Preferential Attachment Graphs are somewhere-dense

Jan Dreier, **Philipp Kuinke**, Peter Rossmanith

TACO 2018

RWTH Aachen University

Motivation

Sparsity

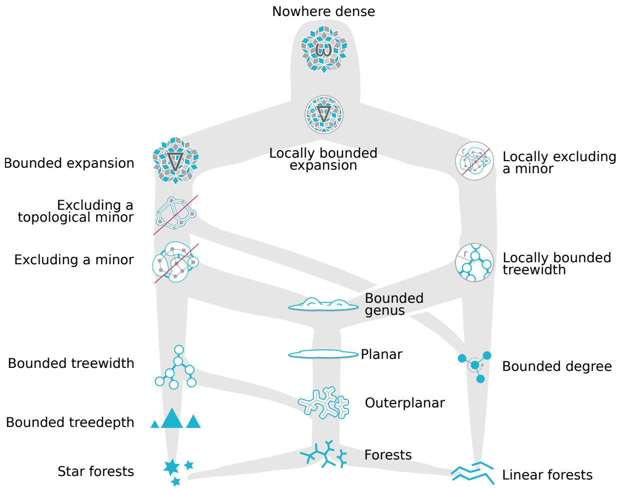
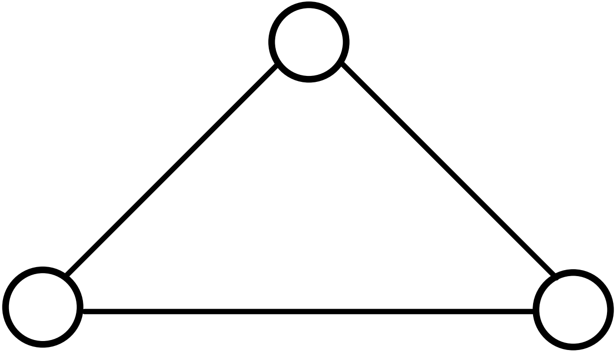
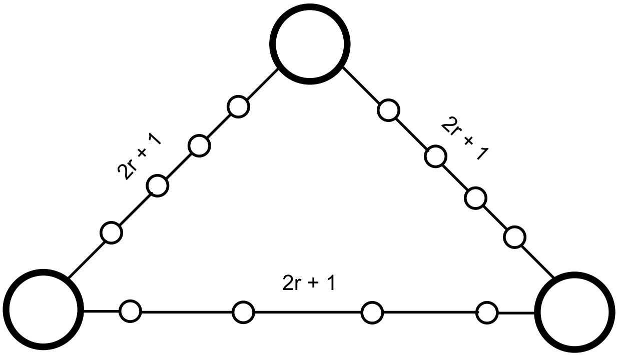


Image by Felix Reidl

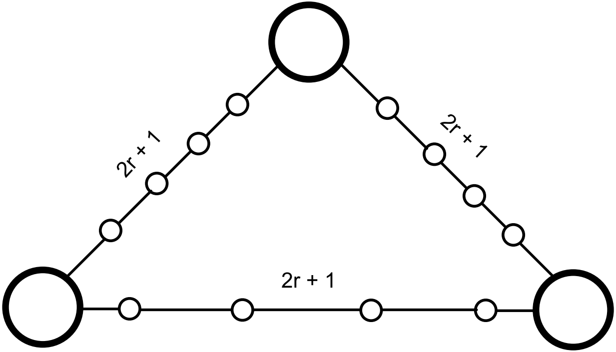
Sparsity: r-shallow topological minor



Sparsity: r-shallow topological minor

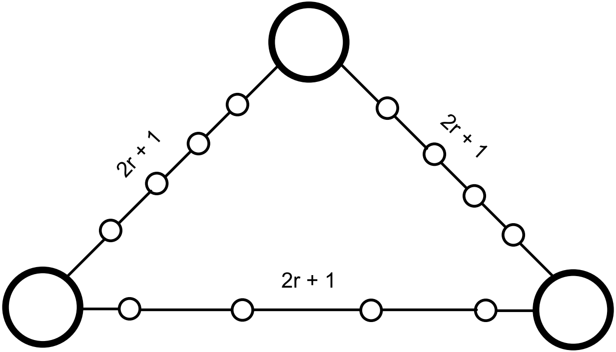


Sparsity: r-shallow topological minor



$G \,\widetilde{\nabla}\, r$:= The set of all r-shallow topological minors of G.

Sparsity: r-shallow topological minor



$G \mathbin{\widetilde{\nabla}} r$:= The set of all r-shallow topological minors of G.

$$\omega(G \mathbin{\widetilde{\nabla}} r) = \max_{H \in G \mathbin{\widetilde{\nabla}} r} \omega(H) \quad \text{(clique size)}$$

Definition (Nowhere-dense)

A graph class $\mathcal{G}$ is nowhere-dense if there exists a function f, such that for all r and all $G \in \mathcal{G}$, $\omega(G \mathbin{\widetilde{\nabla}} r) \leq f(r)$.

Sparsity

Definition (Nowhere-dense)

A graph class $\mathcal{G}$ is nowhere-dense if there exists a function f, such that for all r and all $G \in \mathcal{G}$, $\omega(G \,\widetilde{\nabla}\, r) \leq f(r)$.

Definition (Somewhere-dense)

A graph class $\mathcal{G}$ is somewhere-dense if for all functions f there exists an r and a $G \in \mathcal{G}$, such that $\omega(G \,\widetilde{\nabla}\, r) > f(r)$.

Sparsity

Definition (Nowhere-dense)

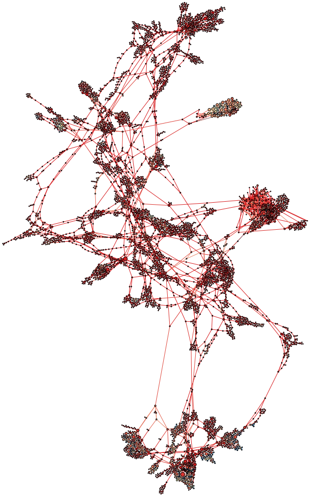
A graph class $\mathcal{G}$ is nowhere-dense if there exists a function f, such that for all r and all $G \in \mathcal{G}$, $\omega(G \,\widetilde{\nabla}\, r) \leq f(r)$.

Definition (Somewhere-dense)

A graph class $\mathcal{G}$ is somewhere-dense if for all functions f there exists an r and a $G \in \mathcal{G}$, such that $\omega(G \,\widetilde{\nabla}\, r) > f(r)$.

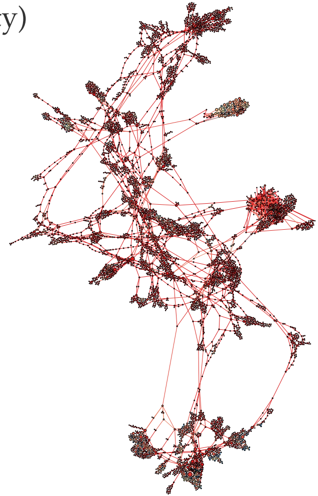
$\mathcal{G}$ is not nowhere-dense $\Leftrightarrow$ $\mathcal{G}$ is somewhere-dense

Typical Properties of Complex Networks



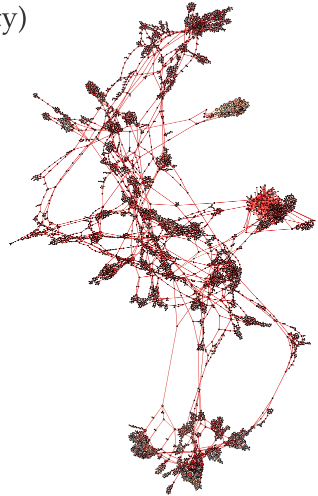
Typical Properties of Complex Networks

- low diameter (small world property)



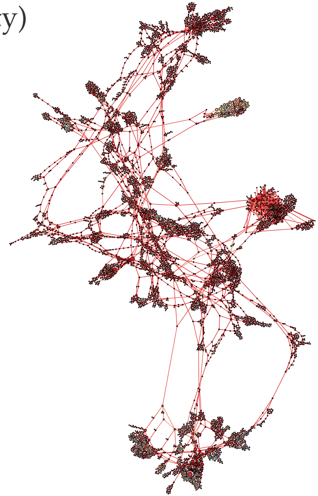
Typical Properties of Complex Networks

- low diameter (small world property)
- locally dense, but globally sparse



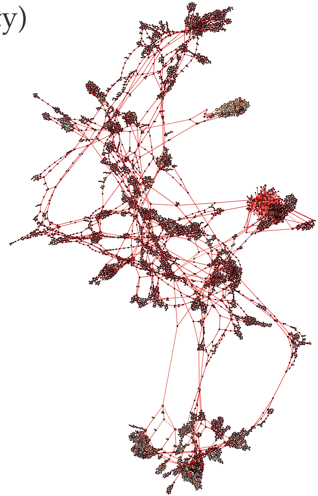
Typical Properties of Complex Networks

- low diameter (small world property)
- locally dense, but globally sparse
- heavy tail degree distribution



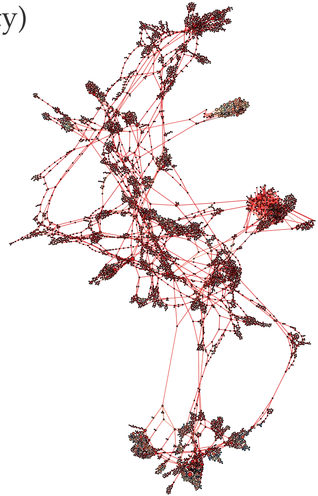
Typical Properties of Complex Networks

- low diameter (small world property)
- locally dense, but globally sparse
- heavy tail degree distribution
- clustering



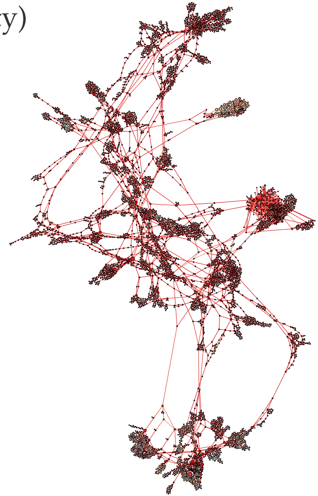
Typical Properties of Complex Networks

- low diameter (small world property)
- locally dense, but globally sparse
- heavy tail degree distribution
- clustering
- community structure

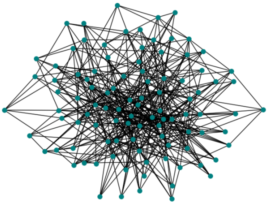


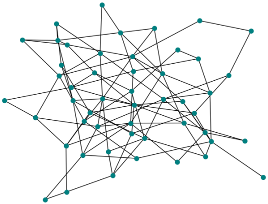
Typical Properties of Complex Networks

- low diameter (small world property)
- locally dense, but globally sparse
- heavy tail degree distribution
- clustering
- community structure
- scale freeness



Random Graphs





Random graphs with the goal of modeling real-world data:

- Mathematically analyzable
- Generation of infinite instances

Sparse in the limit

Definition (a.a.s. nowhere-dense)

A random graph model $\mathcal{G}$ is a.a.s. nowhere-dense if there exists a function f such that for all r

$$\lim_{n \to \infty} \mathrm{P}[\omega(G_n \,\widetilde{\nabla}\, r) \leq f(r)] = 1$$

where G_n is a random variable modeling a graph with n vertices randomly drawn from $\mathcal{G}$.

Sparse in the limit

Definition (a.a.s. somewhere-dense)

A random graph model $\mathcal{G}$ is a.a.s. somewhere-dense if for all functions f there exists an r, such that

$$\lim_{n\to\infty} \mathrm{P}[\omega(G_n \,\widetilde{\nabla}\, r) > f(r)] = 1$$

where G_n is a random variable modeling a graph with n vertices randomly drawn from $\mathcal{G}$.

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

1. $p = 1/n$

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

1. $p = 1/n$ $\rightarrow$ a.a.s. nowhere-dense

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

1. $p = 1/n$ $\rightarrow$ a.a.s. nowhere-dense
2. $p = 1 - 1/n$

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

1. $p = 1/n$ $\rightarrow$ a.a.s. nowhere-dense
2. $p = 1 - 1/n$ $\rightarrow$ a.a.s. somewhere-dense

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

1. $p = 1/n$ $\rightarrow$ a.a.s. nowhere-dense
2. $p = 1 - 1/n$ $\rightarrow$ a.a.s. somewhere-dense
3. $p = 1/2$

Sparse in the limit (not as clear cut)

Assume you have a random graph on n vertices, such that it is with probability p complete and with probability $1 - p$ empty:

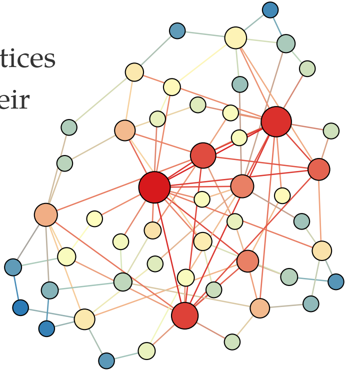
1. $p = 1/n$ $\rightarrow$ a.a.s. nowhere-dense
2. $p = 1 - 1/n$ $\rightarrow$ a.a.s. somewhere-dense
3. $p = 1/2$ $\rightarrow$ Neither!

Preferential attachment graphs

"the rich get richer", "preferential attachment", "Barabási–Albert graphs"

Start with some small fixed graph.

Add vertices. Connect them to m vertices with a probability proportional to their degrees.

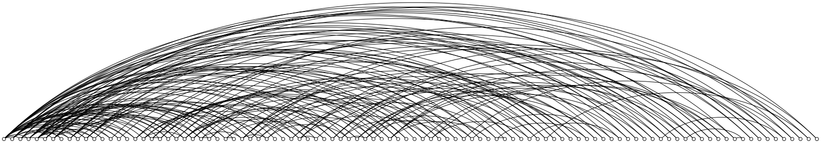


Interesting properties:

- power law degree distribution
- scale free

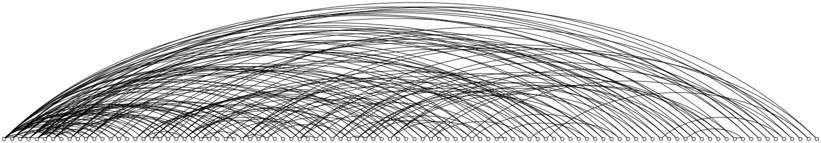
Preferential attachment graphs

$m = 2, n = 100$:



Preferential attachment graphs

$m = 2, n = 100$:



$$\mathrm{E}[d_m^n(v_i)] \sim m\sqrt{n/i}$$

Tail bounds

Degree Concentrations

- Tail bounds exists for number of vertices with degree d. [Bollobás et al. 2001]

Degree Concentrations

- Tail bounds exists for number of vertices with degree d. [Bollobás et al. 2001]
- Via Martingales + Azuma-Hoeffding inequality

Degree Concentrations

- Tail bounds exists for number of vertices with degree d. [Bollobás et al. 2001]
- Via Martingales + Azuma-Hoeffding inequality
- Does not work for large d (i.e. order $\sqrt{n}$)

Degree Concentrations

- Tail bounds exists for number of vertices with degree d. [Bollobás et al. 2001]
- Via Martingales + Azuma-Hoeffding inequality
- Does not work for large d (i.e. order $\sqrt{n}$)
- But we need high degree vertices!

Concentration of a single vertex

No vertex is sharply concentrated!

Concentration of a single vertex

No vertex is sharply concentrated!

$$\mathrm{P}[d_1^n(v_t) = 1]$$

Concentration of a single vertex

No vertex is sharply concentrated!

$$\mathrm{P}[d_1^n(v_t) = 1] = \prod_{i=t}^{n}(1 - \frac{1}{2i - 1})$$

Concentration of a single vertex

No vertex is sharply concentrated!

$$\mathrm{P}[d_1^n(v_t) = 1] = \prod_{i=t}^{n}(1 - \frac{1}{2i-1}) \geq \frac{1}{n}$$

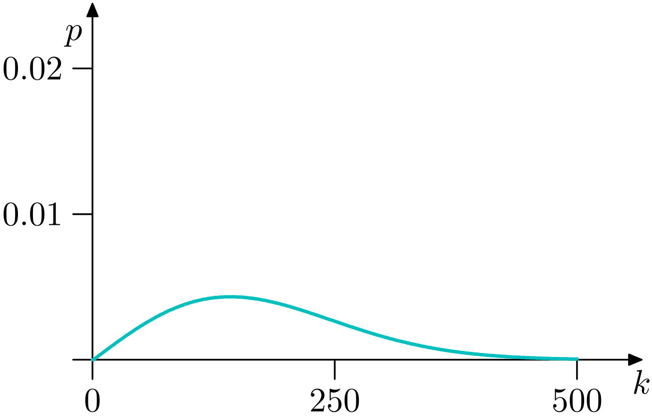
Concentration of a single vertex

No vertex is sharply concentrated!

$$\mathrm{P}[d_1^n(v_t) = 1] = \prod_{i=t}^{n}(1 - \frac{1}{2i-1}) \geq \frac{1}{n}$$

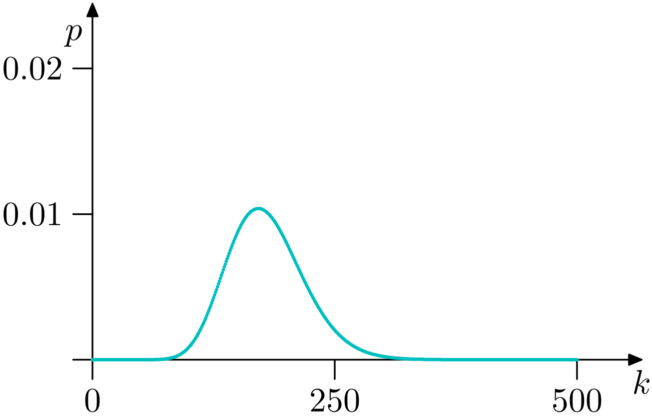
We can not hope for general *exponential* bounds.

Concentration of a single vertex



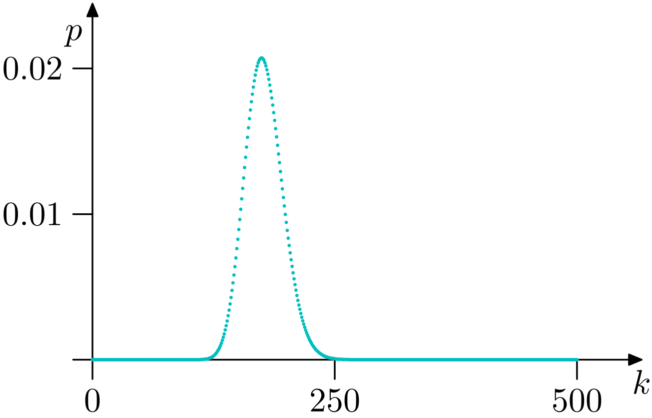
Distribution of $d_1^{10000}(v_1)$.

Concentration of a single vertex



Distribution of $d_1^{10000}(v_1)$ conditioned under $d_1^{100}(v_1) = 18$.

Concentration of a single vertex



Distribution of $d_1^{10000}(v_1)$ conditioned under $d_1^{1000}(v_1) = 56$.

The rich stay rich

Theorem

Let $0 < \varepsilon \leq 1/40$, $t, m, n \in \mathbf{N}$, $t > \frac{1}{\varepsilon^6}$ *and* $S \subseteq \{v_1, \ldots, v_t\}$. *Then*

$$\mathrm{P}\Big[(1-\varepsilon)\sqrt{\frac{n}{t}}d_m^t(S) < d_m^n(S) < (1+\varepsilon)\sqrt{\frac{n}{t}}d_m^t(S) \text{ for all } n \geq t \ \Big|\ d_m^t(S)\Big]$$
$$\geq 1 - \ln(15t)e^{\varepsilon^{-O(1)}d_m^t(S)}.$$

The rich stay rich

Theorem (The approximate version)

Let $\varepsilon \geq 0$*,* $t, m, n \in \mathbf{N}$*, and* $S \subseteq \{v_1, \ldots, v_t\}$*:*

$$\mathrm{P}\Big[(1-\varepsilon)\,\mathrm{E}[d_m^n(S)] < d_m^n(S) < (1+\varepsilon)\,\mathrm{E}[d_m^n(S)] \;\Big|\; d_m^t(S)\Big] \geq 1 - e^{-\varepsilon\, d_m^t(S)}$$

The rich stay rich

Theorem (The approximate version)

Let $\varepsilon \geq 0$, $t, m, n \in \mathbf{N}$, *and* $S \subseteq \{v_1, \ldots, v_t\}$:

$$\mathrm{P}\left[(1-\varepsilon)\,\mathrm{E}[d_m^n(S)] < d_m^n(S) < (1+\varepsilon)\,\mathrm{E}[d_m^n(S)] \,\middle|\, d_m^t(S)\right] \geq 1 - e^{-\varepsilon d_m^t(S)}$$

- The rich stay rich

The rich stay rich

Theorem (The approximate version)

Let $\varepsilon \geq 0$, $t, m, n \in \mathbf{N}$, *and* $S \subseteq \{v_1, \ldots, v_t\}$*:*

$$\mathrm{P}\Big[(1-\varepsilon)\,\mathrm{E}[d_m^n(S)] < d_m^n(S) < (1+\varepsilon)\,\mathrm{E}[d_m^n(S)] \;\Big|\; d_m^t(S)\Big] \geq 1 - e^{-\varepsilon d_m^t(S)}$$

- The rich stay rich
- At first there is chaos

The rich stay rich

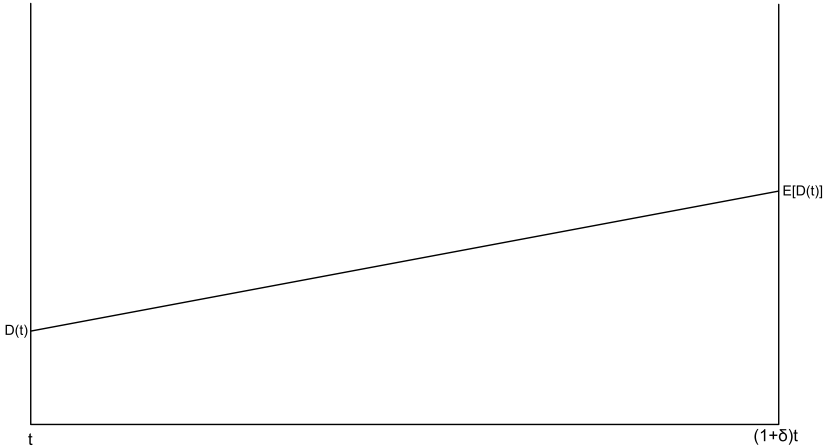
Theorem (The approximate version)

Let $\varepsilon \geq 0$, $t, m, n \in \mathbf{N}$, *and* $S \subseteq \{v_1, \ldots, v_t\}$*:*

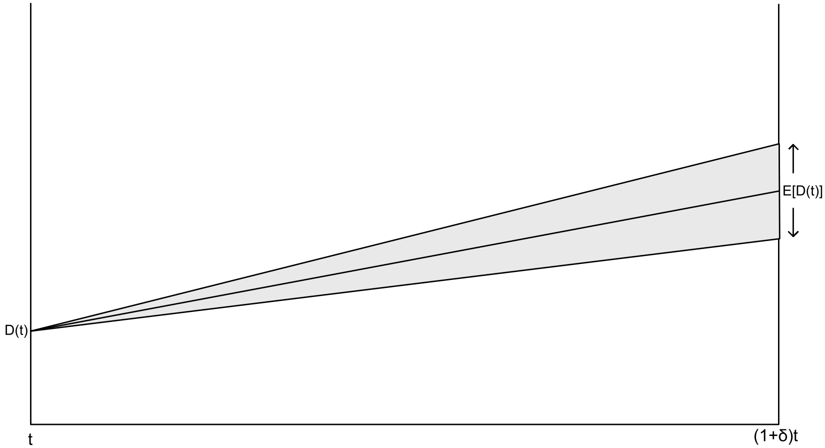
$$\mathrm{P}\Big[(1-\varepsilon)\,\mathrm{E}[d_m^n(S)] < d_m^n(S) < (1+\varepsilon)\,\mathrm{E}[d_m^n(S)] \;\Big|\; d_m^t(S)\Big] \geq 1 - e^{-\varepsilon\, d_m^t(S)}$$

- The rich stay rich
- At first there is chaos
- If we have information for t we can better predict $n > t$

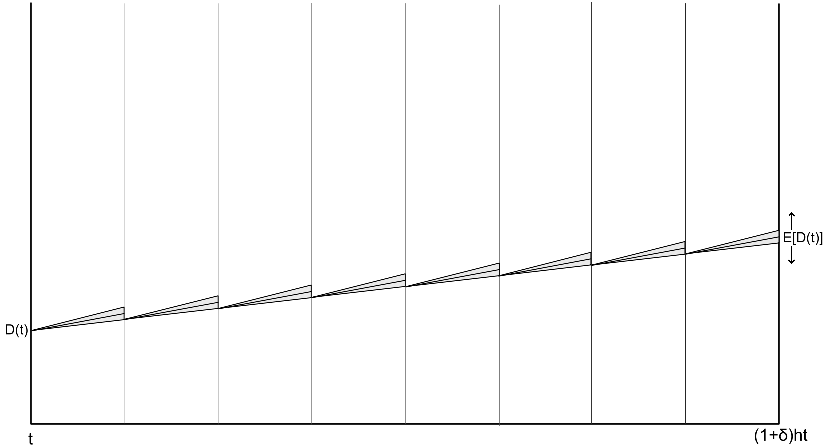
Proving the theorem



Proving the theorem



Proving the theorem



Somewhere-Dense

Large cliques

Theorem

G_m^n contains a.a.s. a one-subdivided clique of size $\sim \log(n)$.

Large cliques

Theorem

G_m^n contains a.a.s. a one-subdivided clique of size $\sim \log(n)$.

Corollary

G_m^n is a.a.s. somewhere-dense for $m \geq 2$.

How we get principals

k sets of vertices

How we get principals

k sets of vertices

If a set of s vertices has degree d one vertex has to have degree at least d/s.

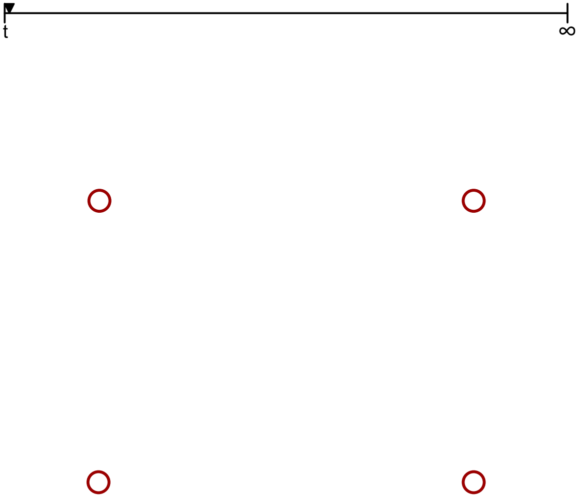
How we get principals

k sets of vertices

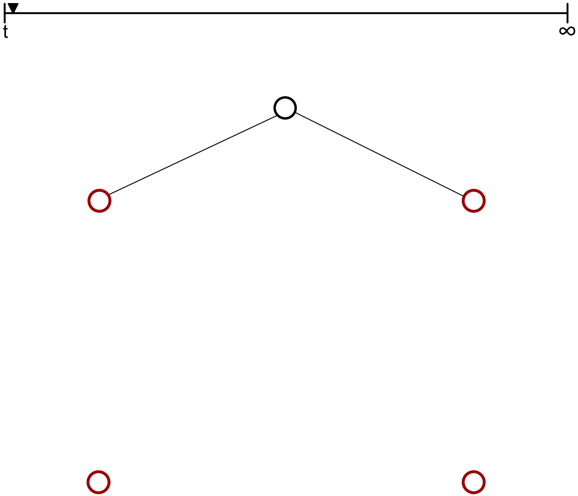
If a set of s vertices has degree d one vertex has to have degree at least d/s.

$\rightarrow$ Ensure with tail bounds it also has high degree in the future.

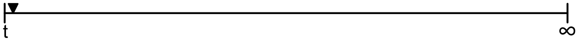
Building cliques



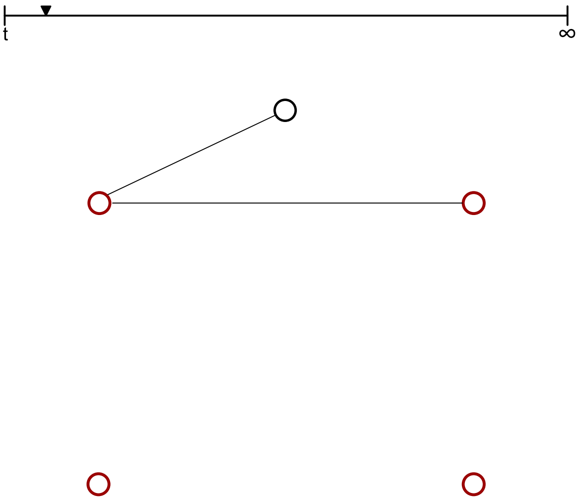
Building cliques



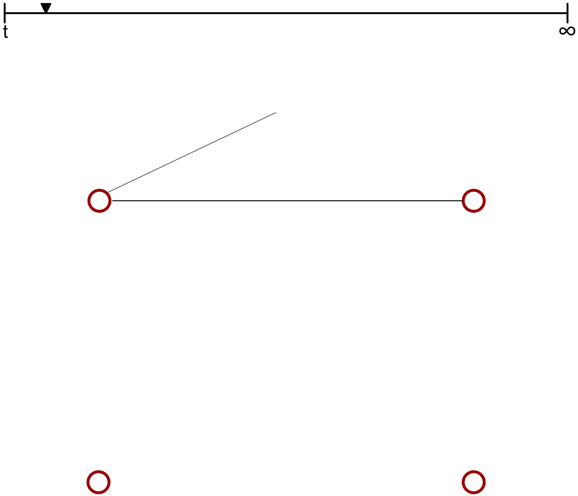
Building cliques



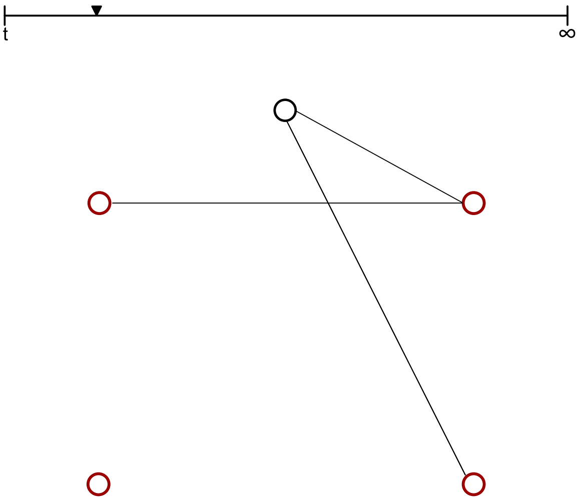
Building cliques



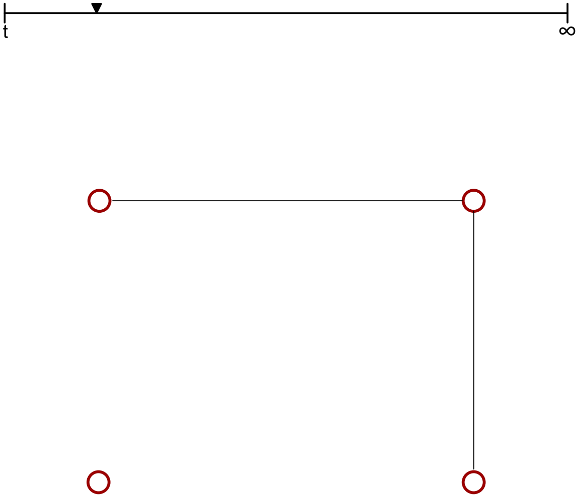
Building cliques



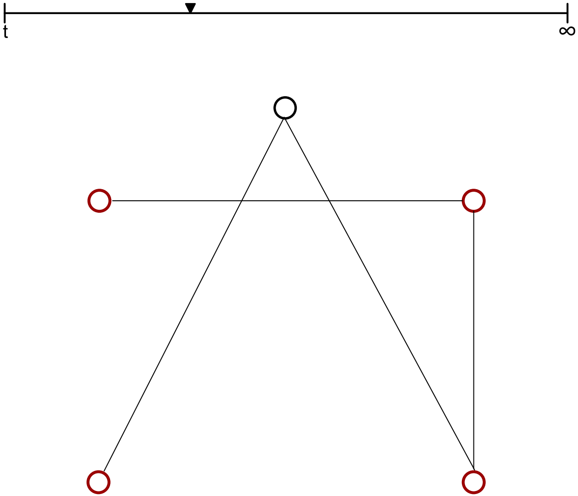
Building cliques



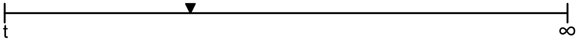
Building cliques

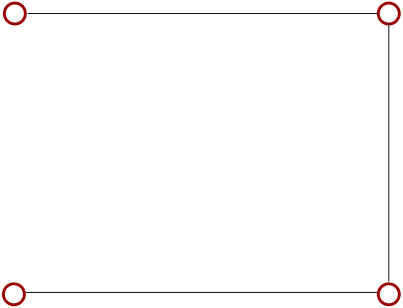


Building cliques

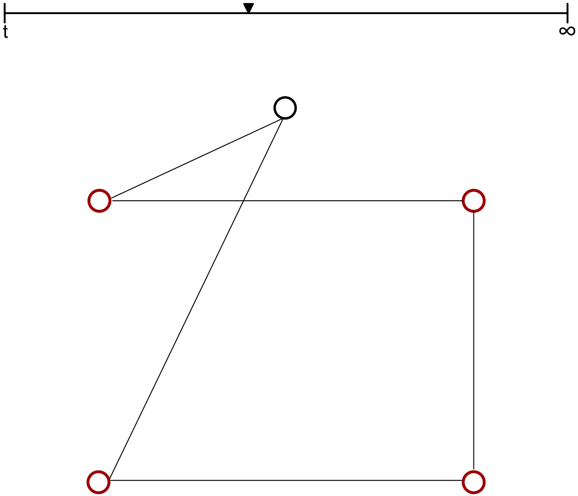


Building cliques

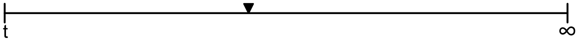


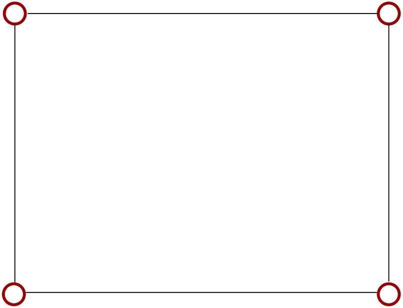


Building cliques

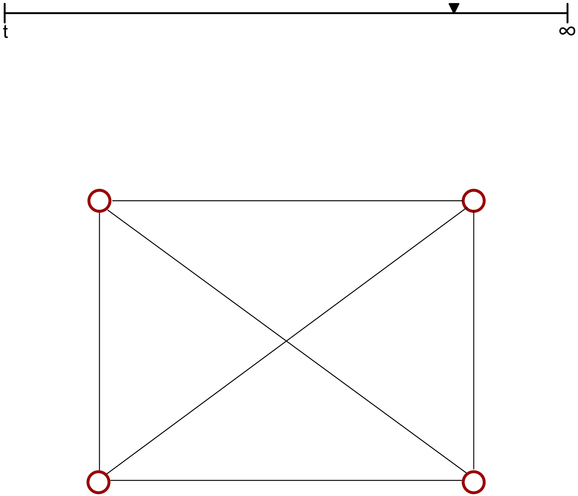


Building cliques





Building cliques



Connecting principals: Why we need $\sqrt{i}$

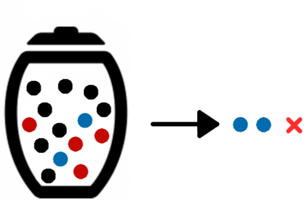
Connecting principals: Why we need $\sqrt{i}$

Step i:

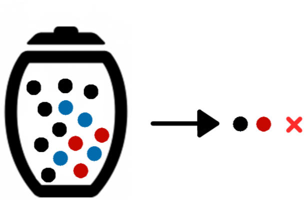
- $\sqrt{i}$ red
- $\sqrt{i}$ blue
- remainder black

Two balls drawn, success if red and blue

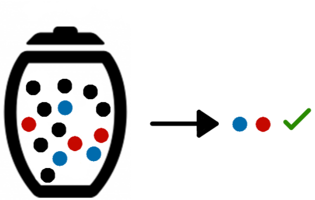
Connecting principals: Why we need $\sqrt{i}$



Connecting principals: Why we need $\sqrt{i}$



Connecting principals: Why we need $\sqrt{i}$



Connecting principals: Why we need $\sqrt{i}$

$$1 - \prod_{i=10}^{\infty}\left(1 - 2\left(\frac{\sqrt{i}}{i}\right)^2\right) = 1$$

Connecting principals: Why we need $\sqrt{i}$

$$1 - \prod_{i=10}^{\infty}\left(1 - 2\left(\frac{\sqrt{i}/\log(i)}{i}\right)^2\right) \neq 1$$

Conclusion

Conclusion

- Tail bounds for vertices where we know an earlier degree

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure
- Preferential Attachment graphs are a.a.s. somewhere-dense

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure
- Preferential Attachment graphs are a.a.s. somewhere-dense
- FO-model checking algorithm not directly applicable

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure
- Preferential Attachment graphs are a.a.s. somewhere-dense
- FO-model checking algorithm not directly applicable
- What about more general PA-graphs with δ parameter?

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure
- Preferential Attachment graphs are a.a.s. somewhere-dense
- FO-model checking algorithm not directly applicable
- What about more general PA-graphs with δ parameter?
  - $\delta = 0$: Our model

Conclusion

- Tail bounds for vertices where we know an earlier degree
- Tail bounds could be used to prove further structure
- Preferential Attachment graphs are a.a.s. somewhere-dense
- FO-model checking algorithm not directly applicable
- What about more general PA-graphs with δ parameter?
  - $\delta = 0$: Our model
  - $\delta = \infty$: Uniform attachment

References

Barabási, Albert-László and Réka Albert (1999). "Emergence of scaling in random networks". In: *Science* 286.5439, pp. 509–512.

Béla Bollobás Oliver Riordan, Joel Spencer and Gábor Tusnády (2001). "The Degree Sequence of a Scale-free Random Graph Process". In: *Random Struct. Algorithms* 18.3, pp. 279–290. ISSN: 1042-9832.

Martin Grohe, Stephan Kreutzer and Sebastian Siebertz (2017). "Deciding First-Order Properties of Nowhere Dense Graphs". In: *Journal of the ACM* 64.3, p. 17.

Nešetřil, Jaroslav and Patrice Ossona de Mendez (2012). *Sparsity*. Springer.

Van Der Hofstad, Remco (2016). *Random graphs and complex networks*. Vol. 1. Cambridge University Press.