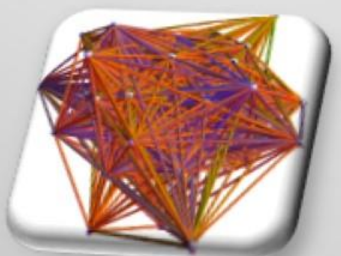
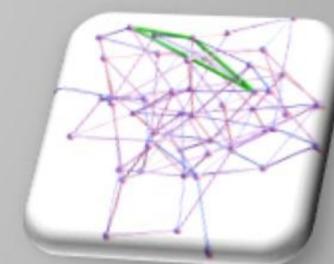
Stanford University, Jan 2014

Random Triangle Removal



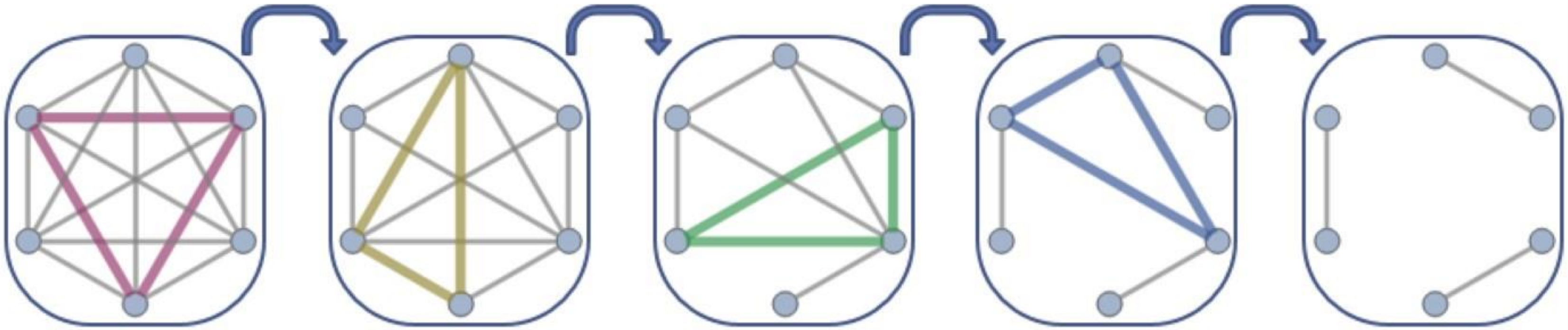
Eyal Lubetzky
Microsoft Research



Joint work with
Tom Bohman and Alan Frieze

Problem definition

- Stochastic process (*random greedy* triangle packing):
  - G_0 = complete graph on n vertices.
  - $G_i \mapsto G_{i+1}$: select a **uniform triangle** in G_i (while $\exists$) and remove its 3 edges.

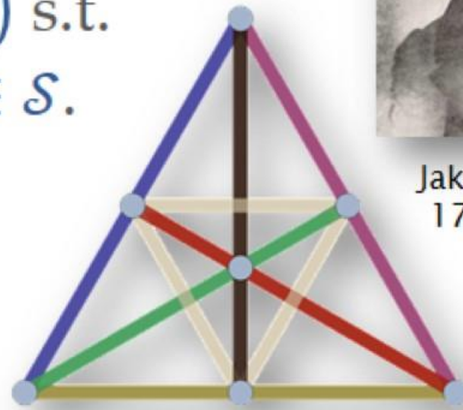


- QUESTION [Bollobás, Erdős (1990)]:

What is the expected number of edges in the final graph?

Motivation I. packing designs

- <u>DEFINITION</u> [Steiner 1853]: a *triple system* of order n is a set of triples $\mathcal{S} \subset \binom{[n]}{3}$ s.t. $\forall$ pair (x, y) belongs to a *unique* $A \in \mathcal{S}$.



Jakob Steiner
1796–1863

- *e.g.* $n = 7$: $\{1,2,4\}, \{2,3,5\}, \{1,3,6\}, \{1,5,7\}, \{3,4,7\}, \{2,6,7\}, \{4,5,6\}$

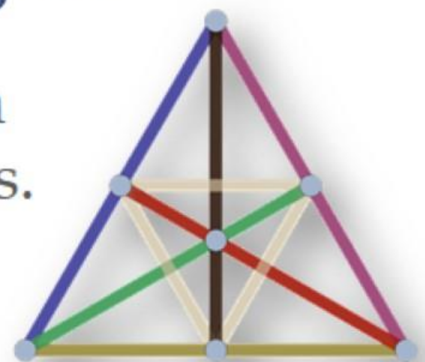
- [Kirkman 1847]: exists iff $n \equiv 1, 3 \pmod 6$.

Thomas
Kirkman
1806–1895

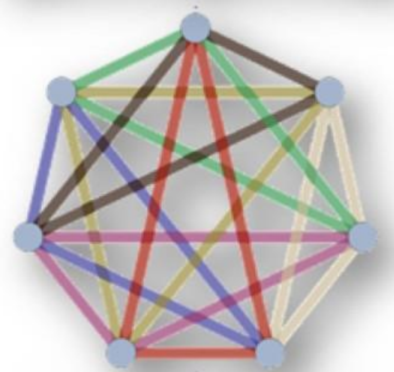
-

n	7	9	13	15	19
# *STS*	1	1	2	80	11084874829

Motivation I. packing designs



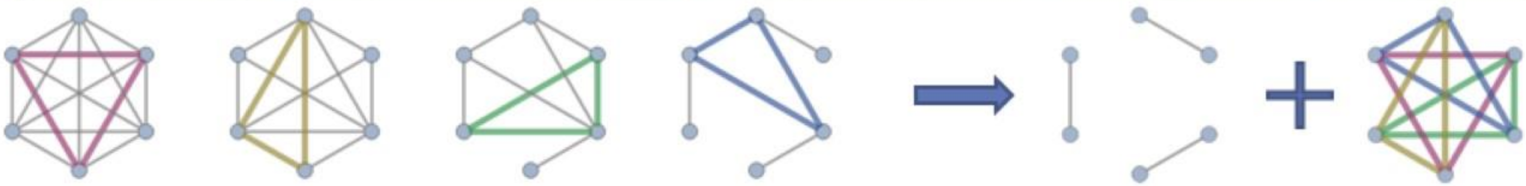
- Steiner triple system ≡ partition edges of a complete graph into edge-disjoint triangles.
  - general Steiner systems: hypergraphs…
  - constructions: group theory, geometry,…



- Recipe for a *near optimal* system (few pairs not covered) – *greedy packing*:
  - Order all $\binom{n}{3}$ triplets uniformly.
  - In this order, add each triplet that does not already intersect an existing one with an edge.
- *Exactly the triangle removal process…*

Packing via triangle removal process

- G_0 = complete graph on n vertices.
- $G_i \mapsto G_{i+1}$: remove edges of a **uniform triangle** in G_i.



$$\tau_0 = \min\{i : \; G(i) \text{ is triangle free}\}$$

- $\forall i$: $|E(G_i)| = \binom{n}{2} - 3i \implies$ equivalent questions:

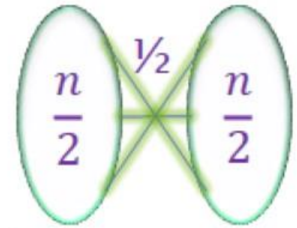
final number of edges

# of steps to terminate

Δ's packed via random greedy

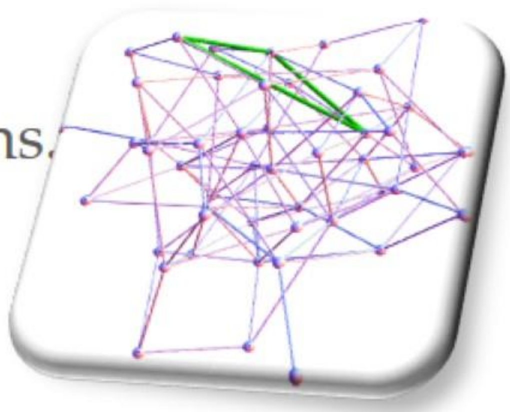
Motivation II. Δ–free networks

- What does a typical Δ-free graph on n vertices look like?
  - [Erdős, Kleitman, Rothschild '76]:
    a.e. such graph: bipartite with about $n^2/8$ edges.



  - [Prömel, Schickinger, Steger '02]:
    a.e. non-bipartite one: can be made bipartite by deleting a vertex.
  - [Osthus, Prömel, Taraz '03]: specifying the # of edges:
    a.e. Δ-free graph with $p\binom{n}{2}$ edges: bipartite as long as $\sqrt{\frac{3}{16}\frac{\log n}{n}} \leq p < 1$.

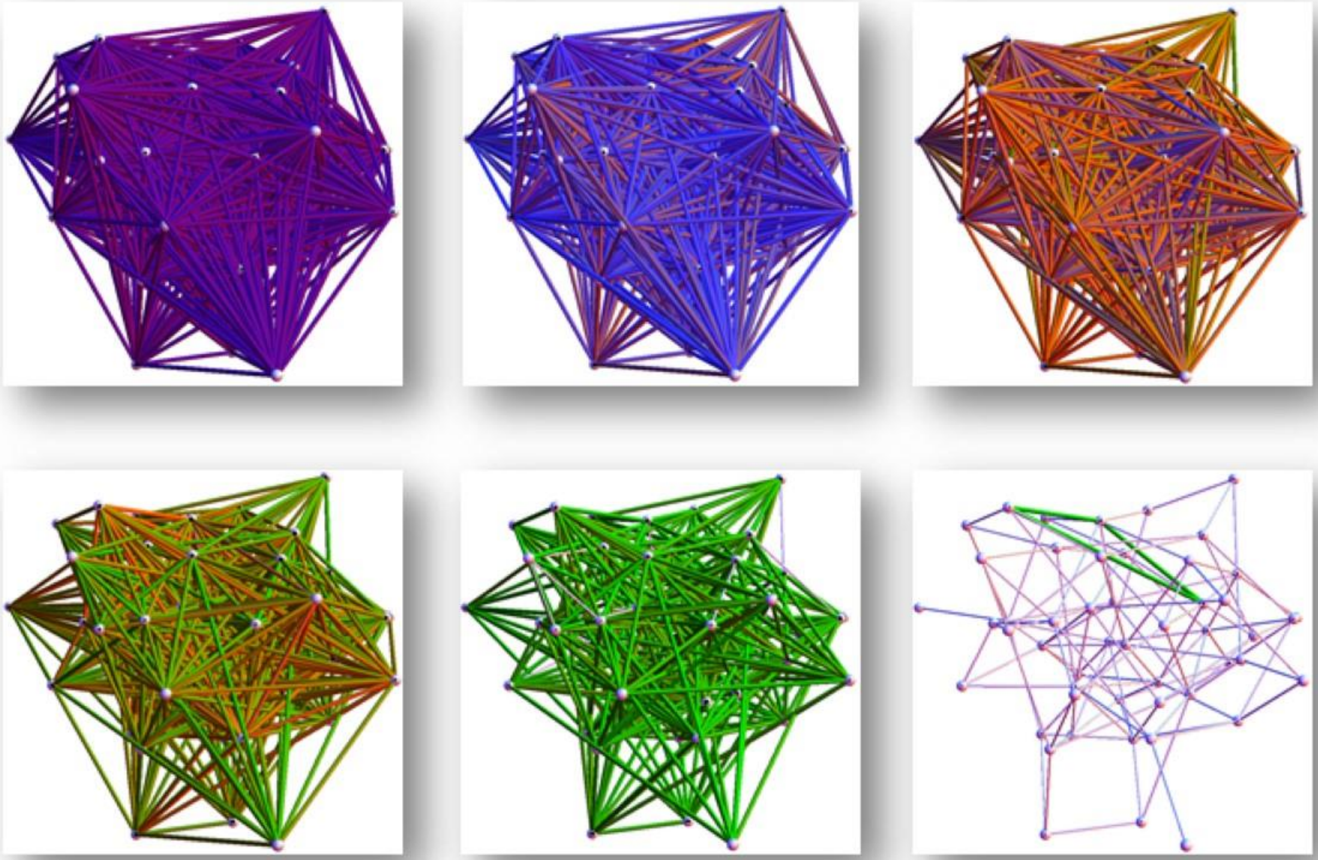
- Output of *Triangle removal process* :
  nontrivial distribution on Δ-free graphs.
  - similar to a random graph $\mathcal{G}(n,p)$;
  - and yet no triangles…



Motivation III. Ramsey theory

- DEFINITION: $R(3,t)$ = minimum n such that $\forall$ Δ-free graph on n vertices $\supset$ independent set I of size t.
  - $(\equiv)$: *How small can the independence number of a Δ-free graph be?*
  - [Erdős '61]: $\exists$ Δ-free graph with **no** I of size $c\sqrt{n}\log n$.
  - [Ajtai, Komlós, Szemerédi '80]: $\forall$ Δ-free graph $\supset$ I of size $c\sqrt{n\log n}$.
  - [Kim '95]: $\exists$ Δ-free graph with **no** I of size $C\sqrt{n\log n}$.
- Triangle removal process:
  - Tractable (hopefully) stochastic process ending with a Δ-free graph having (hopefully) *many* edges.
  - Towards the leading order constant for $R(3,t)$?

Sample run of the process



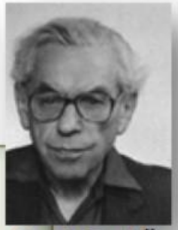
50 vertices, 371 steps, 115 remaining edges

An exponent of 3/2

- CONJECTURE [Bollobás, Erdős (1990)]:

  Expected final # edges has order $n^{3/2}$.

Béla Bollobás



Paul Erdős
1913–1996

- Intuition:
  - G_i should behave like an Erdős-Rényi graph $\mathcal{G}(n,m)$ with $m = |E(G_i)| = \binom{n}{2} - 3i$ edges.
  - At $m = \varepsilon n^{3/2}$ there are $\sim \frac{4}{3}\varepsilon^3 n^{3/2}$ triangles in $\mathcal{G}(n,m)$
  - Removing all triangles still leaves $\asymp n^{3/2}$ edges in G_{τ_0}.

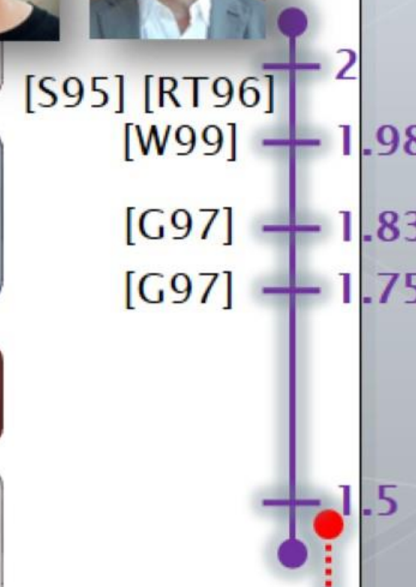
- $o(n^2)$ is already nontrivial: implies that *random greedy* constructs a near-optimal Steiner triple system.

Previous work

- [Spencer (1995)] and [Rödl, Thoma (1996)]:
  Final # edges is $o(n^2)$ with high probability.
- [Grable (1997)]:
  Final # edges is at most $n^{7/4+o(1)}$ w.h.p.

Best known upper bound. No lower bounds known.

- [Gordon, Kuperberg, Patashnik, Spencer ('96)]: simulations supporting the answer $n^{3/2+o(1)}$.
- [Wormald (1999)]:
  Final # edges is at most $n^{2-\frac{1}{57}+o(1)}$ w.h.p.
  (method applies to general case of random greedy packing in k-uniform hypergraphs.)



Previous work (ctd.)



- [Alon, Kim, Spencer (1997)]:
  - semi-random variant of the process finds nearly perfect hypergraph matchings. Specialized to Δ's:

__Variant__ process leaves $O\left(n^{3/2}\log^{3/2} n\right)$ *final edges w.h.p.*

  - Conjectured that *random greedy* matches these results:
    CONJECTURE [Alon, Kim, Spencer (1997)]:

Random greedy for k*-tuples w. pairwise intersections* $\leq k-2$ *has* $\mathbb{E}[\#\textit{uncovered }(k-1)\textit{-tuples}] \leq n^{k-1-\frac{1}{k-1}+o(1)}$.

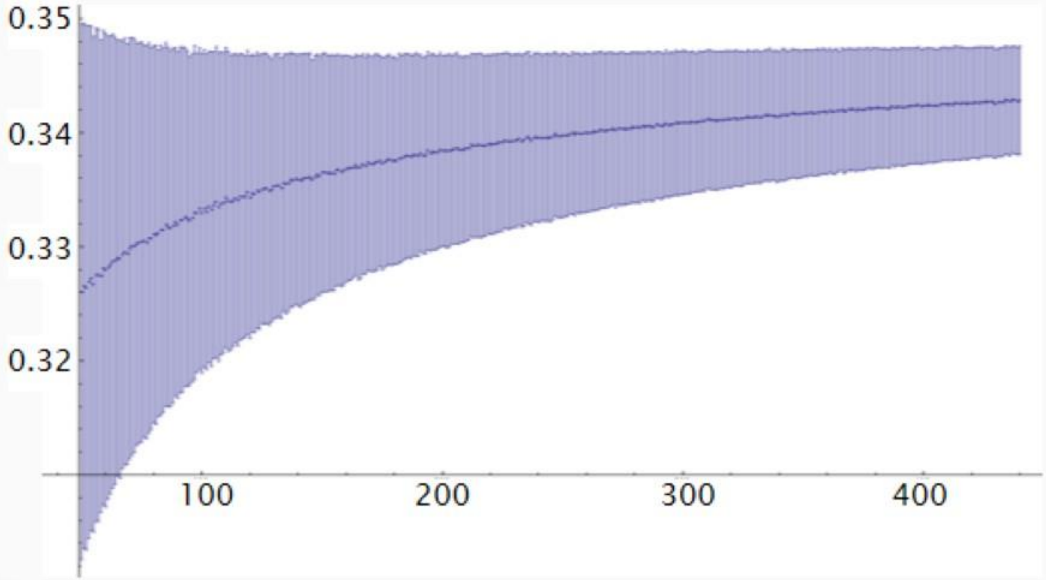
"...at the moment we cannot prove that this is the case even for $k=3$*"*

- Spencer offered \$200 for a proof of $n^{3/2+o(1)}$.

Main result

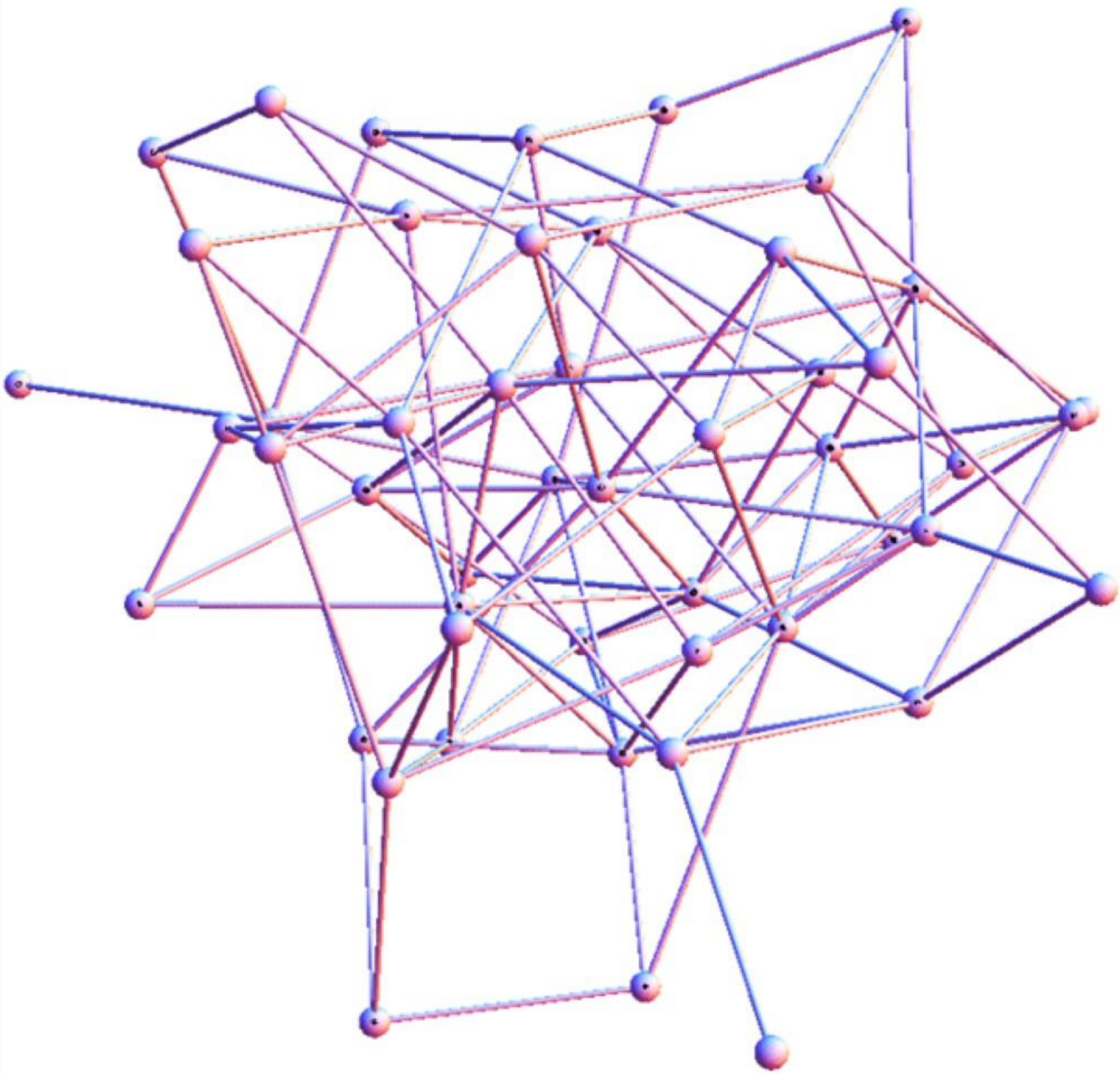
- THEOREM [Bohman, Frieze, L.]:

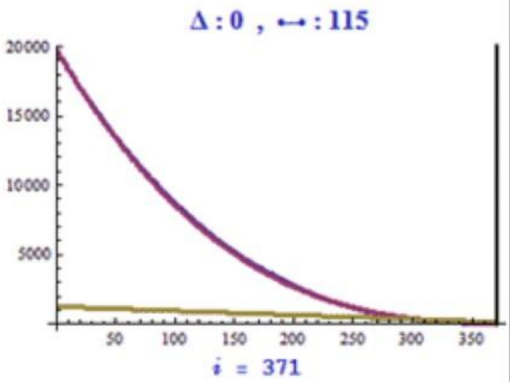
With high probability $\tau_0 = n^2/6 - n^{3/2+o(1)}$, *or equivalently,* $|E(G_{\tau_0})| = n^{3/2+o(1)}$.



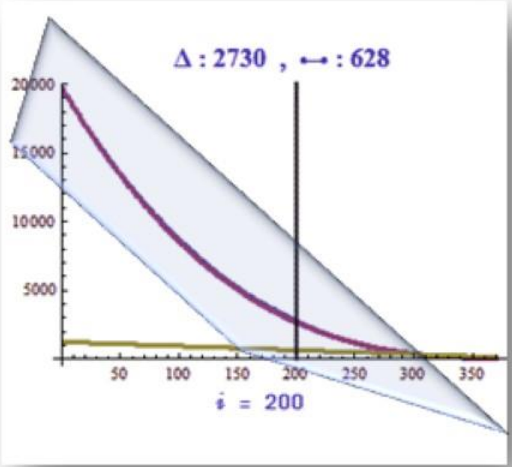
Simulations: Final # edges over $n^{3/2}$

Self correction





Self correction

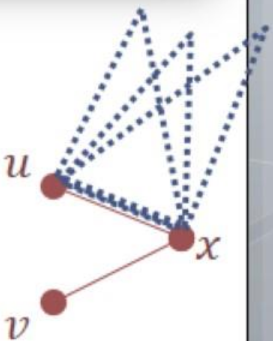


- Goal: maintain concentration of the total number of triangles.
  - Key: co-degrees.

$Q = Q(i) = \#\ \textit{triangles in}\ G_i$

$Y_{uv} = Y_{uv}(i) = \textit{co-degree of}\ u, v\ \textit{in}\ G_i$

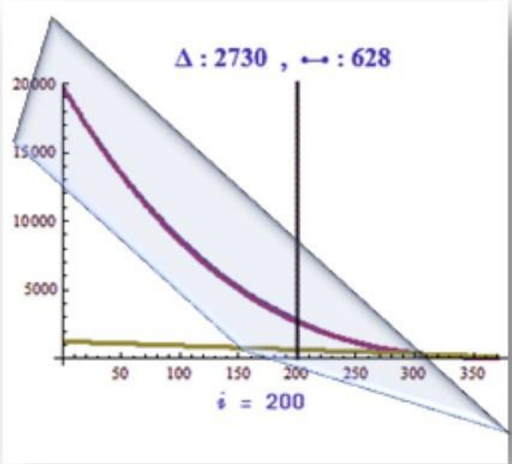
  - Co-degree evolution:



$$\mathbb{E}[\,\Delta Y_{uv} \mid \mathcal{F}_i\,] = -\frac{1}{Q}\sum_{x \in Y_{uv}} (Y_{ux} + Y_{vx} - \mathbb{1}_{u \sim v})$$

- Similar to the form $\mathbb{E}[dX] \le -a\,X$:
  - the larger X is, the bigger the drift towards its mean.

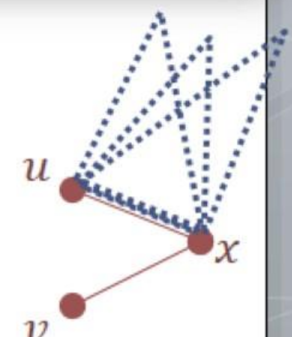
Self correction: the fine print



- Re-examining the key: co-degrees.

$Y_{uv} = Y_{uv}(i) = \textit{co-degree of } u, v \textit{ in } G_i$

1. once edge density drops to $p \approx \frac{1}{2}$ then $Y_{uv} \approx \mathrm{Bin}(n, p^2)$; STDEV $\asymp \sqrt{n}$.
2. this will match our mean of np^2 once $p \asymp n^{-1/4}$
   - Method breaks at $n^{7/4}$ edges…



- Crucial: *error estimates improve over time!*
- New general framework to support this.
  - Later used by [Bohman, Keevash] to improve bounds on $R(3,t)$ to within a factor of 4 (independently proved by [Pontiveros *et al.*]).

Context: the Δ–free process

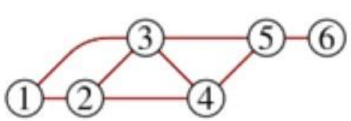
- Adding edges instead of deleting them:
  - G'_0= complete graph on n vertices.
  - $G'_i \mapsto G'_{i+1}$: add a **uniform edge** that does not add a Δ.
- [Erdős-Suen-Winkler (1995)]:
  Final # of edges in Δ-free process $= n^{3/2+o(1)}$ w.h.p.
- [Bohman (2010)]:
  Final # of edges in Δ-free process $\asymp n^{3/2}\sqrt{\log n}$ w.h.p.
- Main differences:
  1. Triangle-removal goes through $n^2/6 - n^{3/2+o(1)}$ steps *vs.* $n^{3/2+o(1)}$ steps in the Δ-free process.
  2. Δ-free is "well behaved" until the very end…

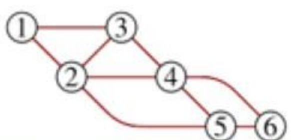
Context: the Δ–free process (ctd.)

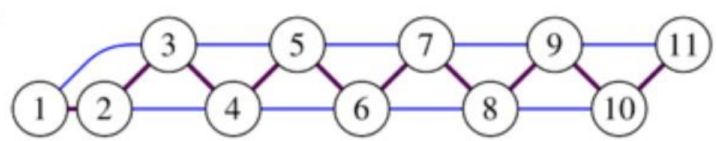
- Δ-free process:
  - Forbidden edges are pairs with a positive co-degree.
  - These are *negligible* until there are $\asymp n^{3/2}$ edges…
  - Coupling to $\mathcal{G}(n,m)$ readily gives a lower bound.
- Triangle removal:
  - Already when the edge density is a small **constant** ε # forbidden triangles ≫ # legal ones ...
  - Tracking the process to $p = n^{-\varepsilon}$ requires delicate control over geometry of remaining triangles.

Proof ingredients

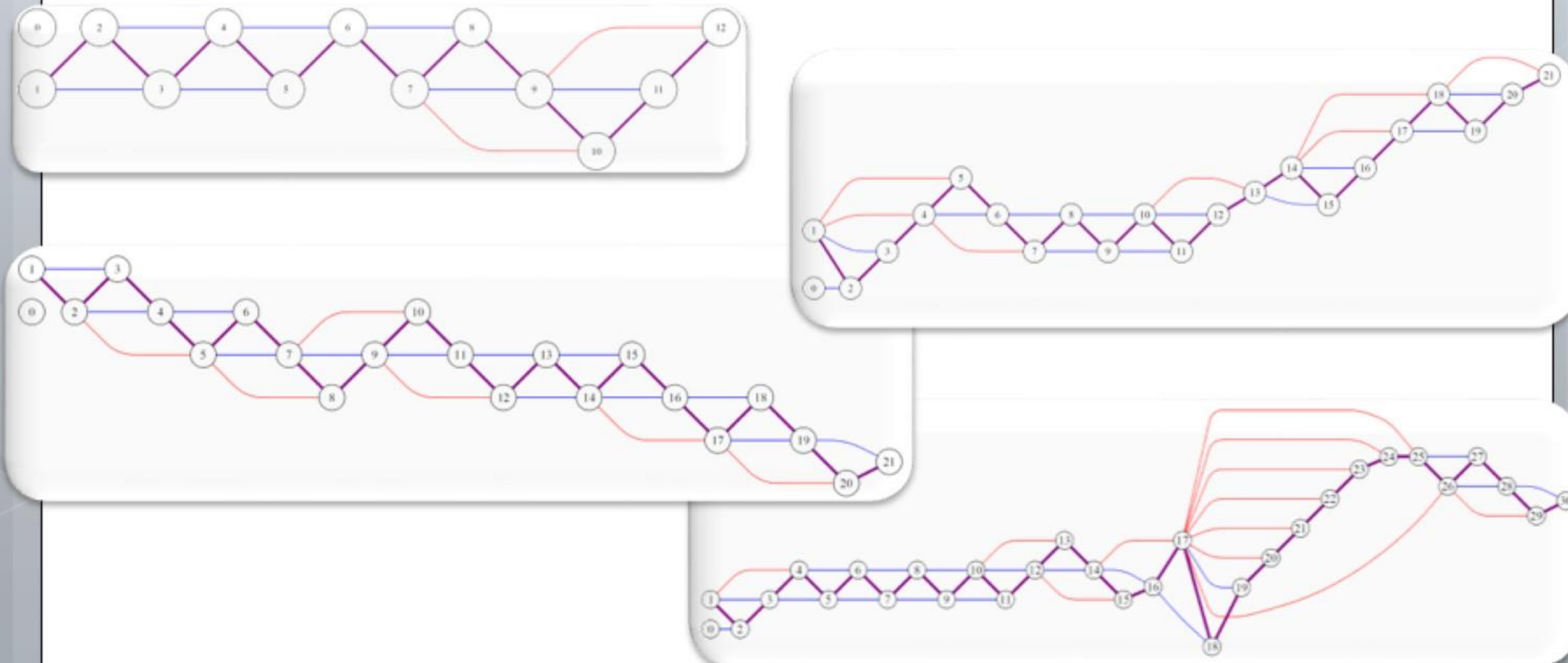
- Starting point: system of martingales tracking the evolution of $\mathrm{poly}(n)$ variables w.r.t. $\mathcal{G}(n,m)$ values.
- Self correction: errors *decrease* as process evolves.
- Objective: track *all* co-degrees followed by the # Δ's.
  - Naïve approach breaks at $7/4$ matching Grable's result via a very different method (physical barrier).
  - More ingredients help push the exponent further, but eventually subgraphs become too sparse to track…
- Construct canonical family of $e^{O(1/\varepsilon)}$ graphs by gluing $O(1/\varepsilon)$ triangles in a prescribed manner; track all graph homomorphisms from them.







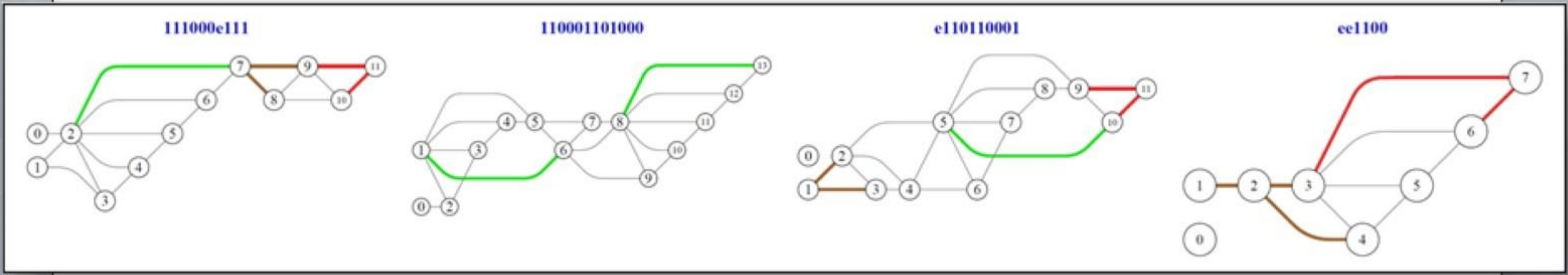
Triangular ladders

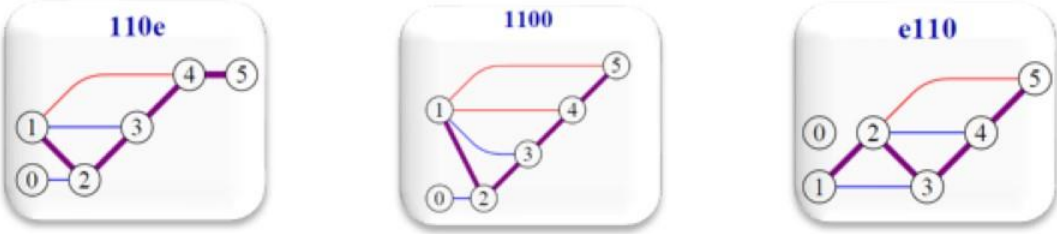


Some of the $\approx 2^{30}$ labeled rooted graphs whose homomorphisms-counts are tracked to imply $\left|E\left(G_{\tau_0}\right)\right| \le n^{3/2+\varepsilon}$ for $\varepsilon = \frac{1}{10}$.

Triangular ladders

- Controlling one ladder is achieved via longer ones.
- End game: crucially relies on the ladder's length…
- Each variable features a self correcting estimate.



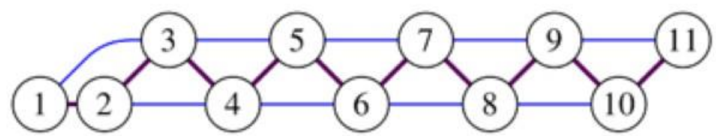


Triangular ladders: example

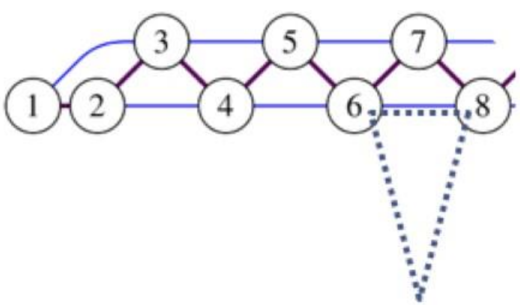
- A simple lemma:

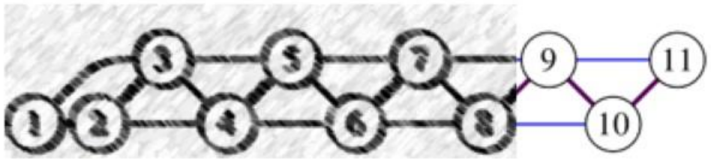
> If $(x_i)_{i\in I}$ and $(y_i)_{i\in I}$ sat. $|x_i - x| \le \delta_1$ and $|y_i - y| \le \delta_2$ for some $x, y \in \mathbb{R}$ and $\forall i \in I$ then
> $$\left|\sum_{i\in I} x_i y_i - \frac{1}{|I|}\left(\sum_{i\in I} x_i\right)\left(\sum_{i\in I} y_i\right)\right| \le 2|I|\delta_1\delta_2.$$

- E.g.., to control:



it suffices to handle:





Simple concrete example

- Recall: $Q(i) \triangleq \#\,\Delta\text{'s}\;;\quad Y_{uv}(i) \triangleq$ *co-degree of* u, v .
- THEOREM:

Suppose $|Y_{uv} - np^2| \le n^{3/4}$ *for all* u, v *and* $i \le i_0$.
Then with high probability $Q \le \frac{1}{6}n^3p^3 + \frac{1}{3}np^2$.

additive error decreases with time!

- Recipe for utilizing self correction:
  - Estimate expected change in terms of Q itself.
  - Set a threshold γ just below desired upper bound Γ.
  - Show that while $Q \in [\gamma, \Gamma]$ it is a supermartingale.
  - Concentrate Q with error probability n^{-100}.

Simple concrete example (ctd.)

- GOAL: $Q \le \frac{1}{6}n^3p^3 + \frac{1}{3}np^2$ given $|Y_{uv} - np^2| \le n^{3/4}$ $\forall u, v$
- PROOF:
  - Analysis of one-step change:

$$\mathbb{E}[\,\Delta Q \mid \mathcal{F}_i\,] = -\frac{1}{Q}\sum_{uvw\in Q}(Y_{uv} + Y_{vw} + Y_{uw} - 2) = 2 - \frac{1}{Q}\sum_{uv\in E} Y_{uv}^2$$

  - Since $\sum_{uv\in E} Y_{uv}^2 \ge 9Q^2/|E|$:

$$\mathbb{E}[\,\Delta Q \mid \mathcal{F}_i\,] = 2 - \frac{1}{Q}\sum_{uv\in E} Y_{uv}^2 \le 2 - \frac{18}{n^2p}\,Q\,.$$

  - Suppose $Q(i) > \frac{1}{6}n^3p^3 + \frac{1}{4}n^2p$. Then

$$\mathbb{E}[\,\Delta Q \mid \mathcal{F}_i\,] \le -3np^2 - \tfrac{5}{2}\,.$$

Simple concrete example (ctd.)

- Suppose Q just entered $[\gamma, \Gamma]$ for $\begin{cases} \gamma = n^2p/4 \\ \Gamma = n^2p/3 \end{cases}$.
- Set $Z = Q - \left(\frac{1}{6}n^3p^3 + \frac{1}{3}n^2p\right)$.
- As $\Delta p = -6/n^2$ the change in the scaling term is
$$\sim \frac{6}{n^2}\left[\frac{1}{2}n^3p^2 + \frac{1}{3}n^2\right] \sim 3np^2 + 2$$
  - Recall: $\mathbb{E}[\,\Delta Q \mid \mathcal{F}_i\,] \leq -3np^2 - \frac{5}{2}$.
- As long as $Q \in [\gamma, \Gamma]$ we get
$$\mathbb{E}\big[\,\Delta Z \mid \mathcal{F}_i\,\big] \leq -\tfrac{1}{2} + o(1) < 0\,,$$
a supermartingale.

Simple concrete example (ctd.)

- Next: concentrate $Z = Q - \left(\frac{1}{6}n^3p^3 - \frac{1}{3}n^2p\right)$.
- Number of steps remaining: $\leq |E| \asymp n^2p$
- Deviation considered: $\frac{1}{12}n^2p$.
- Lipschitz constant for one step:
  - Q changes by some $-(Y_{uv} + Y_{vw} + Y_{uw}) + O(1)$.
  - Scaling term changes by $\sim 3np^2 + 2$.
  - Together: $O\left(n^{3/4}\right)$ thanks to co-degree estimates!
- By Hoeffding's inequality:

$$\mathbb{P}\left[\exists j\colon Z(j) - Z(0) > \tfrac{1}{12}n^2p\right] \leq \exp(-c\sqrt{n}\,p)\ .$$

- W.h.p. we will never cross the $[\gamma, \Gamma]$ interval. ■

Lower bound

- THEOREM:

Suppose $Y_{uv} \sim np^2$ *for* $\forall\, u, v$ *and all*

$$p \geq p_0 = n^{-1/2+\varepsilon}.$$

Then w.h.p. the final # edges is at least $c\, n^{3/2-6\varepsilon}$.

- PROOF:
  - Important ingredient: a variant of the upper bound on Q with the *correct* additive error:

w.h.p. $Q \leq\sim \left[\frac{1}{6}n^3p^3 + \frac{1}{6}np^2\right]$ *at all times.*

  - (we demonstrated an additive error of $\frac{1}{3}np^2$.)

Lower bound (ctd.)

- Assume: $Q \leq \left[\frac{1}{6}n^3p^3 + \frac{1}{6}np^2\right] \; \forall p.$
- Consider time $p = p_1 = \delta/\sqrt{n}$ for small enough $\delta > 0$.
  - # edges: $|\mathrm{E}| \sim \frac{1}{2}n^2p = \frac{1}{2}\delta n^{3/2}$.
  - # triangles: $Q \leq \frac{1}{6}\delta^3 n^{3/2} + \frac{1}{6}\delta n^{3/2}$.

negligible

why $\frac{1}{6}$ mattered: $|E|/3$

- If $Q < \frac{1-\alpha}{6}\delta n^{3/2}$ then $Q < \frac{1-\alpha}{3}|E|$ and necessarily there will be $cn^{3/2}$ edges at the end of the process (done).
  - $\Rightarrow$ may assume: $Q \approx \frac{1}{3}|E|$.

Lower bound (ctd.)

- At time $p = p_1 = \delta/\sqrt{n}$ for small enough $\delta > 0$:
  - # edges: $|\mathrm{E}| = c\, n^{3/2}$; # triangles: $Q \approx \frac{1}{3}|E|$.
- If $cn^{3/2}$ edges have no triangles on them ⇒ done.
  - ⇒ may assume: almost ∀ edge incident to a Δ .
- Combined: **almost all triangles are edge-disjoint**.

At time p_0 every co-degree is $\sim n^{2\varepsilon}$
⇒ every triangle is incident to $\sim 3n^{2\varepsilon}$ others.

At time p_1 there are $cn^{3/2}$ "isolated" triangles.

Lower bound (ctd.)

- Look at triangles just before they became isolated:
  - Mark a triangle once it has an edge with co-degree 1 (no other triangles resting on this *special* edge.)
  - Filter a subset $\mathfrak{X}$ of marked Δ's where if $R \in \mathfrak{X} \Rightarrow$ no $S \sim R$ (incident triangle) is in $\mathfrak{X}$ nor any $T \sim S$.
  - $|\mathfrak{X}| \geq n^{-4\varepsilon} \times cn^{3/2}$ (pay $n^{2\varepsilon}$ per level by co-degrees).
- If an arbitrary neighbor S of $R \in \mathfrak{X}$ is drawn before any of its own neighbors: the *special* edge of R survives!
  - $\mathbb{P}$(this event) $\gtrsim n^{-2\varepsilon}$; events are independent.
  - W.h.p. final # edges is $cn^{-2\varepsilon}|\mathfrak{X}| \geq cn^{3/2-6\varepsilon}$. ■

Open problems

1. Establish the order of the final number of edges.

2. Study the graph properties of the final output (a nontrivial distribution over triangle-free graphs).

3. Compare final output of the *triangle removal process* with that of the *triangle-free process*.

4. Obtain the leading constant for $R(3,t)$.

Thank you