Arboricity and spanning-tree packing in random graphs with an application to load balancing

Xavier Pérez-Giménez[†]

joint work with

Jane (Pu) Gao[*] and Cristiane M. Sato[†]

[†]University of Waterloo

[*]University of Toronto

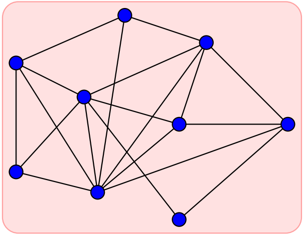
Laboratoire d'Informatique de l'École Polytechnique, Paris, December 2013

Motivation

Spanning-tree packing (STP) number

Definition

$T(G)$ = maximum number of edge-disjoint spanning trees in G.



Example

$T(G) = 2.$

Trivial upper bound

$T(G) \leq \min\left\{\delta, \left\lfloor \frac{\bar{d}}{2} \right\rfloor\right\}$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

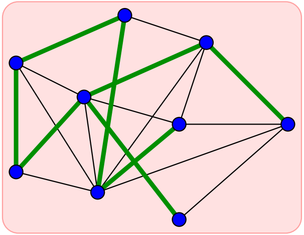
An application

Measure of network strength/vulnerability.

Spanning-tree packing (STP) number

Definition

$T(G)$ = maximum number of edge-disjoint spanning trees in G.



Example

$T(G) = 2$.

Trivial upper bound

$T(G) \leq \min\left\{\delta, \left\lfloor \frac{\bar{d}}{2} \right\rfloor\right\}$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

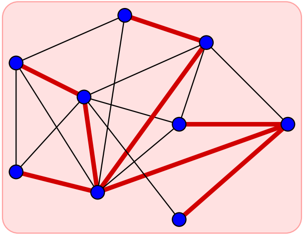
An application

Measure of network strength/vulnerability.

Spanning-tree packing (STP) number

Definition

$T(G)$ = maximum number of edge-disjoint spanning trees in G.



Example

$T(G) = 2$.

Trivial upper bound

$$T(G) \le \min\left\{\delta, \left\lfloor \frac{\bar{d}}{2} \right\rfloor\right\}, \qquad \text{where} \qquad \frac{\bar{d}}{2} = \frac{m}{n-1}.$$

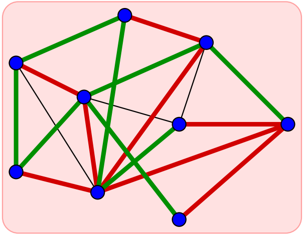
An application

Measure of network strength/vulnerability.

Spanning-tree packing (STP) number

Definition

$T(G)$ = maximum number of edge-disjoint spanning trees in G.



Example

$T(G) = 2.$

Trivial upper bound

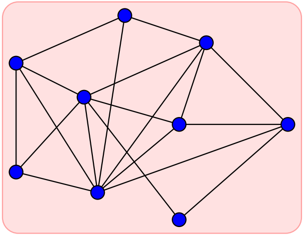
$T(G) \leq \min\left\{\delta, \left\lfloor \frac{\bar{d}}{2} \right\rfloor\right\}$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

An application

Measure of network strength/vulnerability.

Definition

$$A(G) = \text{minimum number of spanning trees covering all edges of } G;$$
$$= \text{minimum number of forests decomposing } E(G).$$



Example

$A(G) = 3.$

Trivial lower bound

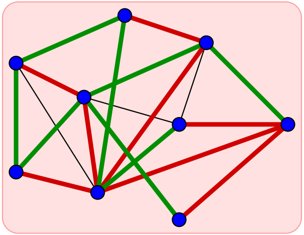
$A(G) \geq \left\lceil \frac{\bar{d}}{2} \right\rceil$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

An application

Measure of density of subgraphs.

Definition

$A(G)$ = minimum number of spanning trees covering all edges of G;
= minimum number of forests decomposing $E(G)$.



Example

$A(G) = 3$.

Trivial lower bound

$A(G) \geq \left\lceil \frac{\bar{d}}{2} \right\rceil$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

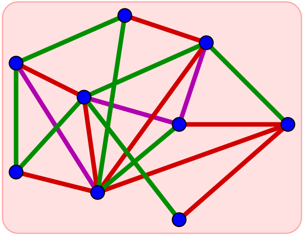
An application

Measure of density of subgraphs.

Arboricity

Definition

$$A(G) = \text{minimum number of spanning trees covering all edges of } G;$$
$$= \text{minimum number of forests decomposing } E(G).$$



Example

$A(G) = 3.$

Trivial lower bound

$A(G) \geq \lceil \frac{\bar{d}}{2} \rceil$, where $\frac{\bar{d}}{2} = \frac{m}{n-1}$.

An application

Measure of density of subgraphs.

k-orientability and load balancing

k-orientability

A graph is k-orientable if it admits an orientation of the edges s.t. the maximum indegree is at most k.

Equivalent formulation:

Load balancing scenario

m balls (jobs) are assigned to n bins (machines) in a way that each ball must pick between two randomly chosen bins. We wish to minimise the load of the bins by allowing at most k balls in each bin.

Connection to arboricity:

k-orientability determined by density of densest subgraph (Hakimi '65).

Problem and previous results

Theorem (Palmer, Spencer '95)

$T\big(\mathscr{G}(n,p)\big) = \delta, \quad$ for δ constant $(p \sim \log n/n)$.

Theorem (Catlin, Chen, Palmer '93)

$$\begin{cases} T\big(\mathscr{G}(n,p)\big) = \left\lfloor \frac{\bar{d}}{2} \right\rfloor \\ A\big(\mathscr{G}(n,p)\big) = \left\lceil \frac{\bar{d}}{2} \right\rceil \end{cases}, \quad \text{for } p = C(\log n/n)^{1/3}.$$

Theorem (Chen, Li, Lian '13+)

$T\big(\mathscr{G}(n,p)\big) = \delta, \quad$ for $p \le 1.1 \log n/n$;
$T\big(\mathscr{G}(n,p)\big) > \delta, \quad$ for $p \ge 51 \log n/n$.

Question (Chen, Li, Lian)

What's the smallest p such that $T\big(\mathscr{G}(n,p)\big) > \delta$?

Our results (i)

Theorem

For every $p = p(n) \in [0,1]$, a.a.s. $T(\mathscr{G}(n,p)) = \min\left\{\delta, \left\lfloor \frac{\bar{d}}{2} \right\rfloor\right\}$.
(Same holds throughout the random graph process.)

Theorem

Let $\beta = 2/\log(e/2) \approx 6.51778$.

- If $p = \frac{\beta\left(\log n - \frac{\log\log n}{2}\right) - \omega(1)}{n-1}$, then a.a.s. $\delta \le \left\lfloor \frac{\bar{d}}{2} \right\rfloor$ in $\mathscr{G}(n,p)$.
- If $p = \frac{\beta\left(\log n - \frac{\log\log n}{2}\right) + \omega(1)}{n-1}$, then a.a.s. $\delta > \left\lfloor \frac{\bar{d}}{2} \right\rfloor$ in $\mathscr{G}(n,p)$.

(Same holds throughout the random graph process.)

Threshold for $T(\mathscr{G}(n,p)) = \begin{cases} \delta \\ \left\lfloor \frac{\bar{d}}{2} \right\rfloor \end{cases}$ at $p \sim \beta \frac{\log n}{n}$.

Our results (ii)

Theorem

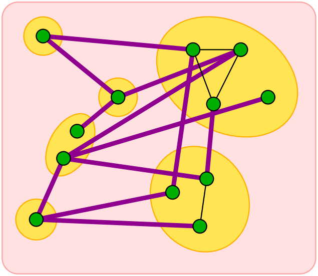
- If $p = o(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \leq 1$.
- If $p = \Theta(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \in \{k(p), k(p)+1\}$.
- If $p = \omega(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \in \{\lceil \frac{\bar{d}}{2} \rceil, \lceil \frac{\bar{d}}{2} \rceil + 1\}$.
  For most values of $p = \omega(1/n)$, a.a.s. $A\big(\mathscr{G}(n,p)\big) = \lceil \frac{\bar{d}}{2} \rceil$.

(Same holds throughout the random graph process.)

Corollary

Threshold for k-orientability of $\mathscr{G}(n,m)$ for $k \to \infty$ at $m \sim kn$.

Tutte and Nash-Williams



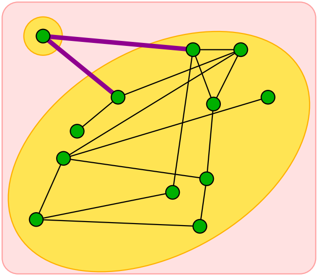
$\mathcal{P}$ partition of $V(G)$;
$m(\mathcal{P})$ = number of edges in $E(G)$ with ends in distinct parts of $\mathcal{P}$.

Theorem (Tutte '61; Nash-Williams '61)

A graph G has t edge-disjoint spanning trees iff every partition $\mathcal{P}$ of $V(G)$ satisfies $m(\mathcal{P}) \geq t(|\mathcal{P}| - 1)$.

$T(G)$ is given by the smallest ratio $\left\lfloor \frac{m(\mathcal{P})}{|\mathcal{P}|-1} \right\rfloor$.

Tutte and Nash-Williams

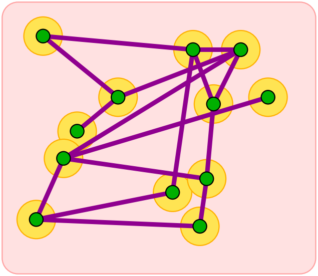


$\mathcal{P}$ partition of $V(G)$;
$m(\mathcal{P})$ = number of edges in $E(G)$ with ends in distinct parts of $\mathcal{P}$.

Theorem (Tutte '61; Nash-Williams '61)

A graph G has t edge-disjoint spanning trees iff every partition $\mathcal{P}$ of $V(G)$ satisfies $m(\mathcal{P}) \geq t(|\mathcal{P}| - 1)$.

$T(G)$ is given by the smallest ratio $\left\lfloor \frac{m(\mathcal{P})}{|\mathcal{P}|-1} \right\rfloor$.



$\mathcal{P}$ partition of $V(G)$;
$m(\mathcal{P})$ = number of edges in $E(G)$ with ends in distinct parts of $\mathcal{P}$.

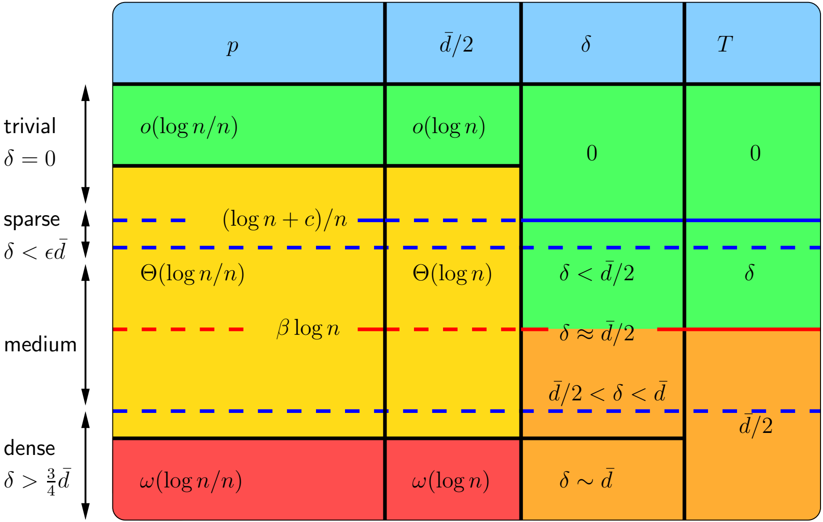
Theorem (Tutte '61; Nash-Williams '61)

A graph G has t edge-disjoint spanning trees iff every partition $\mathcal{P}$ of $V(G)$ satisfies $m(\mathcal{P}) \geq t(|\mathcal{P}| - 1)$.

$T(G)$ is given by the smallest ratio $\left\lfloor \frac{m(\mathcal{P})}{|\mathcal{P}|-1} \right\rfloor$.

p	$\bar{d}/2$	δ	T
$o(\log n/n)$	$o(\log n)$	0	0
$(\log n + c)/n$			
$\Theta(\log n/n)$	$\Theta(\log n)$	$\delta < \bar{d}/2$	δ
$\beta \log n$		$\delta \approx \bar{d}/2$	
		$\bar{d}/2 < \delta < \bar{d}$	$\bar{d}/2$
$\omega(\log n/n)$	$\omega(\log n)$	$\delta \sim \bar{d}$	

Proof STP number $\mathcal{G}(n, p)$



Proof STP number $\mathscr{G}(n, p)$

Different ranges

- Trivial $\quad p \leq 0.9\frac{\log n}{n}$: $\quad \delta = 0.$
- Sparse $\quad 0.9\frac{\log n}{n} \leq p \leq \gamma_1\frac{\log n}{n}$: $\quad \delta < \epsilon \bar{d}.$
- Medium $\quad \gamma_1\frac{\log n}{n} \leq p \leq \gamma_2\frac{\log n}{n}$: $\quad \delta \approx \bar{d}/2$
- Dense $\quad p \geq \gamma_2\frac{\log n}{n}$: $\quad \delta \geq \frac{3}{4}\bar{d}.$

($\gamma_1 > 1$ small, γ_2 large)

Proof STP number $\mathscr{G}(n, p)$ — Sparse range

Recall: $$0.9\frac{\log n}{n} \le p \le \gamma_1 \frac{\log n}{n}, \qquad \gamma_1 > 1 \text{ small}$$

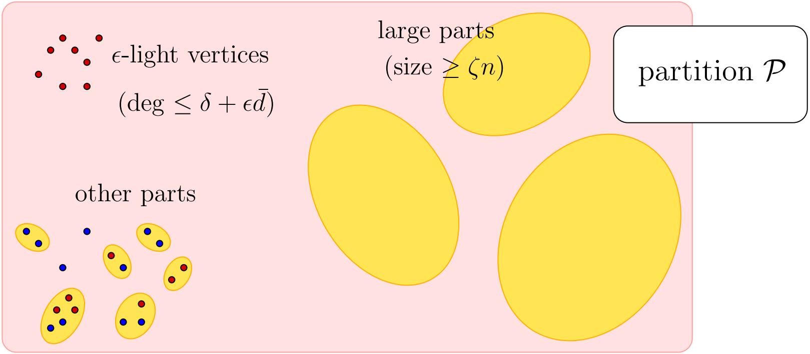
A.a.s. properties (A)

- $\delta < \epsilon \bar{d}$.
- No pair of ϵ-light vertices (degree $\le \delta + \epsilon \bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\ge \zeta n$) has many edges across.

Deterministic argument: $$\forall \mathcal{P}, \quad \frac{m(\mathcal{P})}{|\mathcal{P}| - 1} \ge \delta.$$

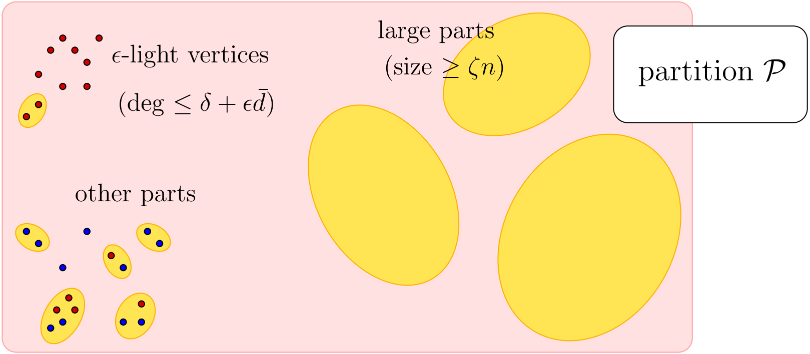
A.a.s. properties (A)

- $\delta < \epsilon \bar{d}$.
- No pair of ϵ-light vertices (degree $\leq \delta + \epsilon \bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.



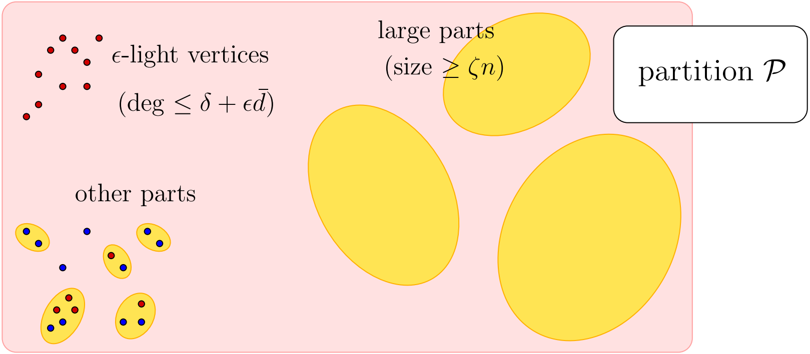
A.a.s. properties (A)

- $\delta < \epsilon \bar{d}$.
- No pair of ϵ-light vertices (degree $\leq \delta + \epsilon \bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.



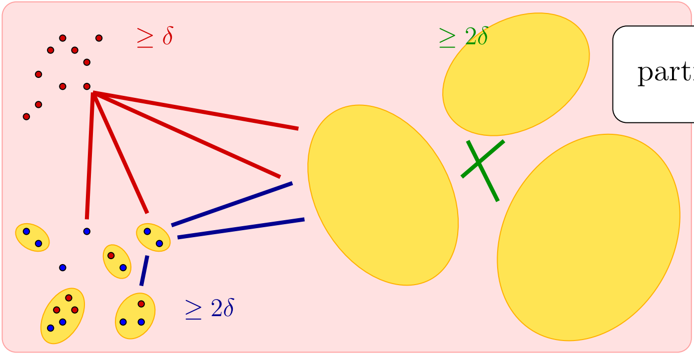
A.a.s. properties (A)

- $\delta < \epsilon \bar{d}$.
- No pair of ϵ-light vertices (degree $\leq \delta + \epsilon \bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.



A.a.s. properties (A)

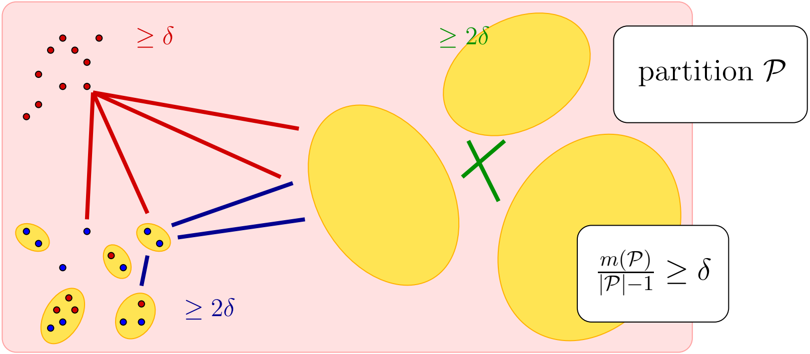
- $\delta < \epsilon\bar{d}$.
- No pair of ϵ-light vertices (degree $\leq \delta + \epsilon\bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.



partition $\mathcal{P}$

A.a.s. properties (A)

- $\delta < \epsilon \bar{d}$.
- No pair of ϵ-light vertices (degree $\leq \delta + \epsilon \bar{d}$) are adjacent or have a common neighbour.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.



Proof STP number $\mathscr{G}(n, p)$ — Dense range

Recall: $p \geq \gamma_2 \dfrac{\log n}{n}$, γ_2 large

A.a.s. properties (B)

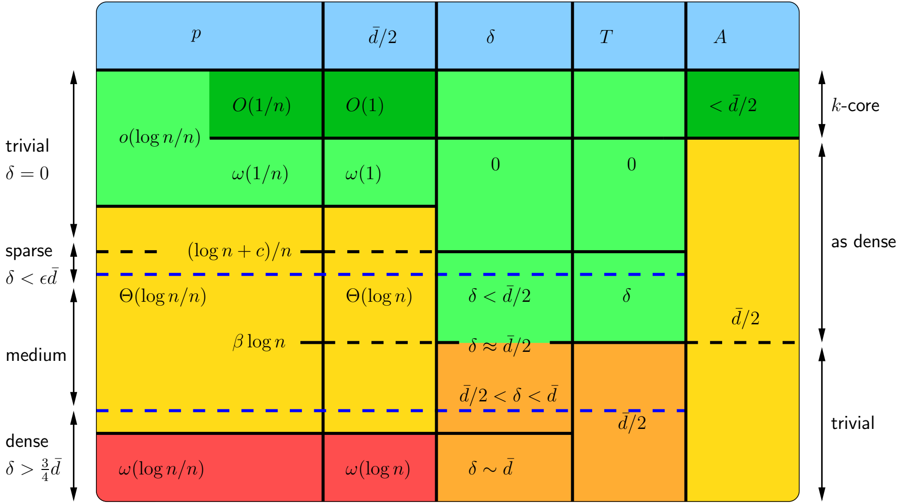
- $\delta \geq \frac{3}{4}\bar{d}$.
- All small sets of vertices (size $< \zeta n$) induce very few edges.
- Every pair of disjoint large sets of vertices (size $\geq \zeta n$) has many edges across.
- Every large set of vertices (size $\geq \zeta n$) has average degree close to $\bar{d}$.
- Every non-trivial set of vertices has at least $\bar{d}/2$ edges in its edge boundary.

Deterministic argument: $\forall \mathcal{P}$, $\dfrac{m(\mathcal{P})}{|\mathcal{P}| - 1} \geq \dfrac{m-1}{n} = \dfrac{\bar{d}}{2}$.

Theorem

- If $p = o(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \le 1$.
- If $p = \Theta(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \in \{k(p), k(p)+1\}$.
- If $p = \omega(1/n)$, then a.a.s. $A\big(\mathscr{G}(n,p)\big) \in \{\lceil \frac{\bar{d}}{2} \rceil, \lceil \frac{\bar{d}}{2} \rceil + 1\}$.
  For most values of $p = \omega(1/n)$, a.a.s. $A\big(\mathscr{G}(n,p)\big) = \lceil \frac{\bar{d}}{2} \rceil$.

(Same holds throughout the random graph process.)



p
$\bar{d}/2$
δ
T
A
trivial
$\delta = 0$
sparse
$\delta < \epsilon \bar{d}$
medium
dense
$\delta > \frac{3}{4}\bar{d}$
$O(1/n)$
$O(1)$
$< \bar{d}/2$
$o(\log n/n)$
$\omega(1/n)$
$\omega(1)$
0
0
$(\log n + c)/n$
$\Theta(\log n/n)$
$\Theta(\log n)$
$\delta < \bar{d}/2$
δ
$\bar{d}/2$
$\beta \log n$
$\delta \approx \bar{d}/2$
$\bar{d}/2 < \delta < \bar{d}$
$\bar{d}/2$
$\omega(\log n/n)$
$\omega(\log n)$
$\delta \sim \bar{d}$
k-core
as dense
trivial

Recall

$T\left(\mathcal{G}(n,p)\right) = \min\left\{\delta, \left\lfloor \frac{m}{n-1} \right\rfloor\right\}$. Threshold at $p \sim \beta \log n/n$.

Supercritical range ($\delta > \left\lfloor \frac{m}{n-1} \right\rfloor$)

- if $n-1 \mid m$, then $T\left(\mathcal{G}(n,p)\right) = \dfrac{m}{n-1} = A(\mathcal{G}(n,p))$.
- if $n-1 \nmid m$, then use coupling

$$A(\mathcal{G}(n,p)) = T\left(\mathcal{G}(n,p)\right) + 1 = \left\lceil \frac{m}{n-1} \right\rceil.$$

Proof Arboricity $\mathscr{G}(n,p)$ — Range $p = \omega(1/n)$

Recall

$T(\mathscr{G}(n,p)) = \min\left\{\delta, \left\lfloor \frac{m}{n-1} \right\rfloor\right\}$. Threshold at $p \sim \beta \log n/n$.

Subcritical range ($\delta \leq \left\lfloor \frac{m}{n-1} \right\rfloor$)

Create G' from $G = \mathscr{G}(n,p)$:

Add $o(n)$ new edges to $\mathscr{G}(n,p)$ so that

- $\delta' > \dfrac{m'}{n-1}$;
- properties (B) are satisfied;
- $n-1 \mid m'$.

Then $T(G') = \dfrac{m'}{n-1} = A(G')$ and

$$A(\mathscr{G}(n,p)) \in \left\{\left\lceil \frac{m}{n-1} \right\rceil, \left\lceil \frac{m}{n-1} \right\rceil + 1\right\}.$$

Proof Arboricity $\mathscr{G}(n,p)$ — Range $p = \Theta(1/n)$

Theorem (Nash-Williams '64)

Edges of G can be covered by t forests iff for every non-empty $S \subseteq V(G)$ we have $|E[S]| \leq t(|S|-1)$.

$$A(G) = \max_{\emptyset \neq S \subseteq V(G)} \left\lceil \frac{|E[S]|}{|S|-1} \right\rceil, \qquad \text{(where } 0/0 := 0\text{)}.$$

Proof Arboricity $\mathscr{G}(n,p)$ — Range $p = \Theta(1/n)$

k-core

Largest subgraph with minimum degree k.

Theorem (Cain, Sanders, Wormald '07)

For $k \geq 2$, if the average degree of the $(k+1)$-core of $\mathscr{G}(n,p)$ is at most $2k - \epsilon$, then a.a.s. $\mathscr{G}(n,p)$ is k-orientable.

Theorem (Hakimi '65)

A graph is k-orientable iff it has no subgraph with average degree $> 2k$.

Corollary

For $k \geq 2$, if the average degree of the $(k+1)$-core of $\mathscr{G}(n,p)$ is at most $2k + o(1)$, then a.a.s. $\mathscr{G}(n,p)$ has no subgraph with average degree more than $2k + o(1)$.

Summary of the argument

If $p \leq 1/n$, then $A \in \{1, 2\}$ (easy).

If $p > 1/n$:

- Find k such that the $(k+1)$-core of $\mathscr{G}(n, p)$ has average degree at most $2k + o(1)$, and the k-core of $\mathscr{G}(n, p)$ has average degree greater than $2(k-1)$.
- Then the densest subgraph of $\mathscr{G}(n, p)$ has average degree $> 2(k-1)$ and $\leq 2k + o(1)$.
- So $A \in \{k, k+1\}$.

Further work

- Extend results to random hypergraphs (easy).
- Extend results to other families of random graphs
  (work in progress for random geometric graphs, sparse graphs with a fix degree sequence).
- Study other graph parameters with similar characterisations following from matroid union
  (work in progress for random directed graphs).

Thank you