# Small-worlds, complex networks and random graphs

Remco van der Hofstad

Spring School Complex Networks
Darmstadt
March 2-6, 2020

Joint work with:

- ▷ G. Hooghiemstra (Delft)
- ▷ P. Van Mieghem (Delft)
- ▷ S. Bhamidi (North Carolina)
- ▷ J. Komjáthy (TU/e)
- ▷ D. Znamensky (Philips Research)
- ▷ H. van den Esker (Delft)
- ▷ S. Dommers (TU/e)

# Plan lectures

Lecture 1:
Real-world networks and random graphs

Lecture 2:
Local weak convergence: theory

Lecture 3:
Local weak convergence of random graphs

Lecture 4:
The giant in random graphs is almost local

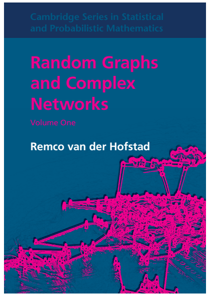
Lecture 5:
Small-world structure of random graphs

# Material

▷ Intro random graphs:
Random Graphs and Complex Networks Volume 1

`http://www.win.tue.nl/~rhofstad/NotesRGCN.html`

Volume 2: in preparation on same site




Treat selected parts of Chapters I.1, I.6–I.8 and II.2–II.8.

Argument are probabilistic, using

▷ first and second moment method;

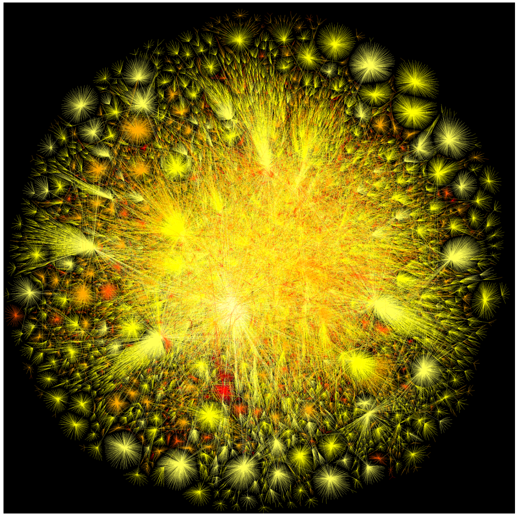
▷ branching process approximations.

# Lecture 1:

# Real-world networks and random graphs

# Complex networks

Yeast protein interaction network[a]



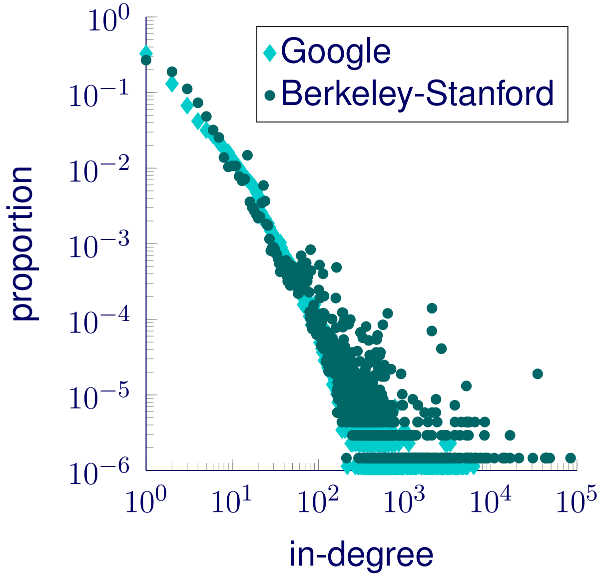
Internet 2010[b]

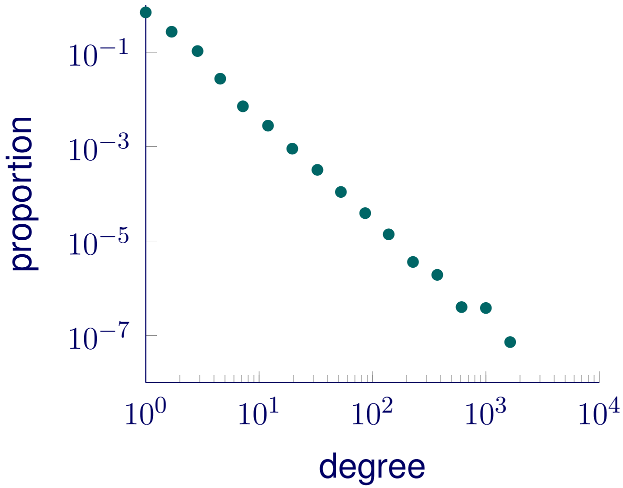
Attention focussing on unexpected commonality.

[a] Barabási & Óltvai 2004

[b] Opte project `http://www.opte.org/the-internet`

# Scale-free paradigm

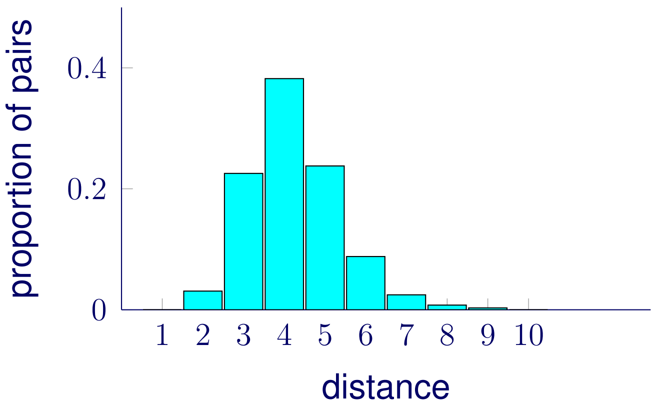


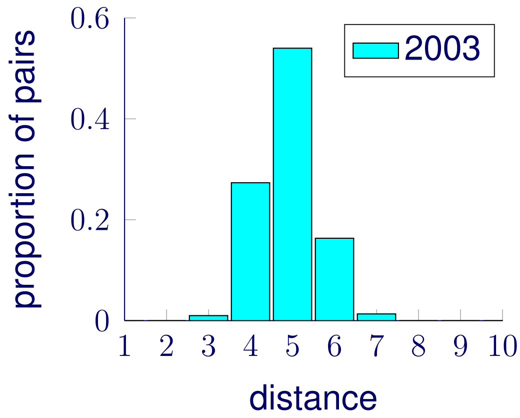




Loglog plot degree sequences WWW in-degree and Internet

- ▷ Straight line: proportion $p_k$ of vertices of degree $k$ satisfies $p_k = ck^{-\tau}$.
- ▷ Empirical evidence: Often $\tau \in (2, 3)$ reported.

# Small-world paradigm




Distances in Strongly Connected Component WWW and IMDb.

# Network science

▷ Complex networks modeled using

random graphs.

▷ Network functionality modeled by stochastic processes on them.

▷ A plethora of examples:

| | |
|---|---|
| Disease spread | Synchronization |
| Information diffusion | Robustness to failures |
| Consensus reaching | Information retrieval |
| Percolation | Random walks... |

▷ Also algorithms on networks important: PageRank, assortativity, community detection,...

▷ Prominent part of applied math for decades to come.

# Models complex networks

▷ Inhomogeneous Random Graphs:
Static random graph, independent edges with inhomogeneous edge occupation probabilities, yielding scale-free graphs.
(Chapters I.6, II.2 and II.5)

[Extensions of Erdős-Rényi random graphs Chapters I.4 and I.5.]

▷ Configuration Model:
Static random graph with prescribed degree sequence.
(Chapters I.7, II.3 and II.6)

▷ Preferential Attachment Model:
Dynamic model, attachment proportional to degree plus constant.
(Chapters I.8, II.4 and II.7)

Universality??

# Erdős-Rényi

Erdős-Rényi random graph is random subgraph of complete graph on $[n] := \{1, 2, \ldots, n\}$ where each of $\binom{n}{2}$ edges is occupied independently with prob. $p$.

Simplest imaginable model of a random graph.

$\triangleright$ Attracted tremendous attention since introduction 1959, mainly in combinatorics community:

Probabilistic method (Spencer, Erdős et al.).

$\triangleright$ Average degree equals $(n-1)p \approx np$, so choose $p = \lambda/n$ to have sparse graph.

$\triangleright$ Egalitarian: Every vertex has equal connection probabilities. Misses hub-like structure of real networks.

# Inhomogeneous random graphs

▷ Extensions of Erdős-Rényi random graph with different vertices.

▷ Chung-Lu: random graphs with prescribed expected degrees:
⋆ Connected component structure (2002)
⋆ Distance results (2002), PNAS
⋆ Book (2006)

▷ Most general:
⋆ Bollobas, Janson and Riordan (2007)
⋆ Söderberg (2007): Phys. Rev. E

We focus on

generalized random graph.

# Generalized random graph

$\triangleright$ Attach edge with probability $p_{ij}$ between vertices $i$ and $j$, where

$$p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j}, \qquad \text{with} \qquad \ell_n = \sum_{i \in [n]} w_i,$$

different edges being independent [Britton-Deijfen-Martin-Löf 05]

$\triangleright$ Resulting graph is denoted by $\mathrm{GRG}_n(\boldsymbol{w})$.

| |
|---|
| Interpretation: $w_i$ is close to expected degree vertex $i$. |

$\star$ Retrieve Erdős-Rényi RG with $p = \lambda/n$ when $w_i = n\lambda/(n-\lambda)$.

$\triangleright$ Related models:

$\star$ Chung-Lu model: $p_{ij} = w_i w_j/\ell_n \wedge 1$;

$\star$ Norros-Reittu model: $p_{ij} = 1 - \mathrm{e}^{-w_i w_j/\ell_n}$.

$\star$ Janson (2010): General conditions for asymptotic equivalence.

# Regularity vertex weights

**Condition I.6.4.** Denote empirical distribution function weight by

$$F_n(x) = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{w_i \leq x\}}, \qquad x \geq 0.$$

(a) Weak convergence of vertex weight. There exists $F$ s.t.

$$W_n \xrightarrow{d} W,$$

where $W_n$ and $W$ have distribution functions $F_n$ and $F$.
(b) Convergence of average vertex weight.

$$\lim_{n \to \infty} \mathbb{E}[W_n] = \mathbb{E}[W] > 0.$$

(c) Convergence of second moment vertex weight.

$$\lim_{n \to \infty} \mathbb{E}[W_n^2] = \mathbb{E}[W^2].$$

# Canonical choice weights

Aim: Proportion of vertices $i$ with $d_i = k$ is close to

$$p_k = \mathbb{P}(D = k),$$

for some random variable $D$.

(A) Take $\boldsymbol{w} = (w_1, \ldots, w_n)$ as i.i.d. random variables with distribution function $F$.

(B) Take $\boldsymbol{w} = (w_1, \ldots, w_n)$ as

$$w_i = [1 - F]^{-1}(i/n).$$

Interpretation: Proportion of vertices $i$ with $w_i \leq x$ is close to $F(x)$.

▷ Power-law example: $F(x) = [1 - (a/x)^{\tau-1}]\mathbb{1}_{\{x \geq a\}}$, for which

$$[1 - F]^{-1}(u) = a(1/u)^{-1/(\tau-1)}, \qquad \text{so that} \qquad w_j = a(n/j)^{1/(\tau-1)}.$$

# Degree structure GRG

Denote proportion of vertices with degree $k$ by

$$P_k^{(n)} = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{D_i = k\}},$$

where $D_i$ is degree of $i \in [n]$. Then [Bollobás-Janson-Riordan (07)]

$$P_k^{(n)} \xrightarrow{\mathbb{P}} p_k = \mathbb{E}\Big[\mathrm{e}^{-W} \frac{W^k}{k!}\Big],$$

where $W$ is a random variable having distribution function $F$. †

Recognize limit $(p_k)_{k \geq 0}$ as probability mass function of Poisson random variable with random parameter $W \sim F$. In particular,

$$\sum_{l \geq k} p_l \sim c k^{-(\tau-1)} \quad \text{iff} \quad \mathbb{P}(W \geq k) \sim c k^{-(\tau-1)}.$$

# Configuration model

$\triangleright$ Invented by Bollobás (80) EJC

to study number of graphs with given degree sequence.

Inspired by Bender+Canfield (78) JCT(A)

Giant component: Molloy, Reed (95)

Popularized by Newman-Strogatz-Watts (01)

$\triangleright$ In configuration model $\mathrm{CM}_n(\boldsymbol{d})$ degree sequence is prescribed:

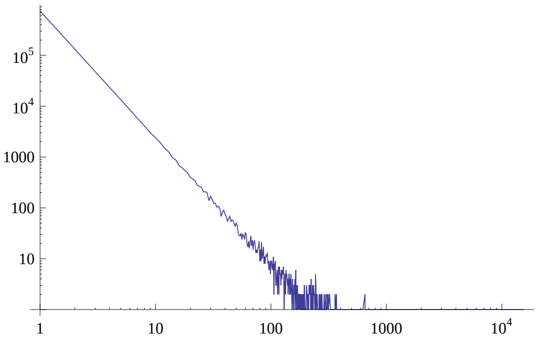
> $\triangleright$ $n$ number of vertices;
> $\triangleright$ $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ sequence of degrees is given.

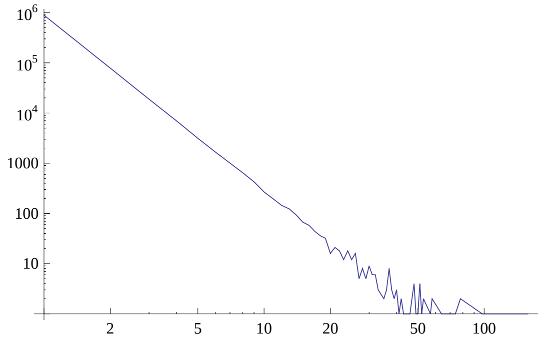
Often $(d_i)_{i\in[n]}$ taken to be i.i.d.

$\triangleright$ Special attention to power-law degrees, i.e., for $\tau > 1$ and $c_\tau$

$$\mathbb{P}(d_1 \geq k) = c_\tau k^{-\tau+1}(1 + o(1)).$$

# Power laws CM




Loglog plot of degree sequence CM with i.i.d. degrees $n = 1,000,000$ and $\tau = 2.5$ and $\tau = 3.5$, respectively.

# Graph construction CM

▷ Assign $d_j$ half-edges to vertex $j$. Assume total degree

$$\ell_n = \sum_{i \in [n]} d_i$$

is even.

▷ Pair half-edges to create edges as follows:
Number half-edges from $1$ to $\ell_n$ in any order.
First connect first half-edge at random with one of other $\ell_n - 1$ half-edges.

▷ Continue with second half-edge (when not connected to first) and so on, until all half-edges are connected.

▷ Resulting graph is denoted by $\mathrm{CM}_n(\boldsymbol{d})$.

# Regularity vertex degrees

**Condition I.7.8.** Denote empirical distribution function degrees by

$$F_n(x) = \frac{1}{n} \sum_{i\in[n]} \mathbb{1}_{\{d_i \leq x\}}, \qquad x \geq 0.$$

(a) Weak convergence of vertex degrees. There exists $F$ s.t.

$$D_n \xrightarrow{d} D,$$

where $D_n$ and $D$ have distribution functions $F_n$ and $F$.
(b) Convergence of average vertex weight.

$$\lim_{n\to\infty} \mathbb{E}[D_n] = \mathbb{E}[D] > 0.$$

(c) Convergence of second moment vertex degrees.

$$\lim_{n\to\infty} \mathbb{E}[D_n^2] = \mathbb{E}[D^2] < \infty.$$

# Canonical choice degrees

Aim: Proportion of vertices $i$ with $d_i = k$ is close to

$$F(k) - F(k-1) = p_k = \mathbb{P}(D = k),$$

where $D$ has distribution function $F$.

$\star$ Power-law degrees: precise structure of large degrees crucial.

(A) Take $\boldsymbol{d} = (d_1, \ldots, d_n)$ as i.i.d. rvs with distribution function $F$.

Double randomness!

(B) Take $\boldsymbol{d} = (d_1, \ldots, d_n)$ such that $d_i = [1 - F]^{-1}(i/n)$, with $F$ distribution function on $\mathbb{N}$.

**Power-law degrees:**

$$[1 - F](k) \approx ck^{-(\tau-1)}, \quad \text{so that} \quad d_j \approx a(n/j)^{1/(\tau-1)}.$$

# Simple CMs

**Proposition I.7.7.** Let $G = (x_{ij})_{i,j\in[n]}$ be multigraph on $[n]$ s.t.

$$d_i = x_{ii} + \sum_{j\in[n]} x_{ij}.$$

Then, with $\ell_n = \sum_{v\in[n]} d_v$,

$$\mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) = G) = \frac{1}{(\ell_n - 1)!!} \frac{\prod_{i\in[n]} d_i!}{\prod_{i\in[n]} 2^{x_{ii}} \prod_{1\leq i\leq j\leq n} x_{ij}!}.$$

Consequently, number of simple graphs with degrees $\boldsymbol{d}$ equals

$$N_n(\boldsymbol{d}) = \frac{(\ell_n - 1)!!}{\prod_{i\in[n]} d_i!} \mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \text{ simple}),$$

and, conditionally on $\mathrm{CM}_n(\boldsymbol{d})$ simple,

$\mathrm{CM}_n(\boldsymbol{d})$ is uniform random graph with degrees $\boldsymbol{d}$.

# Relation GRG and CM

**Theorem I.6.15.** The $\mathrm{GRG}_n(\boldsymbol{w})$ with edge probabilities $(p_{ij})_{1\leq i<j\leq n}$ given by

$$p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j},$$

conditioned on its degrees $\{d_i(X) = d_i \forall i \in [n]\}$ is uniform over all graphs with degree sequence $(d_i)_{i\in[n]}$.

Consequently, conditionally on degrees, $\mathrm{GRG}_n(\boldsymbol{w})$ has the same distribution as $\mathrm{CM}_n(\boldsymbol{d})$ conditioned on simplicity.

Allows properties of $\mathrm{GRG}_n(\boldsymbol{w})$ to be proved through $\mathrm{CM}_n(\boldsymbol{d})$ by showing that degrees $\mathrm{GRG}_n(\boldsymbol{w})$ satisfy right asymptotics.

Inspires Degree Regularity Condition.[†]

# Self-loops + multi-edges

▷ CM can have cycles and multiple edges, but these are relatively scarce compared to the number of edges. [Theorem I.7.10 and Prop. I.7.11]

▷ Let $D_n$ denote degree of uniformly chosen vertex. Condition I.7.8(a): $D_n$ converges in distribution to limiting random variable $D$.

▷ When $\mathbb{E}[D_n^2] \to \mathbb{E}[D^2] < \infty$, then numbers of self-loops and multiple edges converge in distribution to two independent Poisson variables with parameters $\nu/2$ and $\nu^2/4$, respectively, where

$$\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}.$$

[Theorem I.7.12, Prop. I.7.13]

▷ Proof: moment method (Bollobás 80, Janson 09) or Chen-Stein method (Angel-Holmgren-vdH 16).

# Preferential attachment model

▷ Albert-Barabási (1999):
Emergence of scaling in random networks (Science).
34013 cit. (12-08-2019).

▷ Bollobás, Riordan, Spencer, Tusnády (2001):
The degree sequence of a scale-free random graph process (RSA)
852 cit. (12-08-2019).

[Yule (1925) and Simon (1955) already introduced similar models.]

| |
|---|
| In preferential attachment models, network is growing in time, in such a way that new vertices are more likely to be connected to vertices that already have high degree. |

**Rich-get-richer model.**

# Preferential attachment model

▷ Albert-Barabási (1999):
Emergence of scaling in random networks (Science).
34013 cit. (12-08-2019).
▷ Bollobás, Riordan, Spencer, Tusnády (2001):
The degree sequence of a scale-free random graph process (RSA)
852 cit. (12-08-2019).

[Yule (1925) and Simon (1955) already introduced similar models.]

In preferential attachment models, network is growing in time, in such a way that new vertices are more likely to be connected to vertices that already have high degree.

**Old-get-richer model.**

# Preferential attachment

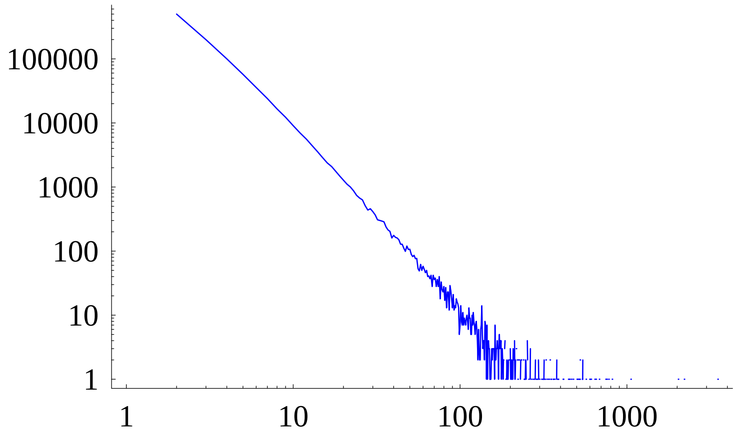
At time $n$, single vertex is added with $m$ edges emanating from it. Probability that edge connects to $i$th vertex is proportional to

$$D_i(n-1)+\delta,$$

where $D_i(n)$ is degree vertex $i$ at time $n$, $\delta > -m$ is parameter.

Yields power-law degree sequence with exponent $\tau = 3 + \delta/m > 2$.

Bol-Rio-Spe-Tus 01 $\delta = 0$, DvdEvdHH09,...




$m = 2, \delta = 0, \tau = 3, n = 10^6$

# Degrees in PAM

Bollobás-Riordan-Spencer-Tusnády 01: First to give proof for $\delta = 0$. Tons of subsequent proofs, many of which follow same key steps:

▷ A clever Doob martingale:

$$M_n = \mathbb{E}[N_k(t) \mid \mathrm{PA}_n],$$

where $N_k(t)$ is number of vertices of degree $k$ at time $t$, combined with Azuma-Hoeffding to prove concentration. See Section I.8.4 for details.

▷ Analysis of means: Identify asymptotics $\mathbb{E}[N_k(t)]$ and prove that

$$\frac{\mathbb{E}[N_k(t)]}{t} \to p_k.$$

Many different ways to do this. See Section I.8.5 for details.

# Albert-László Barabási

"...the scale-free topology is evidence of organizing principles acting at each stage of the network formation. (...) No matter how large and complex a network becomes, as long as preferential attachment and growth are present it will maintain its hub-dominated scale-free topology."

# Conclusion networks

Many real-world networks share important features:

scale-free and small-world paradigms.

Often, suggestion of infinite-variance degrees.

Models invented to describe properties:

Configuration model and generalized random graph.

Models are flexible in their degree structure.

# Lecture 2:

# Local weak convergence: theory

# Network models I

▷ **Configuration model with clustering:**

Input per vertex $i$ is number of simple edges, number of triangles, number of squares, etc. Then connect uniformly at random.

Result: Random graph with (roughly) specified degree, triangle, square, etc distribution over graph.

Application: Social networks?

▷ **Small-world model:**

Start with $d$-dimensional torus (=circle $d = 1,$ donut $d = 2,$ etc). Put in nearest-neighbor edges. Add few edges between uniform vertices, either by rewiring or by simply adding.

Result: Spatial random graph with high clustering, but degree distribution with thin tails.

Application: None? Often used by neuroscientists.

# Network models II

▷ **Random intersection graph:**
Specify collection of groups. Vertices choose group memberships. Put edge between any pairs of vertices in same group.

Result: Flexible collection of random graphs, with high clustering, communities by groups, tunable degree distribution.

Application: Collaboration graphs?

▷ **Spatial preferential attachment model:**
First give vertex uniform location. Let it connect to close by vertices with probability proportionally to degree.

Result: Spatial random graph with scale-free degrees and high clustering.

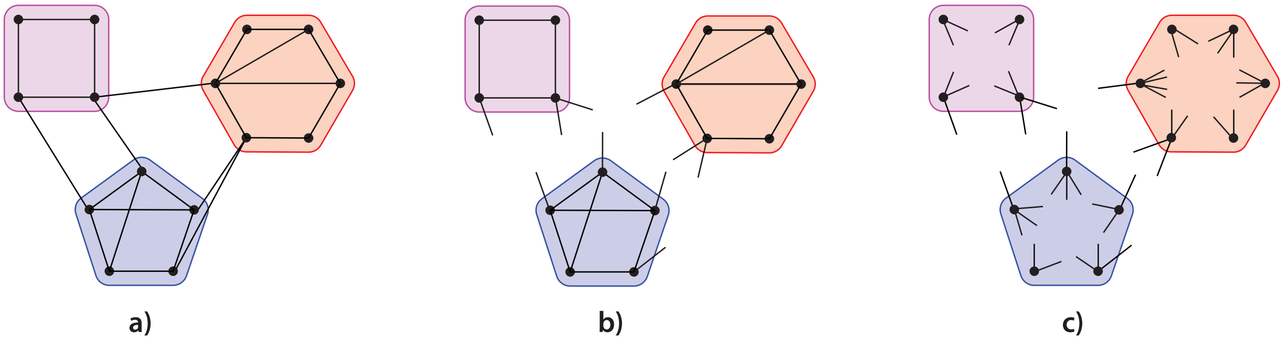
Application: Social networks, WWW?

# Hierarchical CM

Vertex $i$ is blown up to represent small community graph.
Connect inter-community half-edges uniformly at random.

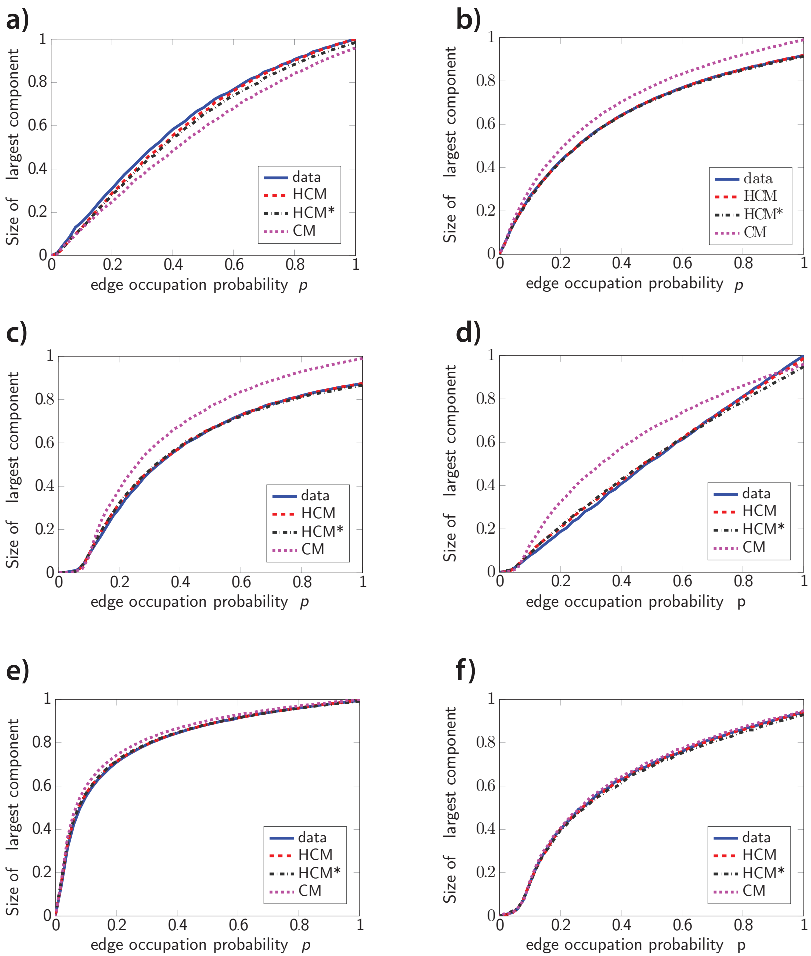
Result: Random graph with (roughly) specified communities.

Application: Many real-world networks on mesoscopic scale.
Stegehuis+vdH+vL16 Scientific Reports, Phys. Rev. E.




Network HCM HCM$^*$

# Percolation on HCM



# Local weak convergence

$\triangleright$ Key technique in analyzing sparse graphs is

local weak convergence.

Makes statement that local neighborhoods in CM are like BP exact. See Chapter II.2 for intro LWC and Section II.4.2 for LWC CM.$^\dagger$

$\triangleright$ Applies much more generally:
- General IRG: Section II.3.2.
- PAM: Berger-Borgs-Chayes-Saberi (14) and Section II.5.2.

$\triangleright$ LWC holds when

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{B_r(i)\simeq H_\star\}} \xrightarrow{\mathbb{P}} \mathbb{P}(B_r(\varnothing)\simeq H_\star),$$

for any rooted graph $H_\star$, where $B_r(i)$ is $r$-neighborhood of $i\in[n]$ and $B_r(\varnothing)$ is $r$-neighborhood of $\varnothing$ in some limiting rooted random graph.

# Overview local weak convergence

Local weak convergence implies that

$\triangleright |\mathcal{C}_{\max}|/n$ is at most $\mathbb{P}(B_r(\varnothing) > 0 \forall r)$ (=one-sided LLN);
$\triangleright$ proportion neighborhoods of specific shape converges;
$\triangleright$ various continuous functionals in local weak convergence topology converge as well.

Examples include log partition function Ising model, PageRank distribution, spectral distribution and through somewhat more work and under more restrictions, densest subgraph.

$\triangleright$ Many global graph parameters, such as proportion vertices in giant component or graph distances do not directly converge, but

LWC gives good starting point analysis.

# Local weak convergence: theory

$\star$ Literature:
Aldous+Steele (2004): Objective Method.
Benjamini-Schramm (2001): Recurrence of random walks.
Lovasz (2012): More combinatorial perspective.

$\triangleright$ Metric on rooted graphs in Section II.2.1.

$\triangleright$ Local weak convergence of deterministic graphs in Section II.2.2.

$\triangleright$ Local weak convergence of random graphs in Section II.2.3.

$\triangleright$ Consequences of local weak convergence in Section II.2.4?

# Lecture 3:

# Local weak convergence of random graphs

# Neighborhoods in CM

$\triangleright$ Important ingredient in proof is description local neighborhood of uniform vertex $U_1 \in [n]$. Its degree has distribution $D_{U_1} \stackrel{d}{=} D$.

$\triangleright$ Take any of $D_{U_1}$ neighbors $a$ of $U_1$. Law of number of forward neighbors of $a$, i.e., $B_a = D_a - 1$, is approximately

$$\mathbb{P}(B_a = k) \approx \frac{(k+1)}{\sum_{i\in[n]} d_i} \sum_{i\in[n]} \mathbb{1}_{\{d_i=k+1\}} \xrightarrow{\mathbb{P}} \frac{(k+1)}{\mathbb{E}[D]} \mathbb{P}(D = k+1).$$

Equals size-biased version of $D$ minus 1. Denote this by $D^\star - 1$.

# Local tree-structure CM

$\triangleright$ Forward neighbors of neighbors of $U_1$ are close to i.i.d. Also forward neighbors of forward neighbors have asymptotically same distribution...

$\triangleright$ Conclusion: Neighborhood looks like branching process with offspring distribution $D^\star - 1$ (except for root, which has offspring $D.$)

$\triangleright$ Tool to make this precise is

local weak convergence.

$\triangleright$ Give proof in Section II.4.2.

# Local weak conv. PAM

$\triangleright$ Pólya urn: Start with $r_0, b_0$ red and blue balls. Draw

red ball w.p. proportional to number of red balls plus $a_r$,
blue ball w.p. proportional to number of blue balls plus $a_b$.

Replace by two balls of same color. Then number of red balls at time $n$ equals

$$R_n \sim r_0 + \mathsf{Bin}(n, U),$$

where $U$ is Beta random variable with parameters $(r_0 + a_r, b_0 + a_b)$.

$\triangleright$ Pólya urns: Can give a Pólya urn description of

ratio degree of vertex $k$ compared to total degree vertices $[k]$.

$\triangleright$ Gives Pólya urn description of PAM at time $n$ that gives precise law in terms of $n$ Beta variables and independent edges.

$\triangleright$ Allows to give local weak limit of PAM in terms of multitype BP with continuous types (Ber-Bor-Cha-Sab 14)

# Lecture 4:

# The giant component in random graphs is almost local

# Phase transition CM

Let $\mathcal{C}_{\max}$ denote largest connected component in $\mathrm{CM}_n(\boldsymbol{d})$.

**Theorem 1.** [Mol-Ree 95, Jan-Luc 07, Theorem II.4.4]. When Conditions I.7.8(a-b) hold,

$$\frac{1}{n}|\mathcal{C}_{\max}| \xrightarrow{\mathbb{P}} \zeta,$$

where $\zeta > 0$ precisely when $\nu > 1$ with $\nu = \mathbb{E}[D(D-1)]/\mathbb{E}[D]$.

$\triangleright$ Note: $\zeta > 0$ always true when $\nu = \infty$ : **Robustness!**

$\triangleright$ $d_{\min} = \min_{i\in[n]} d_i \geq 3 : \mathrm{CM}_n(\boldsymbol{d})$ with high probability connected.
Wormald (81), Luczak (92).

$\triangleright$ $d_{\min} = \min_{i\in[n]} d_i \geq 2 : n - |\mathcal{C}_{\max}| \xrightarrow{d} X$ for non-trivial $X$.
Luczak (92), Federico-vdH (17).

# Phase transition for GRG

Let $\mathcal{C}_{\max}$ denote largest connected component in $\mathrm{GRG}_n(\boldsymbol{w})$.

**Theorem 2.** [Chu-Lu 03, Bol-Jan-Rio 07]. When Conditions I.6.4(a-b) hold, there exists $\zeta < 1$ such that

$$\frac{1}{n}|\mathcal{C}_{\max}| \xrightarrow{\mathbb{P}} \zeta,$$

where $\zeta > 0$ precisely when $\nu > 1$, where

$$\nu = \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}.$$

$\triangleright$ Note: $\zeta > 0$ always true when $\nu = \infty$ : **Robustness!**

$\triangleright$ Bol-Jan-Rio 07 much more general.

# Giant is almost local

$\triangleright$ Giant is almost local in Section II.2.5, specifically Corollary II.2.19, and Theorems II.2.20 and II.2.22.

Discussion of Erdős-Rényi random graph in Section II.2.5.1.

$\triangleright$ Intuitive explanation how this can be extended to CM.

Lecture 5:

Small-world phenomenon
on random graphs

Joint work with:

▷ Henri van den Esker (TU Delft)
▷ Gerard Hooghiemstra (TU Delft)
▷ Piet Van Mieghem (TU Delft)
▷ Dmitri Znamenski (Eurandom, now Philips Research)
▷ Alessandro Garavaglia (TU/e)
▷ Francesco Caravenna (Biccoca Milano)

# Giant is almost local

$\triangleright$ Giant is almost local in Section II.2.5, specifically Corollary II.2.19, and Theorems II.2.20 and II.2.22.

Discussion of Erdős-Rényi random graph in Section II.2.5.1.

$\triangleright$ Intuitive explanation how this can be extended to CM.

# Phase transition CM

Let $\mathcal{C}_{\max}$ denote largest connected component in $\mathrm{CM}_n(\boldsymbol{d})$.

**Theorem 1.** [Mol-Ree 95, Jan-Luc 07, Theorem II.4.4]. When Conditions I.7.8(a-b) hold,

$$\frac{1}{n}|\mathcal{C}_{\max}| \xrightarrow{\mathbb{P}} \zeta,$$

where $\zeta > 0$ precisely when $\nu > 1$ with $\nu = \mathbb{E}[D(D-1)]/\mathbb{E}[D]$.

$\triangleright$ Note: $\zeta > 0$ always true when $\nu = \infty$ : **Robustness!**

$\triangleright$ $d_{\min} = \min_{i\in[n]} d_i \geq 3 : \mathrm{CM}_n(\boldsymbol{d})$ with high probability connected.
Wormald (81), Luczak (92).

$\triangleright$ $d_{\min} = \min_{i\in[n]} d_i \geq 2 : n - |\mathcal{C}_{\max}| \xrightarrow{d} X$ for non-trivial $X$.
Luczak (92), Federico-vdH (17).

# Connectivity PAM

**Theorem 3.** [Theorem II.4.16] Let $m \geq 2$. Then, there exists a random time $T < \infty$, such that the preferential attachment model is connected for all times after $T$.

$\triangleright$ Not necessarily true when $m = 1$:
Depends on precise PA rule.

$\triangleright$ Analogy: $\mathrm{CM}_n(\boldsymbol{d})$ with high probability connected when $d_{\min} \geq 3$.

# Graph distances CM

$H_n$ is graph distance between uniform pair of vertices in graph.

**Theorem 4.** [vdHHVM05, Theorem II.7.1]. When Conditions I.7.8(a-c) hold and $\nu = \mathbb{E}[D(D-1)]/\mathbb{E}[D] > 1$, conditionally on $H_n < \infty$,

$$\frac{H_n}{\log_\nu n} \xrightarrow{\mathbb{P}} 1.$$

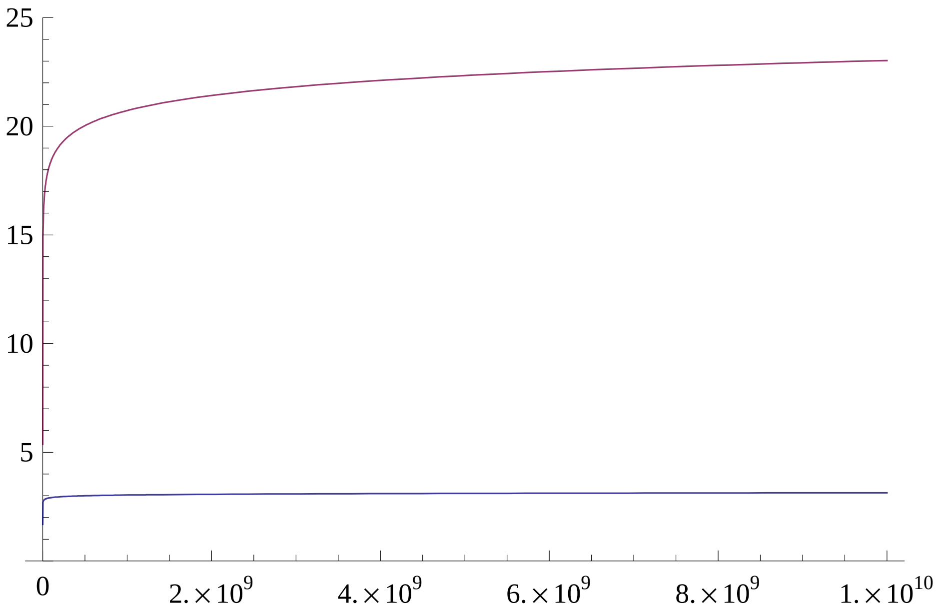
▷ For i.i.d. degrees having at most power-law tails, fluctuations are bounded.

**Theorem 5.** [vdHHZ07, Norros-Reittu 04, Theorem II.7.2]. Let Conditions I.7.8(a-b) hold. When $\tau \in (2,3)$, conditionally on $H_n < \infty$,

$$\frac{H_n}{\log \log n} \xrightarrow{\mathbb{P}} \frac{2}{|\log(\tau-2)|}.$$

▷ vdH-Komjáthy16: For power-law tails, fluctuations are bounded and do not converge in distribution.

# Six degrees of separation revisited




Plot of $x \mapsto \log x$ and $x \mapsto \log \log x$.

# Diameter CM

**Theorem 6.** [Fernholz-Ramachandran 07, Theorem II.7.16]. Under Conditions I.7.8(a-b), there exists $b$ s.t.

$$\frac{\mathrm{diam}(\mathrm{CM}_n(\boldsymbol{d}))}{\log n} \xrightarrow{\mathbb{P}} \frac{1}{\log(\nu)} + 2b.$$

Here $b > 0$ precisely when $\mathbb{P}(D \leq 2) > 0.$

**Theorem 7.** [Caravenna-Garavaglia-vdH 17, Theorem II.7.17]. Under Conditions I.7.8(a-b), when $\tau \in (2,3)$ and $\mathbb{P}(D \geq 3) = 1$ and with $d_{\min} = \min\{d_v \colon v \in [n]\}$,

$$\frac{\mathrm{diam}(\mathrm{CM}_n(\boldsymbol{d}))}{\log\log n} \xrightarrow{\mathbb{P}} \frac{2}{|\log(\tau-2)|} + \frac{2}{\log(d_{\min}-1)}.$$

# Graph distances GRG

**Theorem 8.** [Chung-Lu 03, Bol-Jan-Rio 07, vdEvdHH08, Thm. II.6.2] When Conditions I.6.3(a-c) hold and $\nu = \mathbb{E}[W^2]/\mathbb{E}[W] > 1$, conditionally on $H_n < \infty$,

$$\frac{H_n}{\log_\nu n} \xrightarrow{\mathbb{P}} 1.$$

Under somewhat stronger conditions, fluctuations are bounded.

**Theorem 9.** [Chung-Lu 03, Norros-Reittu 06, Theorem II.6.3]. When $\tau \in (2,3)$, and Conditions I.6.3(a-b) hold, under certain further conditions on $F_n$, and conditionally on $H_n < \infty$,

$$\frac{H_n}{\log\log n} \xrightarrow{\mathbb{P}} \frac{2}{|\log(\tau-2)|}.$$

$\triangleright$ Similar extensions for diameter as for CM (always logarithmic.) Again Bol-Jan-Rio 07 prove Theorem 7 in highly general setting.

# Distances PA models

▷ Results CM and GRG are very alike, with CM having more general behavior (e.g., connectivity). Sign of wished for universality.

> Non-rigorous physics literature predicts that scaling distances in preferential attachment models similar to the one in configuration model with equal
>
> power-law exponent degrees.

▷ General question still wide open, but signs point in this direction.

▷ PAM tends to be much harder to analyze, due to

time dependence.

# Distances PA models

**Theorem 10** [Bol-Rio 04]. For all $m \geq 2$ and $\tau = 3$,

$$\mathrm{diam}(\mathrm{PA}_{m,0}(n)) = \frac{\log n}{\log\log n}(1 + o_{\mathbb{P}}(1)), \qquad H_n = \frac{\log n}{\log\log n}(1 + o_{\mathbb{P}}(1)).$$

**Theorem 11** [Dommers-vdH-Hoo 10]. For all $m \geq 2$ and $\tau \in (3, \infty)$,

$$\mathrm{diam}(\mathrm{PA}_{m,\delta}(n)) = \Theta(\log n), \qquad H_n = \Theta(\log n).$$

**Theorem 12** [Dommers-vdH-Hoo 10, Der-Mon-Mor 12, Car-Gar-vdH17]. For all $m \geq 2$ and $\tau \in (2, 3)$,

$$\frac{H_n}{\log\log n} \xrightarrow{\mathbb{P}} \frac{4}{|\log(\tau - 2)|}, \qquad \frac{\mathrm{diam}(\mathrm{PA}_{m,\delta}(n))}{\log\log n} \xrightarrow{\mathbb{P}} \frac{4}{|\log(\tau - 2)|} + \frac{2}{\log m}.$$

# Structure local limit CM

$\triangleright$ $\mathbb{E}[D^2] < \infty$ : Finite-mean BP, which has exponential growth of generation sizes:

$$\nu^{-k} Z_k \xrightarrow{a.s.} M \in (0, \infty),$$

on event of survival.

$\star$ Explains why distances random graph grow logarithmically.

$\triangleright$ $\tau \in (2, 3)$ : Infinite-mean BP, which has double exponential growth of generation sizes:

$$(\tau - 2)^k \log(Z_k \vee 1) \xrightarrow{a.s.} Y \in (0, \infty),$$

on event of survival.

$\star$ Explains why distances grow doubly logarithmically.

$\triangleright$ Indication of proof...$^\dagger$

# Structure local limit CM

$\triangleright$ Logarithmic upper bound on graph distances CM in Theorem II.7.5.

$\triangleright$ Branching processes with infinite mean in Theorem II.7.14.

$\triangleright$ Diameter of core in CM in Theorem II.7.9.

# Conclusion small-worlds

Many real-world networks share important features:

scale-free and small-world paradigms.

Often, suggestion of infinite-variance degrees.

Models invented to model/explain properties:

Configuration model, generalized random graph and preferential attachment.

Distances are remarkably similar across models.

# Literature CM


[1] Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combin.* **1**(4): 311–316, (1980).

[2] van der Hofstad, Hooghiemstra and Van Mieghem. Distances in random graphs with finite variance degrees. *RSA* **27**(1): 1042–9832, (2005).

[3] Fernholz and Ramachandran. The diameter of sparse random graphs. *RSA* **31**(4): 482–516, (2007).

[4] van der Hofstad, Hooghiemstra and Znamenskyi. Distances in random graphs with finite mean and infinite variance degrees. *Elec. J. Prob.* **12**(25): 703–766, (2007).

[5] Molloy and Reed. A critical point for random graphs with a given degree sequence. *RSA* **6**: 161–179, (1995).

[6] Newman, Strogatz and Watts. Random graphs with arbitrary degree distribution and their application. *Phys. Rev. E* **64**: 026118, (2000).

# Literature IRG


[1] Bollobás, Janson and Riordan. The phase transition in inhomogeneous random graphs. *RSA* **31**(1): 3–122, (2007).

[2] Britton, Deijfen and Martin-Löf. Generating simple random graphs with prescribed degree distribution. *JSP* **124**(6): 1377–1397, (2006).

[3] Chung and Lu. The average distance in a random graph with given expected degrees. *Internet Math.* **1**(1): 91–113, (2003).

[4] Chung and Lu. Complex graphs and networks. CBMS Regional Conference Series in Mathematics. (2006).

[5] van den Esker, van der Hofstad and Hooghiemstra. Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133**(1): 169–202, (2006).

[6] Norros and Reittu. On a conditionally Poissonian graph process. *Adv. Appl. Probab.* **38**(1): 59–75, (2006).

# Literature PAM

[1] Barabási and Albert. Emergence of scaling in random networks. *Science* **286**(2): 509–512, (1999).

[2] Bollobás and Riordan. The diameter of a scale-free random graph. *Combinatorica* **1**(2): 215–225, (2004).

[3] Bollobás, Riordan, Spencer and Tusnády. The degree sequence of a scale-free random graph process. *RSA* **18**(3): 279–290, (2001).

[4] Deijfen, van den Esker, van der Hofstad and Hooghiemstra. A preferential attachment model with random initial degrees. *Ark. Mat.* **47**(1): 41–72, (2009).

[5] Dereich, Mönch and Mörters. Typical distances in ultrasmall random networks. *Journ. Stat. Phys.* **139**(1): 72–107, (2010).

[6] Dommers, van der Hofstad and Hooghiemstra. Diameters in preferential attachment graphs. *Adv. Appl. Probab.* **44**(2): 583–601, (2012).