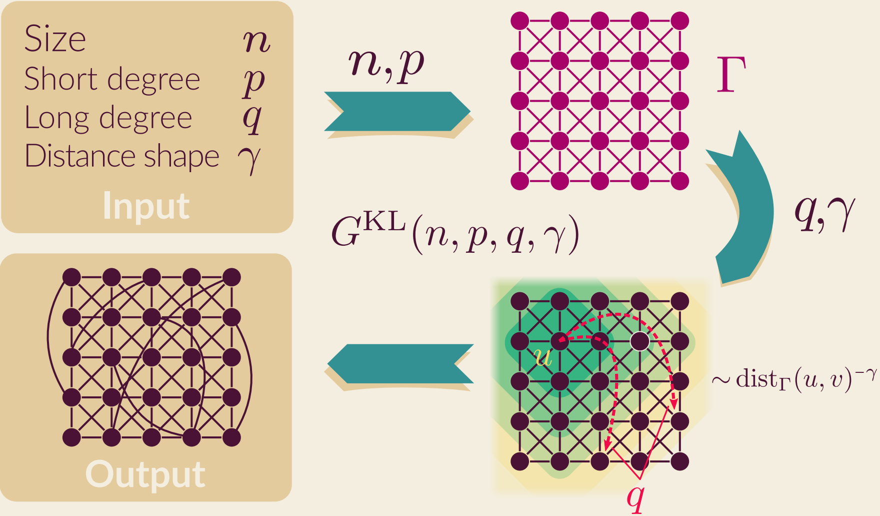
based on

Clauset A, Shalizi CR, Newman ME. **Power-law distributions in empirical data.** SIAM review. 2009 Nov 6;51(4):661-703.

…what about Kevin Bacon?

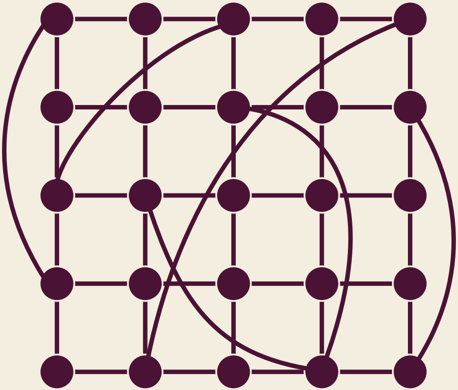
The Kleinberg model

Generate graphs that admit *greedy routing*



Greedy routing

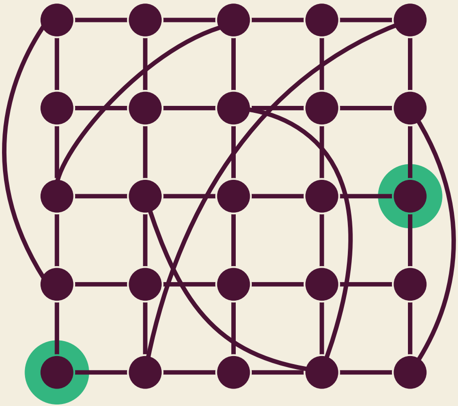
Small world: short average path length
Greedy routing: short paths can be *found* with a *local* routing protocol



Greedy routing

Small world: short average path length

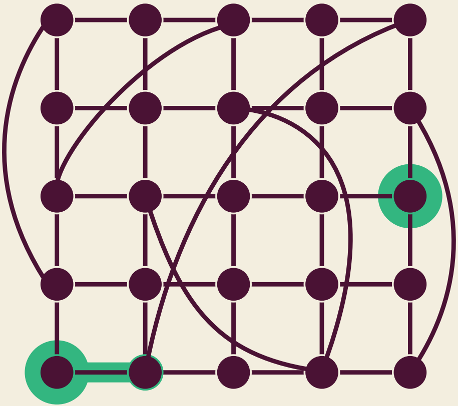
Greedy routing: short paths can be *found* with a *local* routing protocol



Greedy routing

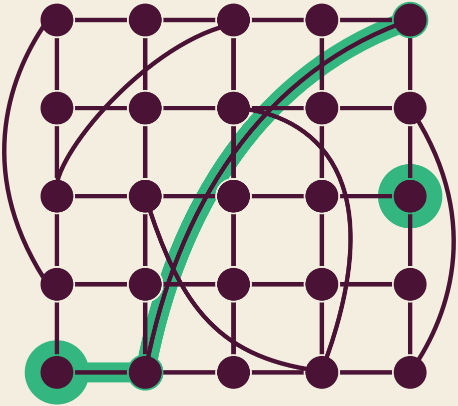
Small world: short average path length

Greedy routing: short paths can be *found* with a *local* routing protocol



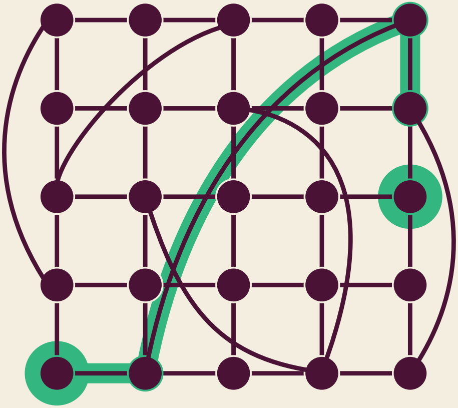
Greedy routing

Small world: short average path length
Greedy routing: short paths can be *found* with a *local* routing protocol



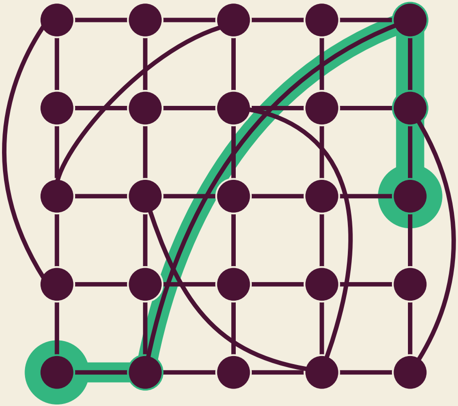
Greedy routing

Small world: short average path length
Greedy routing: short paths can be *found* with a *local* routing protocol

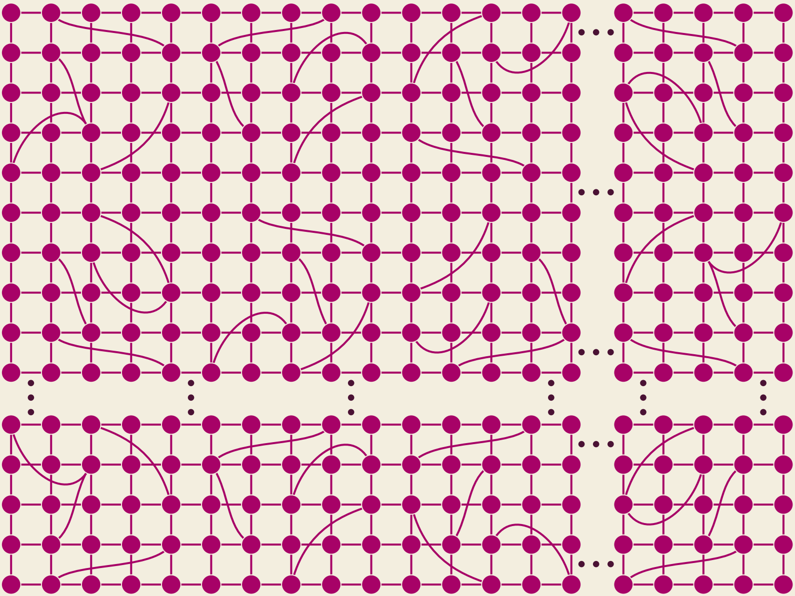


Greedy routing

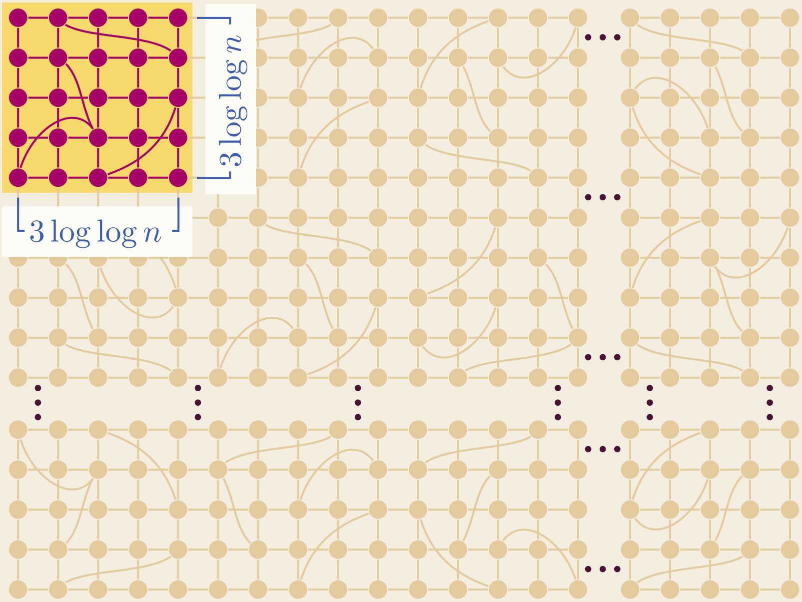
Small world: short average path length
Greedy routing: short paths can be *found* with a *local* routing protocol



The Kleinberg model is dense

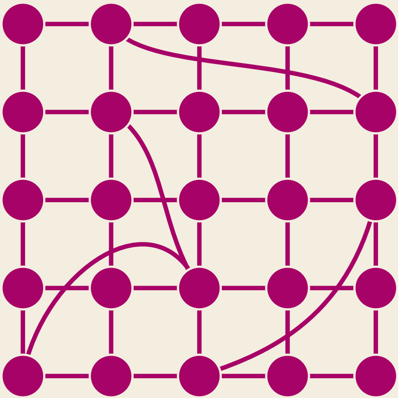


The Kleinberg model is dense



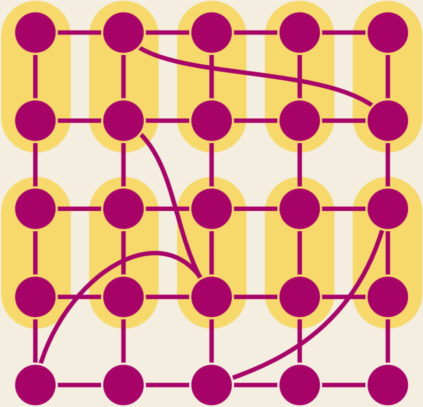
The Kleinberg model is dense

$3 \log \log n \times 3 \log \log n$



The Kleinberg model is dense

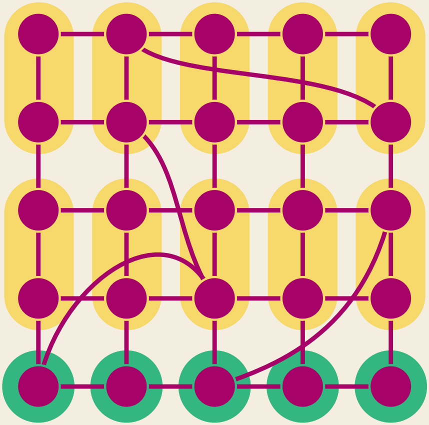
$3 \log \log n \times 3 \log \log n$



Select $(\log \log n)^2$ disjoint edges

The Kleinberg model is dense

$3 \log \log n \times 3 \log \log n$

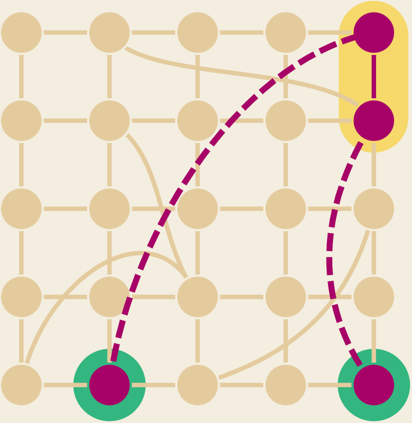


Select $(\log \log n)^2$ disjoint edges

Select $3 \log \log n$ vertices

The Kleinberg model is dense

$3 \log \log n \times 3 \log \log n$

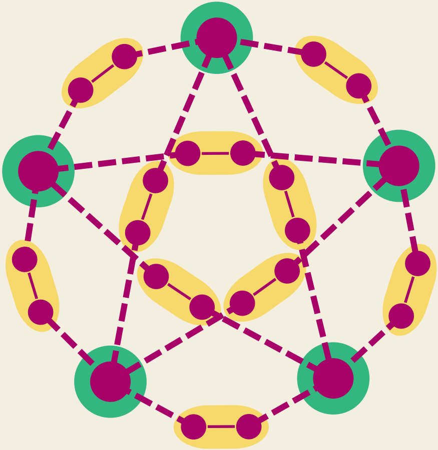


Probability that a *specific* pair of vertices are both connected to a *specific* edge:

$$\frac{1}{d_1^2} \cdot \frac{1}{d_2^2} \cdot \frac{1}{\log^2 n}$$
$$\geqslant \frac{1}{(3 \log \log n)^4 \log^2 n}$$

The Kleinberg model is dense

$3 \log \log n \times 3 \log \log n$

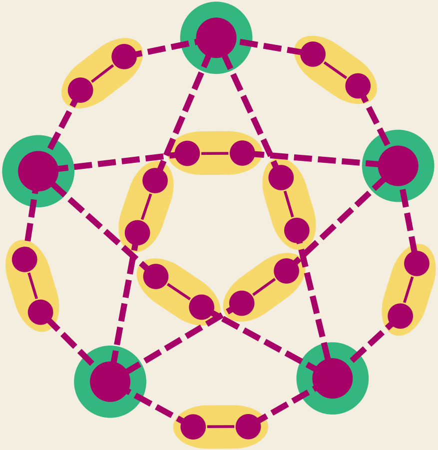


Probability that *every* pair of chosen vertices is connected by a specific edge each:

$$\left(\frac{1}{(3 \log \log n)^4 \log^2 n}\right)^{(\log \log n)^2}$$

The Kleinberg model is dense

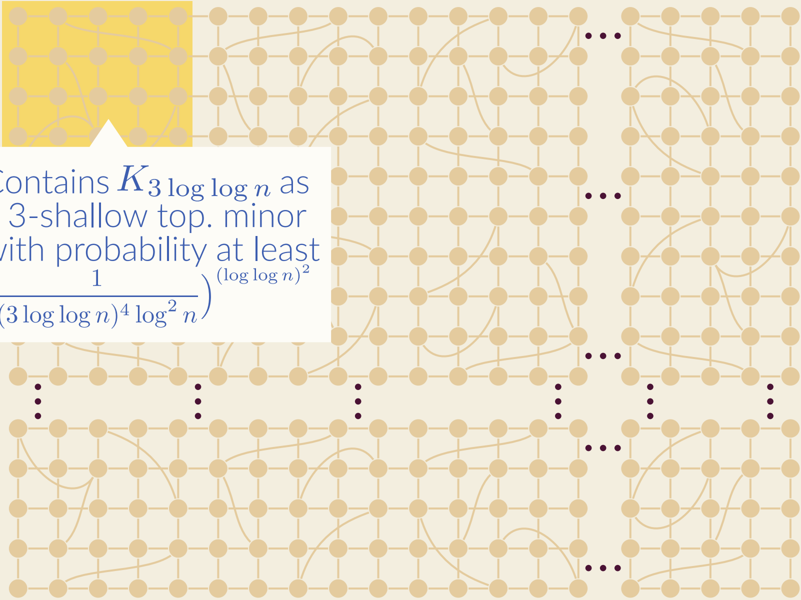
$3 \log \log n \times 3 \log \log n$



Probability that *every* pair of chosen vertices is connected by a specific edge each:

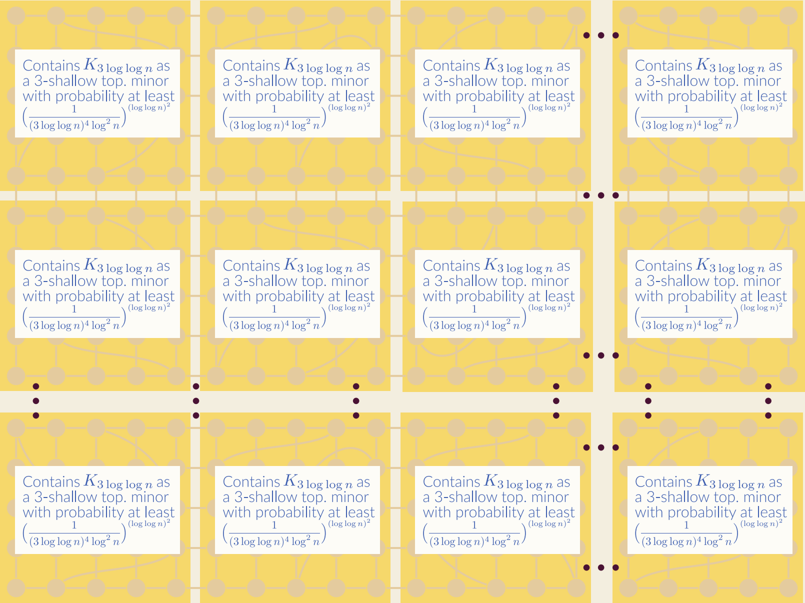
$$\left(\frac{1}{(3 \log \log n)^4 \log^2 n}\right)^{(\log \log n)^2}$$

The Kleinberg model is dense



Contains $K_{3 \log \log n}$ as a 3-shallow top. minor with probability at least $\left(\frac{1}{(3 \log \log n)^4 \log^2 n}\right)^{(\log \log n)^2}$

The Kleinberg model is dense



The Kleinberg model is dense

Flip a coin with success probability

$$\left(\frac{1}{(3\log\log n)^4 \log^2 n}\right)^{(\log\log n)^2}$$

a total of $\frac{n^2}{(3\log\log n)^2}$ times.

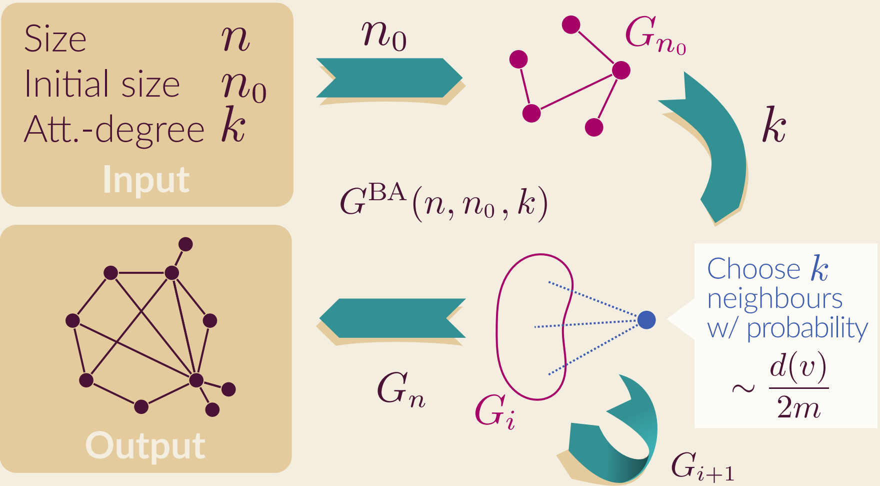
Asymptotically, the probability that none of flips wins is at most e^{-n}.

The Kleinberg model is somewhere dense w.h.p.

and this… but that… or this…

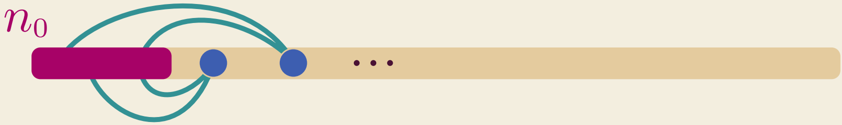
Attachment models

Plausible network formation that generates powerlaw degree distributions



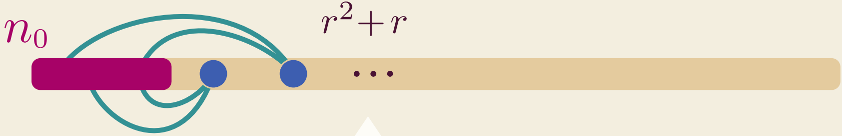
Clique minors in G^{BA}

Lem. $G^{\mathrm{BA}}(n, n_0, k)$ with $k \geqslant 2$ contains, for any $r \leqslant \sqrt{n/2}$, the complete graph K_r with some probability $p(r) > 0$.



Clique minors in G^{BA}

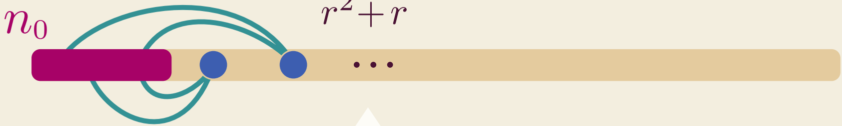
Lem. $G^{\mathrm{BA}}(n, n_0, k)$ with $k \geqslant 2$ contains, for any $r \leqslant \sqrt{n/2}$, the complete graph K_r with some probability $p(r) > 0$.



Probability of forming K_r is some function of r.

Clique minors in G^{BA}

Lem. $G^{\text{BA}}(n, n_0, k)$ with $k \geqslant 2$ contains, for any $r \leqslant \sqrt{n/2}$, the complete graph K_r with some probability $p(r) > 0$.

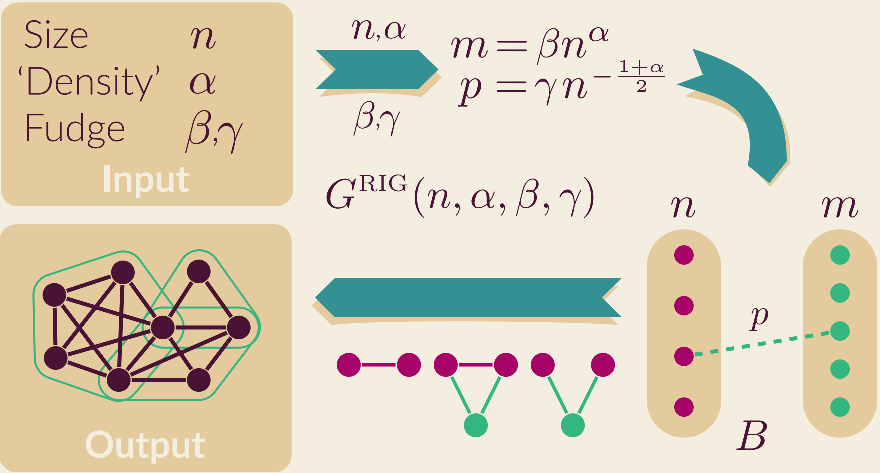


Probability of forming K_r is some function of r.

Attachment models are *not a.a.s nowhere dense!*

Random intersection model

Projection of shared attributes onto binary relationship.



Random intersection model

$$G^{\mathrm{RIG}}(n, \alpha, \beta, \gamma) \qquad \begin{aligned} m &= \beta n^{\alpha} \\ p &= \gamma n^{-\frac{1+\alpha}{2}} \end{aligned}$$

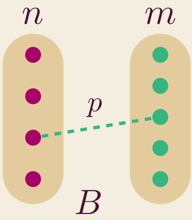
regime	$\alpha < 1$	$\alpha = 1$	$\alpha > 1$
degeneracy	$\Omega\left(\gamma n^{(1-\alpha)/2}\right)$	$\Omega\left(\frac{\log n}{\log \log n}\right)$	$O(1)$
sparsity	Somewhere dense		B.E.

Random intersection model

$G^{\mathrm{RIG}}(n, \alpha, \beta, \gamma)$

$$m = \beta n^{\alpha}$$
$$p = \gamma n^{-\frac{1+\alpha}{2}}$$

regime	$\alpha < 1$	$\alpha = 1$	$\alpha > 1$
degeneracy	$\Omega(\gamma n^{(1-\alpha)/2})$	$\Omega(\frac{\log n}{\log \log n})$	$O(1)$
sparsity	Somewhere dense		B.E.



Contains a high-degree vertex

Random intersection model

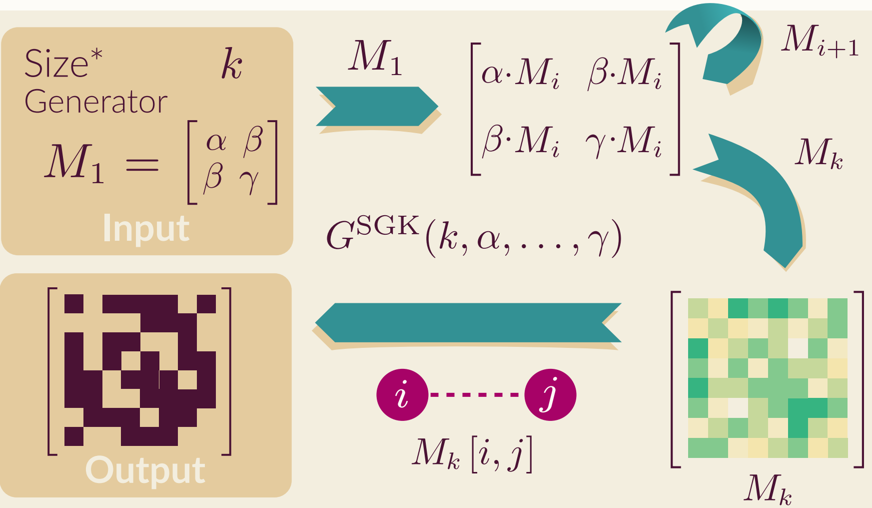
$G^{\mathrm{RIG}}(n, \alpha, \beta, \gamma)$ $m = \beta n^{\alpha}$, $p = \gamma n^{-\frac{1+\alpha}{2}}$

regime	$\alpha < 1$	$\alpha = 1$	$\alpha > 1$
degeneracy	$\Omega(\gamma n^{(1-\alpha)/2})$	$\Omega(\frac{\log n}{\log \log n})$	$O(1)$
sparsity	Somewhere dense	Somewhere dense	B.E.

Regime with ‘tunable’ clustering

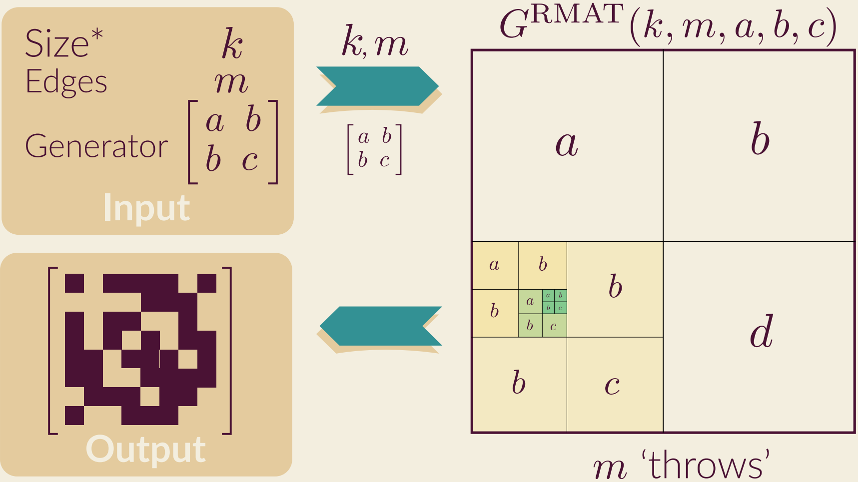
Kronecker models

Hierarchical generation of community-like structures.



RMAT models

Hierarchical generation of community-like structures.

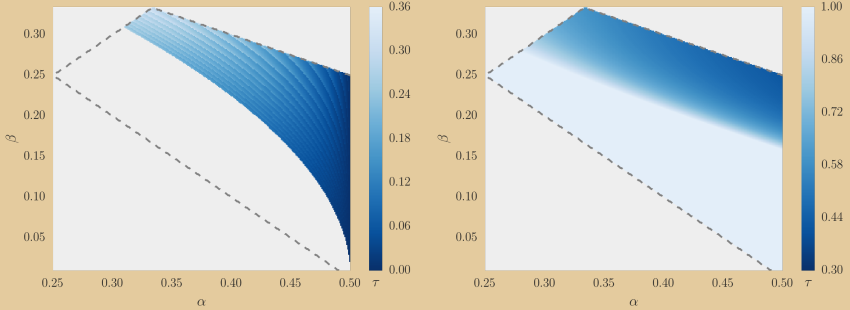


Kronecker models

For $\alpha \geqslant \frac{1}{2}$ it is easy to show that the models are somewhere dense.

Kronecker models

For $\alpha \geqslant \frac{1}{2}$ it is easy to show that the models are somewhere dense.

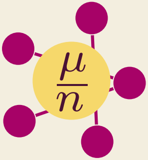


Dense

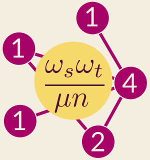
Degenerate ‘slices’

Random model sparsity

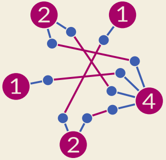
$G(n, \frac{\mu}{n})$



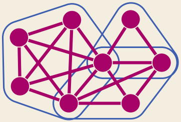
$G^{\mathrm{CL}}(D_n)$



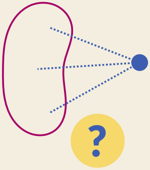
$G^{\mathrm{CF}}(D_n)$



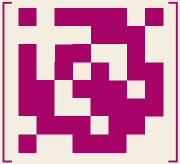
$G^{\mathrm{RIG}}(n, \alpha, \beta, \gamma)$



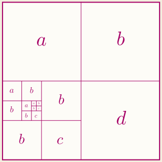
$G^{\mathrm{BA}}(n, n_0, k)$



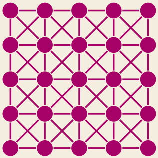
$G^{\mathrm{SGK}}(k, \alpha, \ldots, \gamma)$



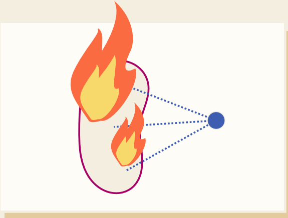
$G^{\mathrm{RMAT}}(k, m, a, b, c)$



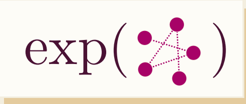
$G^{\mathrm{KL}}(n, p, q, \gamma)$



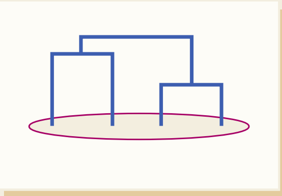
So many models!



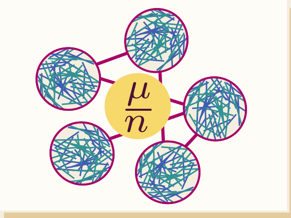
Forest fire



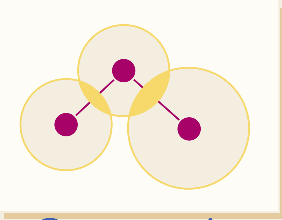
Exponential random graphs



Hierarchical tree model

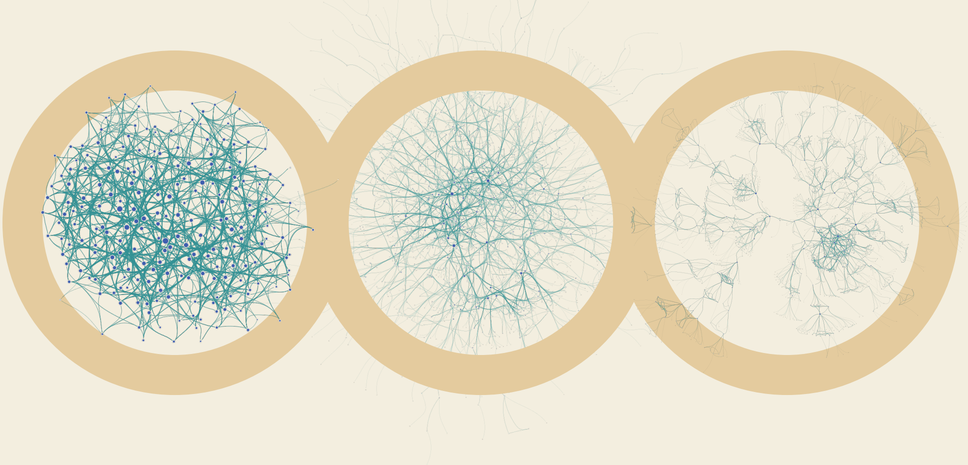


BTER



Geometric random graphs

The big question



Which ones have
bounded expansion?