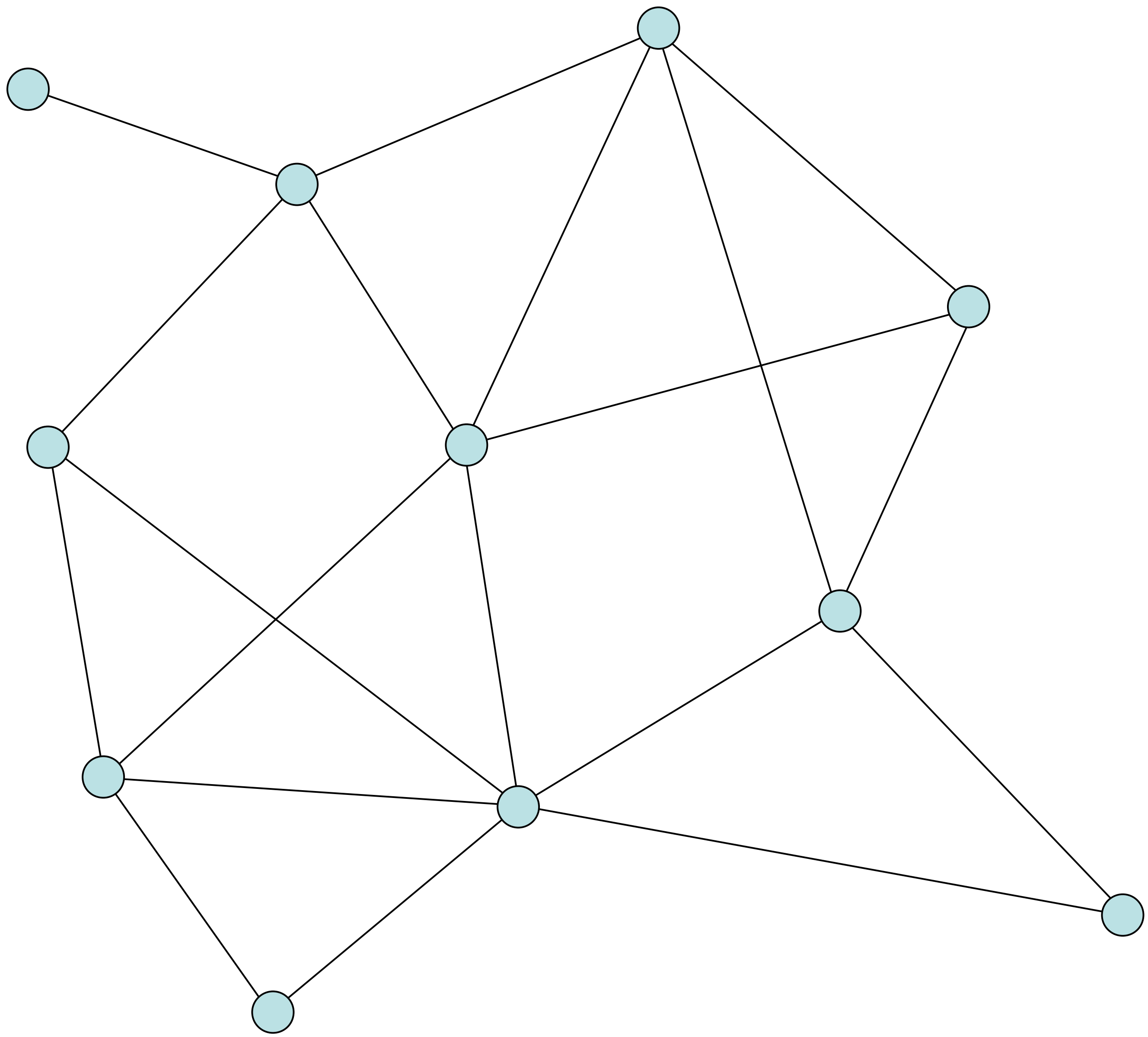
Studying large graphs from left and right

László Lovász

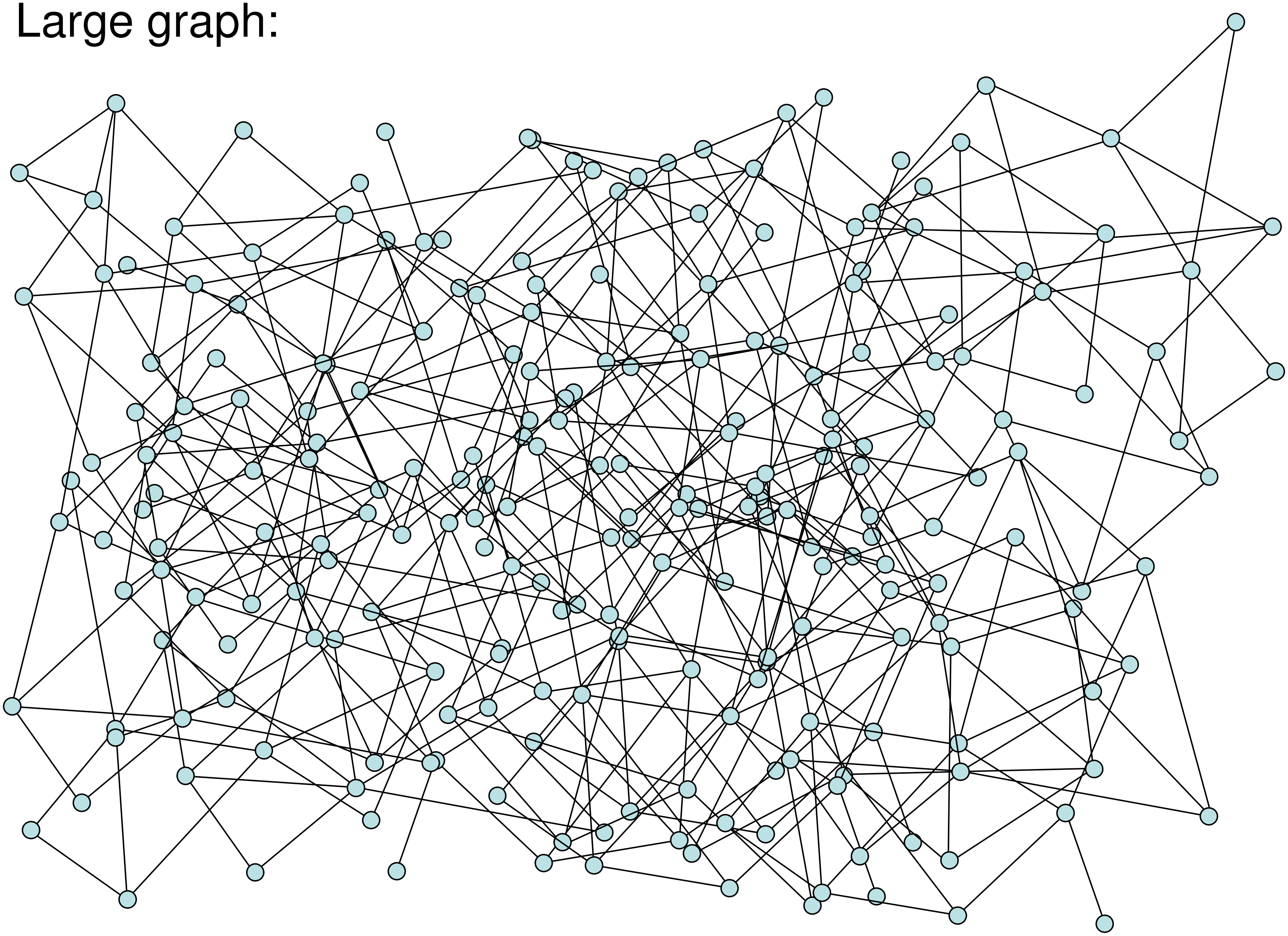
Eötvös Loránd University, Budapest

With: Christian Borgs, Jennifer Chayes, Balázs Szegedy, Vera Sós, Katalin Vesztergombi

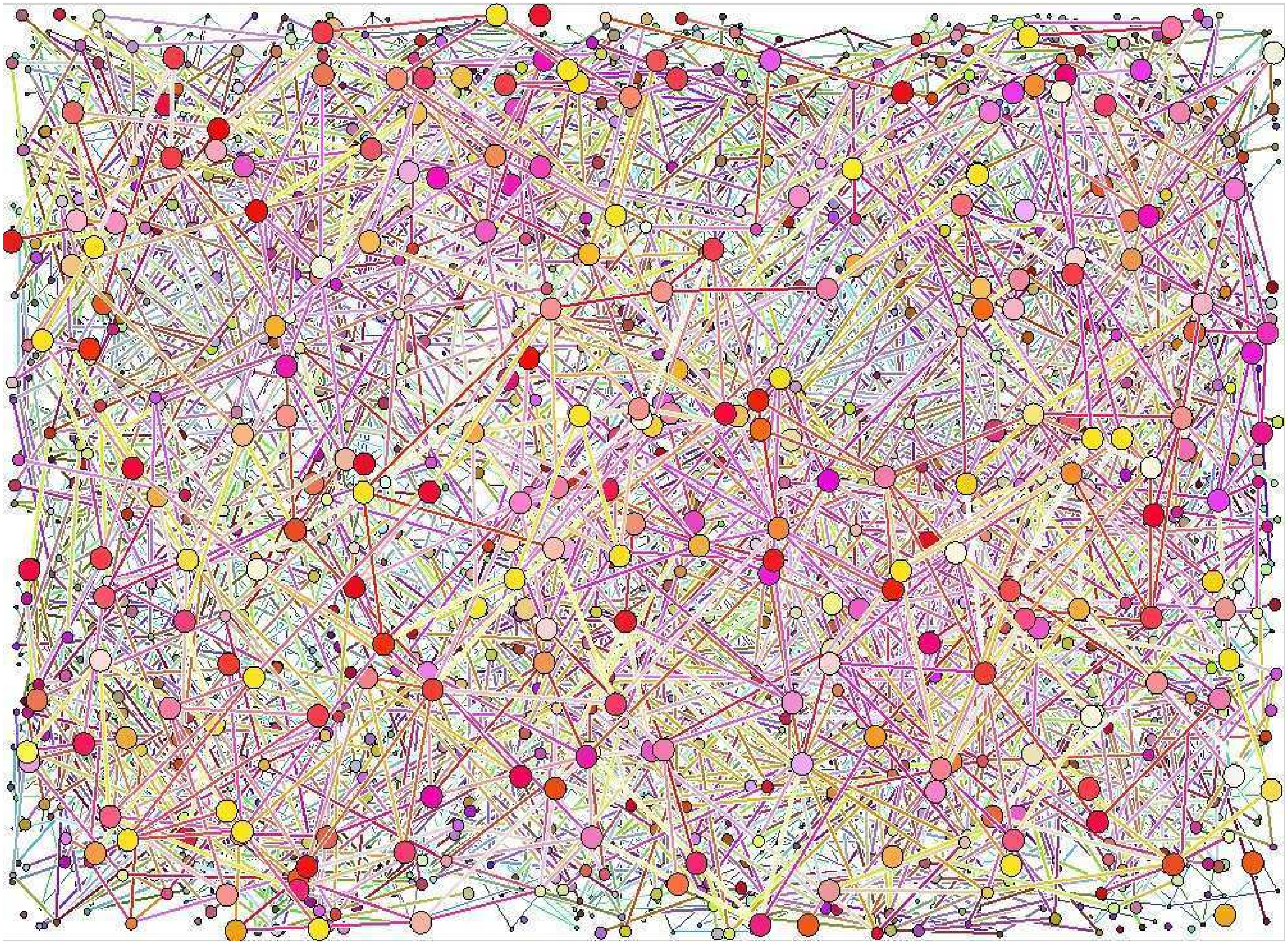
Graph (network):



Large graph:

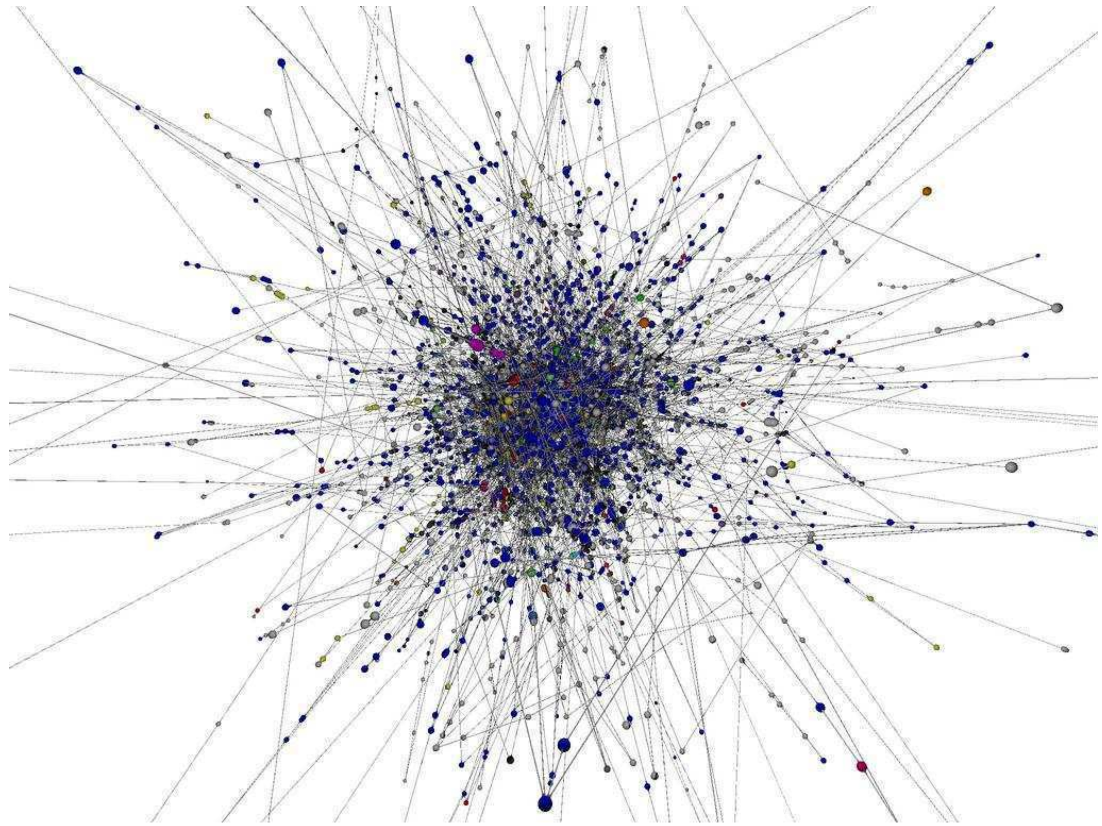


Rather large graph:



Very large graphs:

-Internet



@Stephen Coast

Very large graphs:

- -Internet
- -Social networks
- -Ecological systems
- -VLSI
- -Statistical physics
- -Brain

What properties to study?

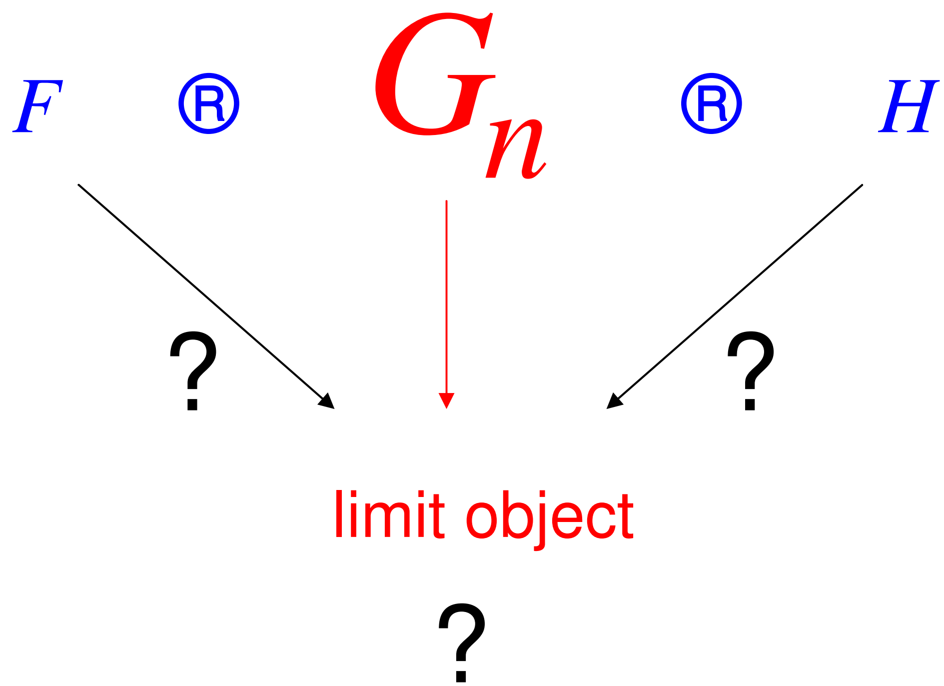
- -Does it have an even number of nodes?
- -How dense is it (average degree)?
- -Is it connected?

F ® G ® H

Very large graph

counting edges,
triangles,
...
spectra,
...

counting colorations,
stable sets,
...
statistical physics,
...
connectivity,
maximum cut,
...



F ® G_n ® H
?
?
limit object
?

When are two graphs "close"?

- Two graphs on the same n nodes that differ in $o(n^2)$ edges are "close".

- Two large random graphs with the same edge-density are "close".

- A random graph with edge-density $1/2$ and a complete bipartite graph are not "close".

Distance of graphs

cut distance

(a) $V(G) = V(G')$

$$d_\square(G,G') = \max_{S,T \subseteq V(G)} \frac{|e_G(S,T) - e_{G'}(S,T)|}{n^2}$$

(b) $|V(G)| = |V(G')| = n$

$$\delta_\square(G,G') = \min_{G \leftrightarrow G'} d_\square(G,G')$$

(c) $|V(G)| = n \neq n' = |V(G')|$

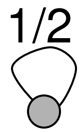
blow up nodes, or fractional overlay

Examples:

$$\delta_\square\left(K_{n,n}, \mathbb{G}(n, \tfrac{1}{2})\right) \approx \frac{1}{8}$$

$$\delta_\square\left(\mathbb{G}_1(n, \tfrac{1}{2}), \mathbb{G}_2(n, \tfrac{1}{2})\right) = o(1)$$

$$\delta_\square\Big(\mathbb{G}_1(n, \tfrac{1}{2}),$$



$$\Big) = \delta_\square\left(\mathbb{G}_1(n, \tfrac{1}{2}), 1/2\right) = o(1)$$

The "weak" Regularity Lemma (Szemerédi, Frieze-Kannan):

For every graph G and $\varepsilon > 0$ there is a graph H with $\le 2^{2/\varepsilon^2}$ nodes such that $\delta_\square(G, H) \le \varepsilon$.

$\hom(G,H) :=$ # of homomorphisms of G into H

Weighted version:

$H = (V, E, \alpha, \beta)$, $\alpha : V \to \mathbb{R}_+$, $\beta : E \to \mathbb{R}_+$

$$\hom(G,H) := \sum_{\varphi: V(G)\to V(H)} \prod_{i\in V(G)} \alpha_{\varphi(i)} \prod_{ij\in E(G)} \beta_{\varphi(i)\varphi(j)}$$

$$\hom^*(G,H) := \sum_{\substack{\varphi: V(G)\to V(H) \\ |\varphi^{-1}(v)|=\alpha_v |V(G)|}} \prod_{ij\in E(G)} \beta_{\varphi(i)\varphi(j)}$$

$$t(F,G)=\frac{\hom(F,G)}{|V(G)|^{|V(F)|}}$$

Probability that random map $V(F)\to V(G)$ is a hom

$$s(G,H)=\frac{\log \hom^*(G,H)}{|V(G)|^2}$$

Which sequences are convergent?

(i) $(G_1, G_2, \ldots)$ convergent: Cauchy in the d_X -metric.

(ii) $(G_1, G_2, \ldots)$ convergent: $"F \quad t(F, G_n)$ is convergent

(iii) $(G_1, G_2, \ldots)$ convergent: $"H \quad t(G_n, H)$ is convergent

(i), (ii) and (iii) are equivalent

Example: random graphs

$$t(F, \mathbb{G}(n, \tfrac{1}{2})) \to \left(\tfrac{1}{2}\right)^{|E(F)|} \qquad \text{with probability 1}$$

$$d_\square\left(\mathbb{G}(n, \tfrac{1}{2}), \mathbb{G}(m, \tfrac{1}{2})\right) \to 0 \qquad (n, m \to \infty)$$

"Counting lemma": $|t(F,G) - t(F,H)| \le |E(F)|\, d_\square(G,H)$

We want to describe completion:
limits of graph sequences

Limits of sequences of graphs with bounded degree:

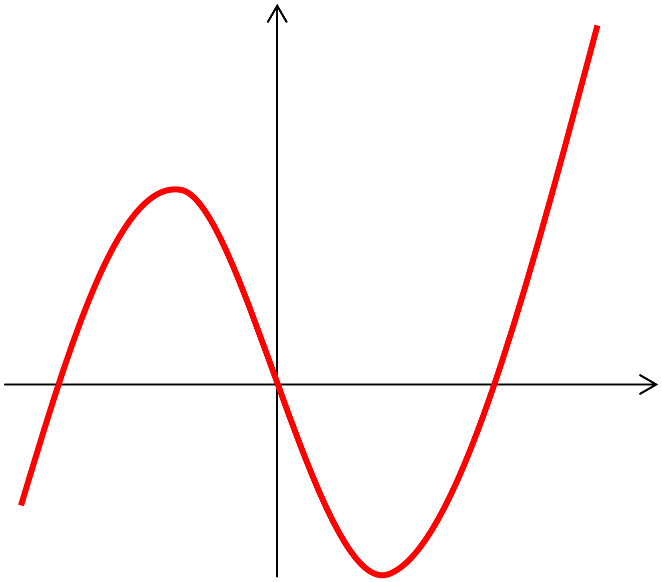
Aldous, Benjamini-Schramm, Lyons, Elek

Limits of sequences of dense graphs:

Borgs, Chayes, L, Sós, B.Szegedy, Vesztergombi

Real numbers

Minimize $x^3 - 6x$ over $x \geq 0$

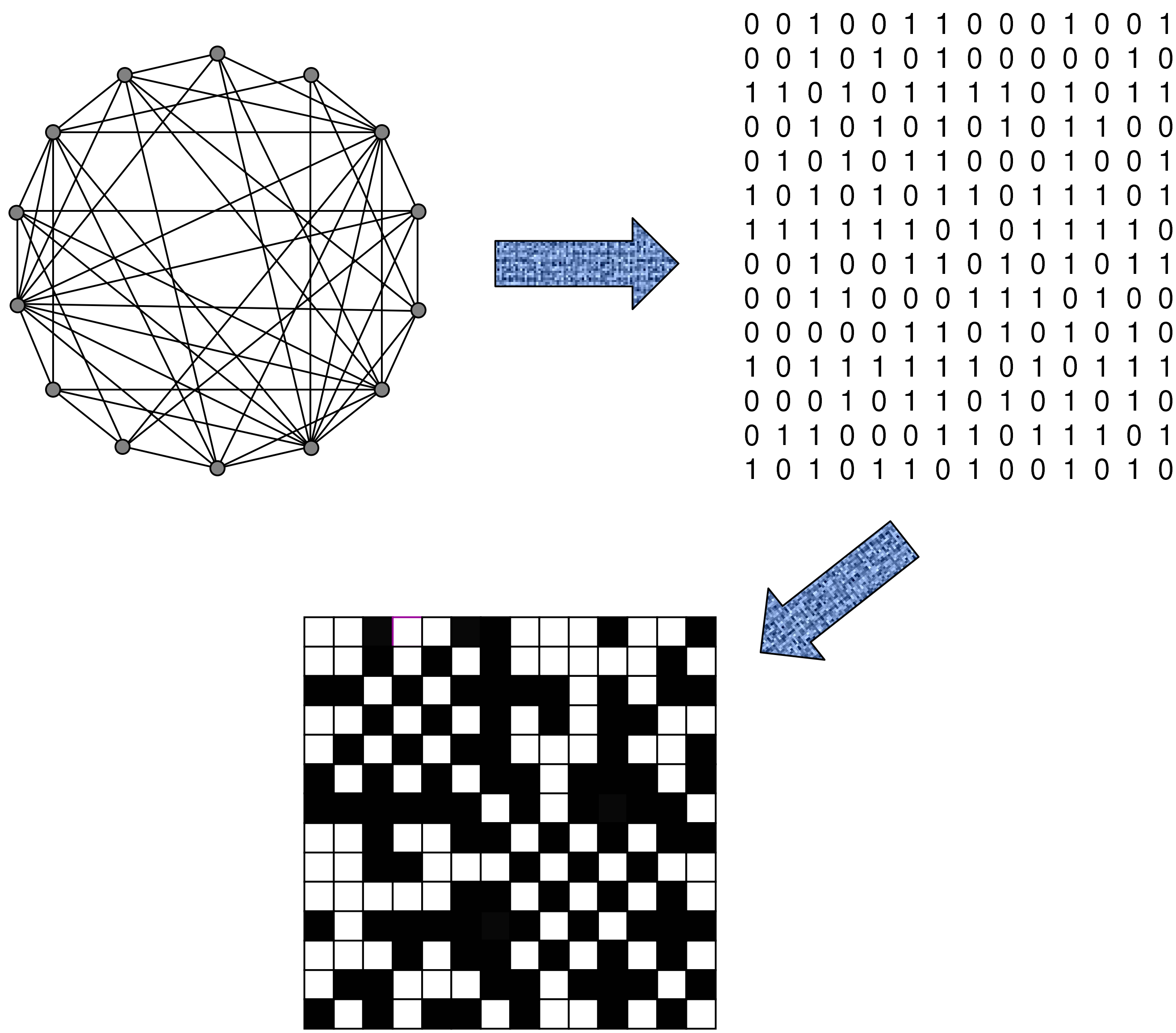


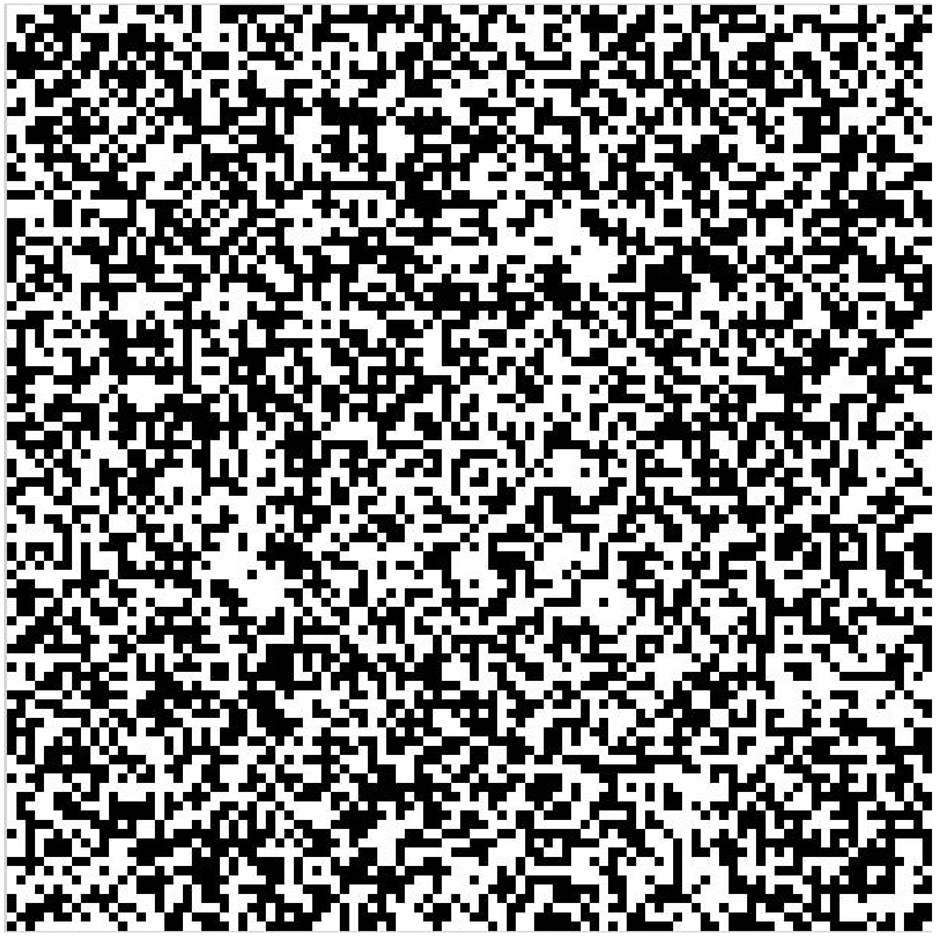
minimum is not attained
in rationals

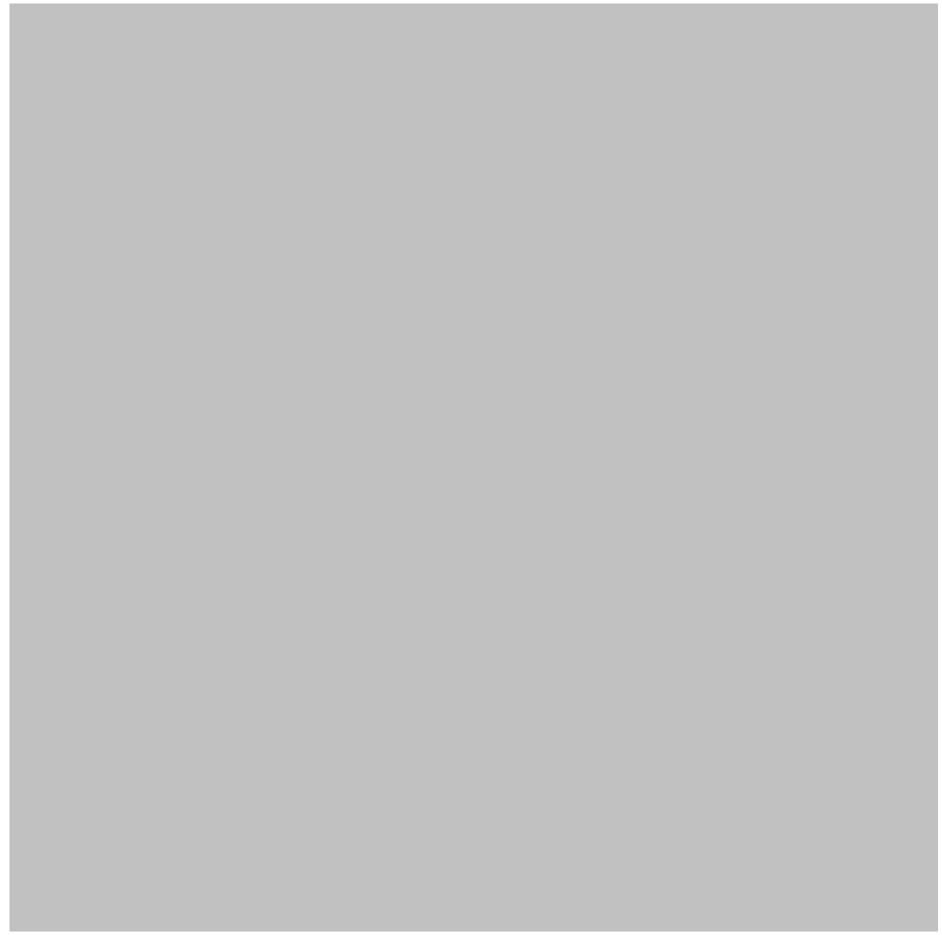
Minimize density of 4-cycles in a graph
with edge-density 1/2

always >1/16,
arbitrarily close for random graphs

minimum is not attained
among graphs


0 0 1 0 0 1 1 0 0 0 1 0 0 1
0 0 1 0 1 0 1 0 0 0 0 0 1 0
1 1 0 1 0 1 1 1 1 0 1 0 1 1
0 0 1 0 1 0 1 0 1 0 1 1 0 0
0 1 0 1 0 1 1 0 0 0 1 0 0 1
1 0 1 0 1 0 1 1 0 1 1 1 0 1
1 1 1 1 1 1 0 1 0 1 1 1 1 0
0 0 1 0 0 1 1 0 1 0 1 0 1 1
0 0 1 1 0 0 0 1 1 1 0 1 0 0
0 0 0 0 0 1 1 0 1 0 1 0 1 0
1 0 1 1 1 1 1 1 0 1 0 1 1 1
0 0 0 1 0 1 1 0 1 0 1 0 1 0
0 1 1 0 0 0 1 1 0 1 1 1 0 1
1 0 1 0 1 1 0 1 0 0 1 0 1 0

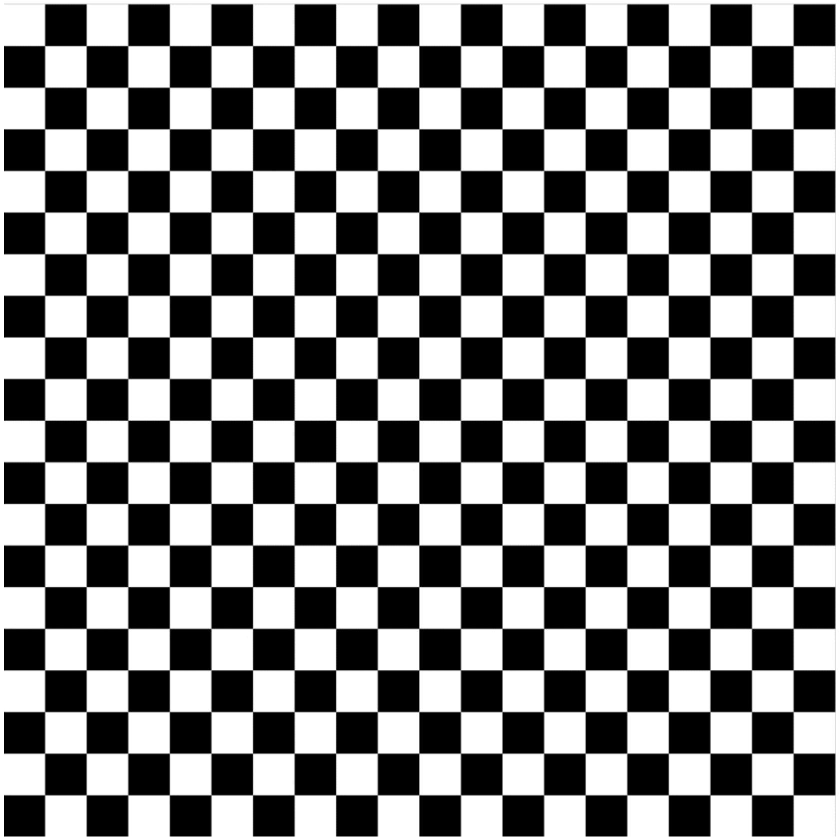


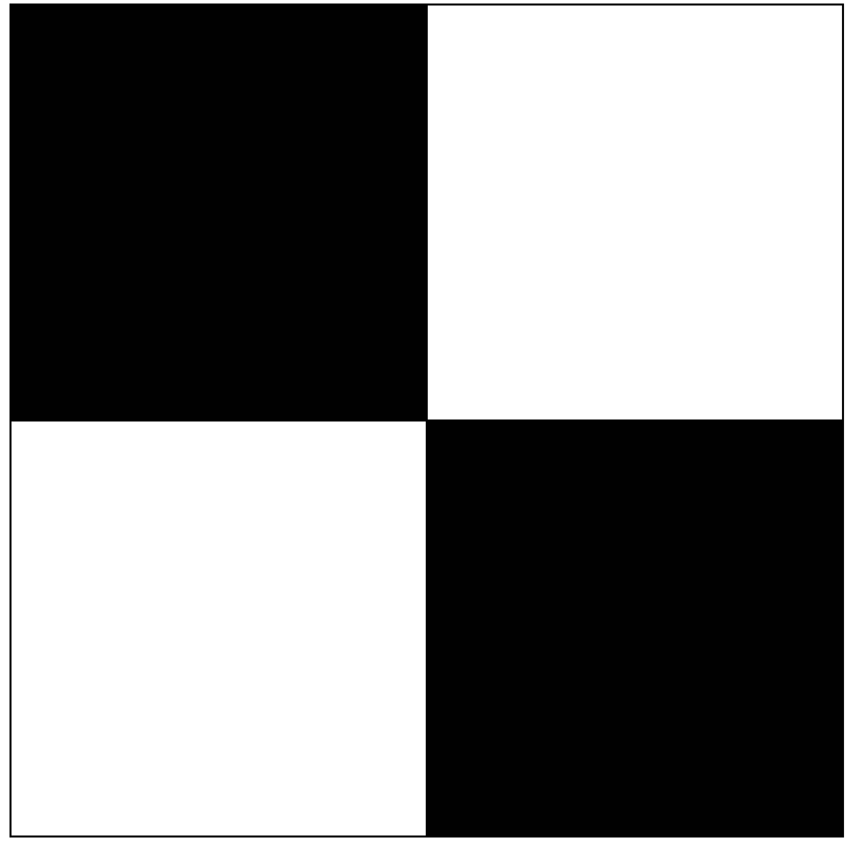


A random graph

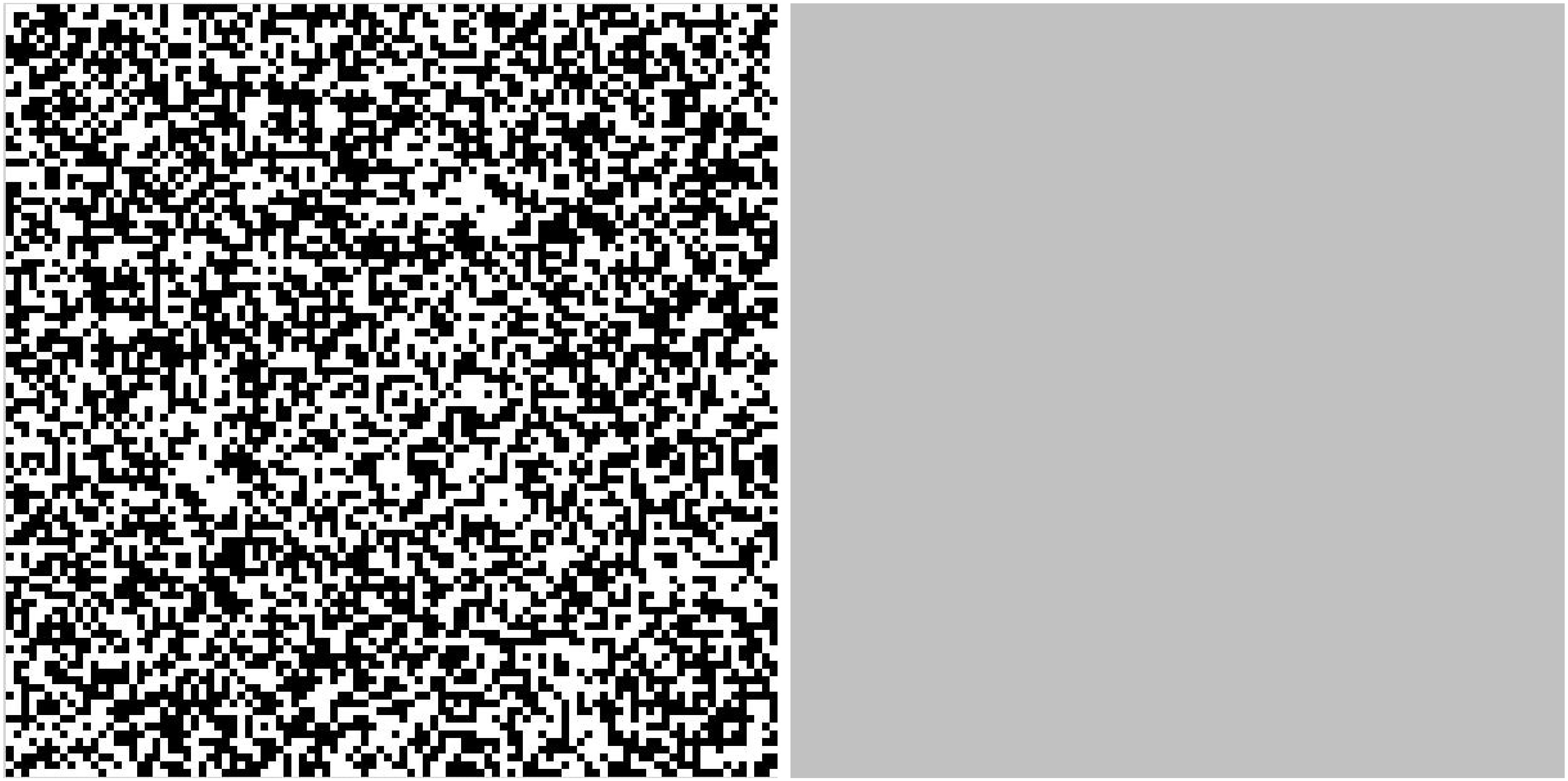
with 100 nodes and with 2500 edges

1/2





Rearranging the rows and columns



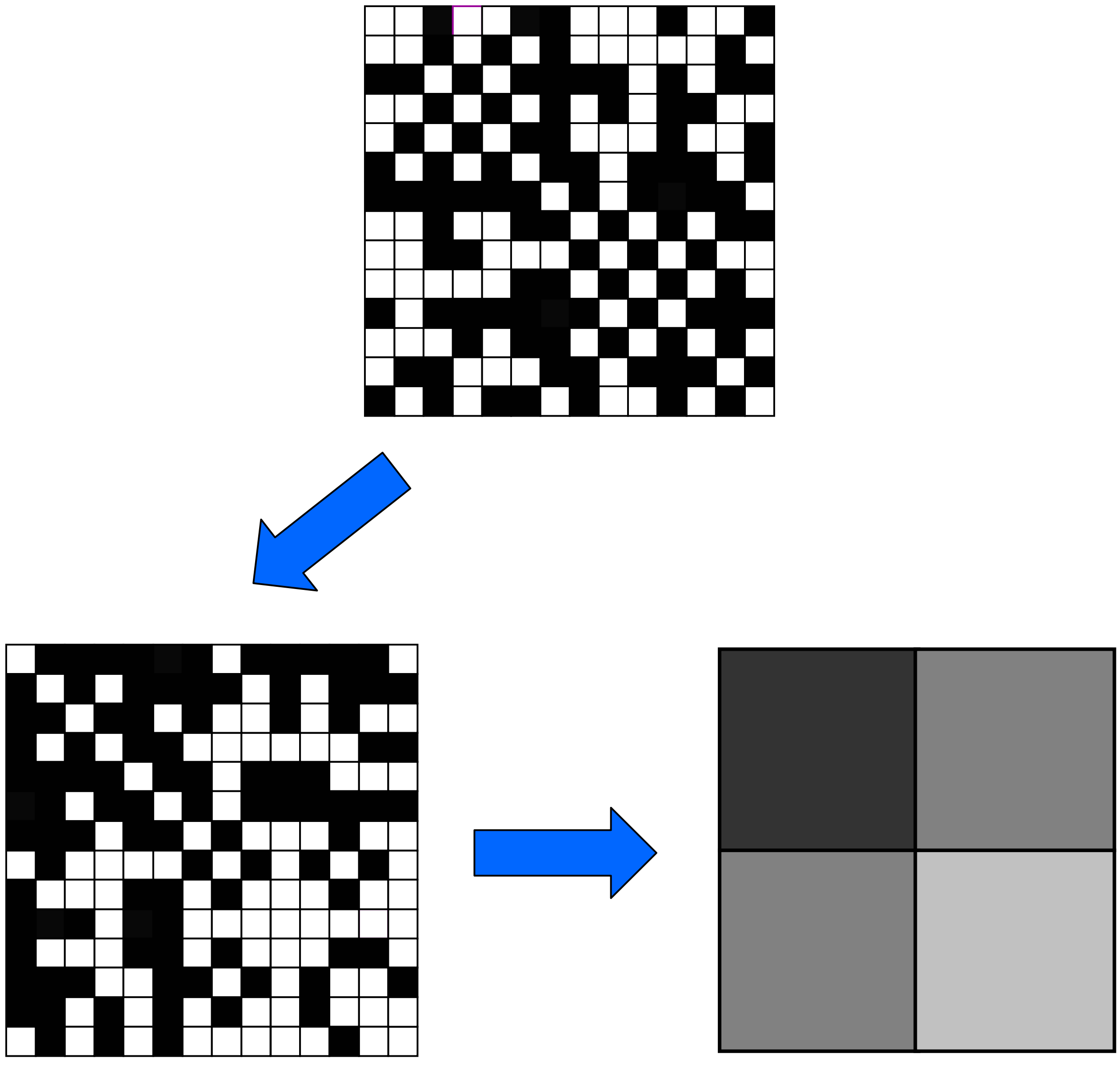
A random graph

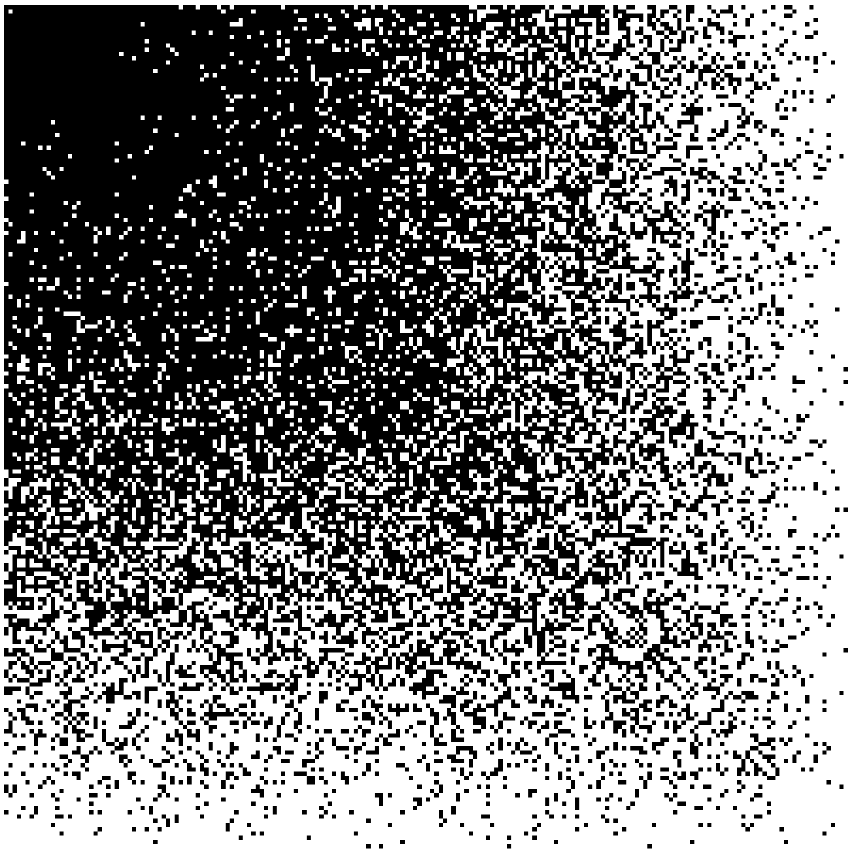
with 100 nodes and with 2500 edges

1/2

(no matter how you reorder the nodes)

Szemerédi's Regularity Lemma

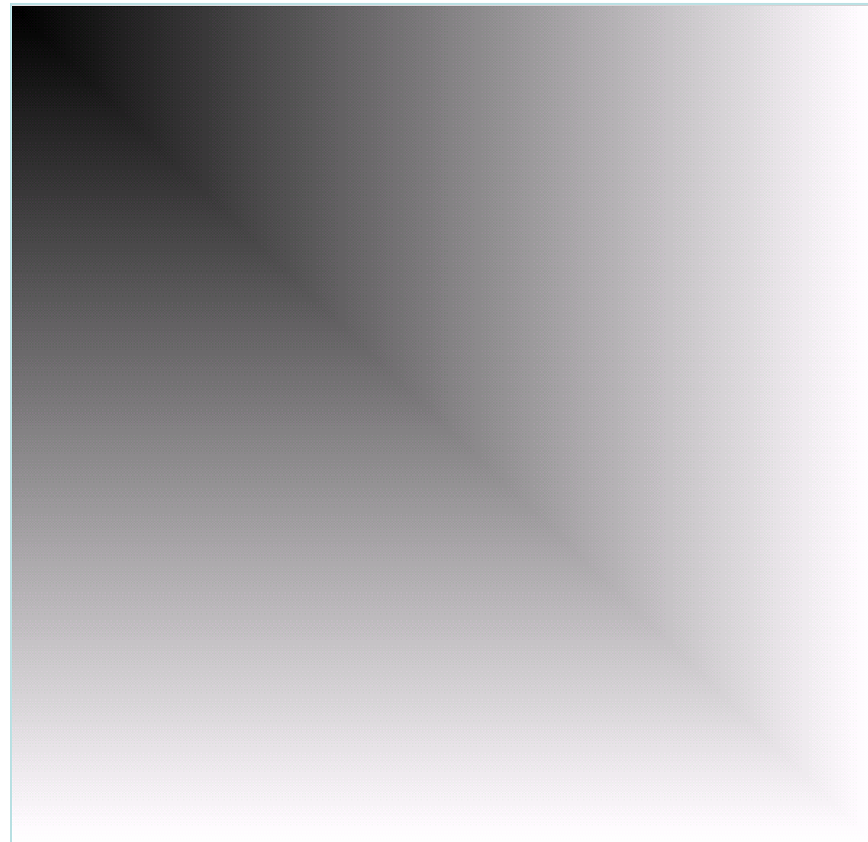




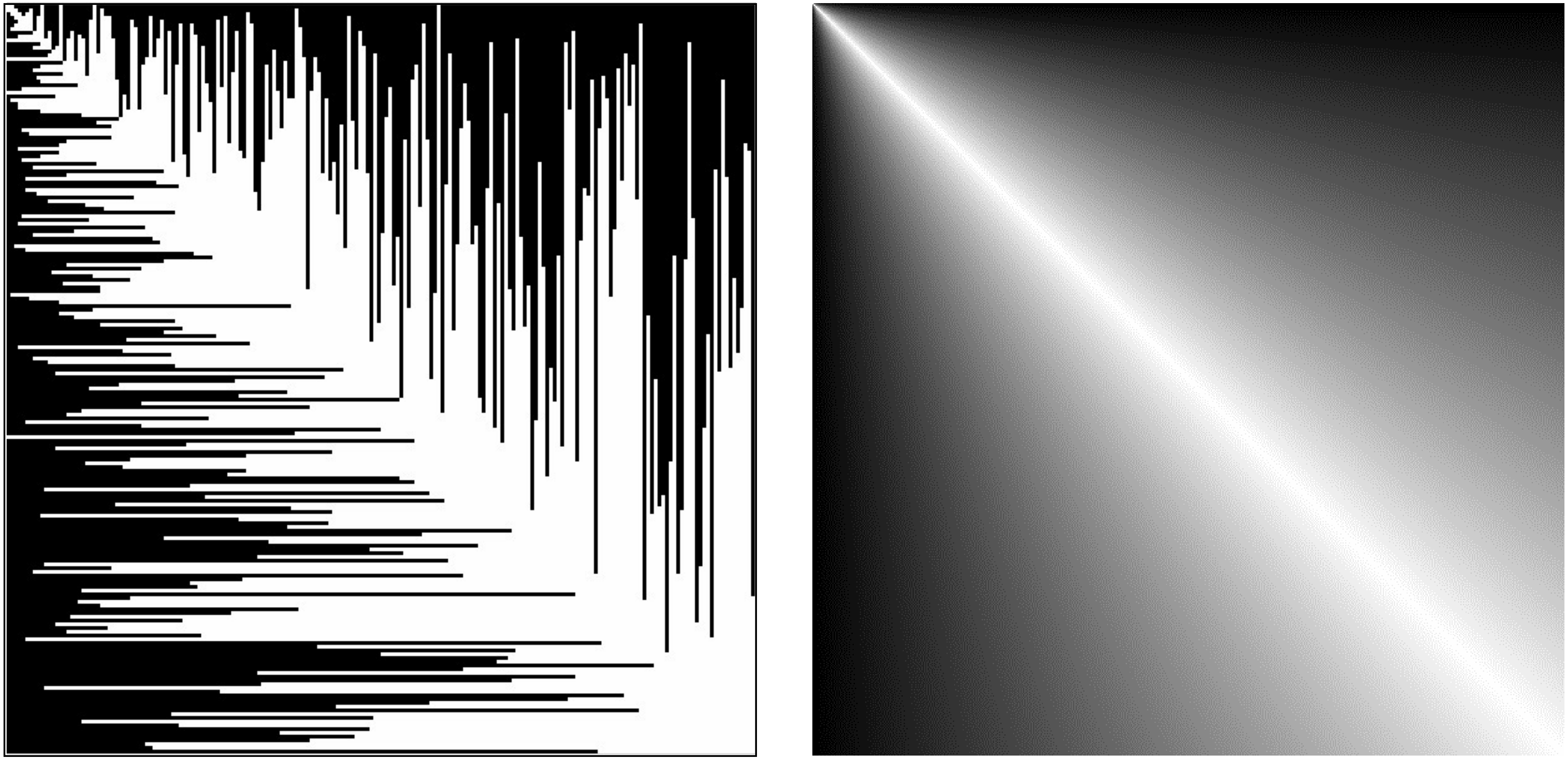
A randomly grown

uniform attachment graph

with 200 nodes

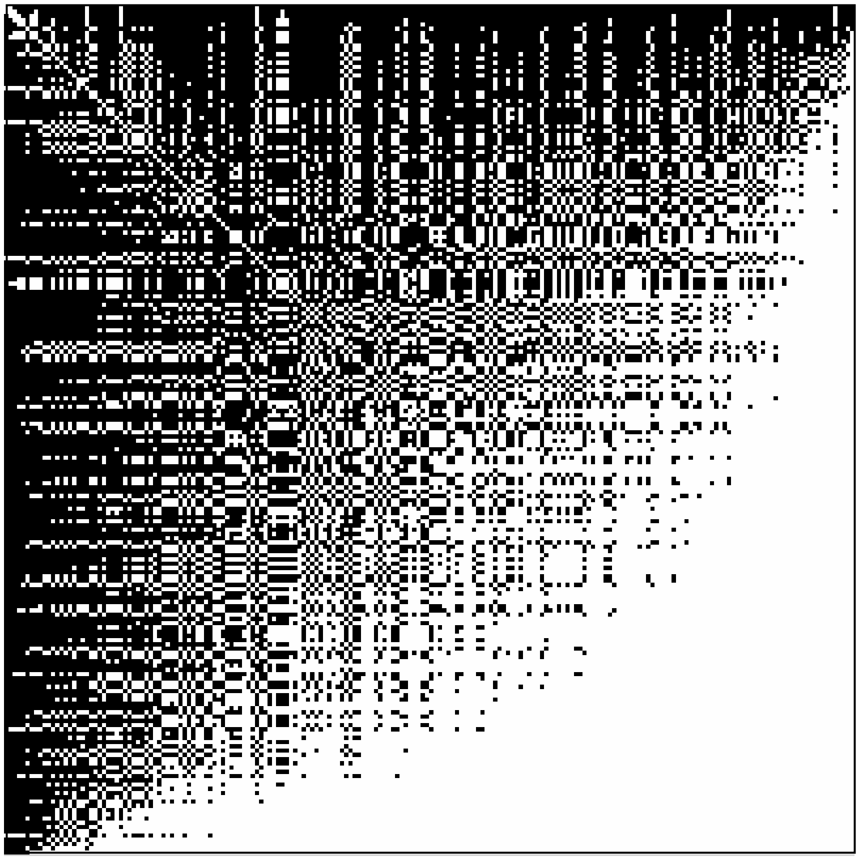


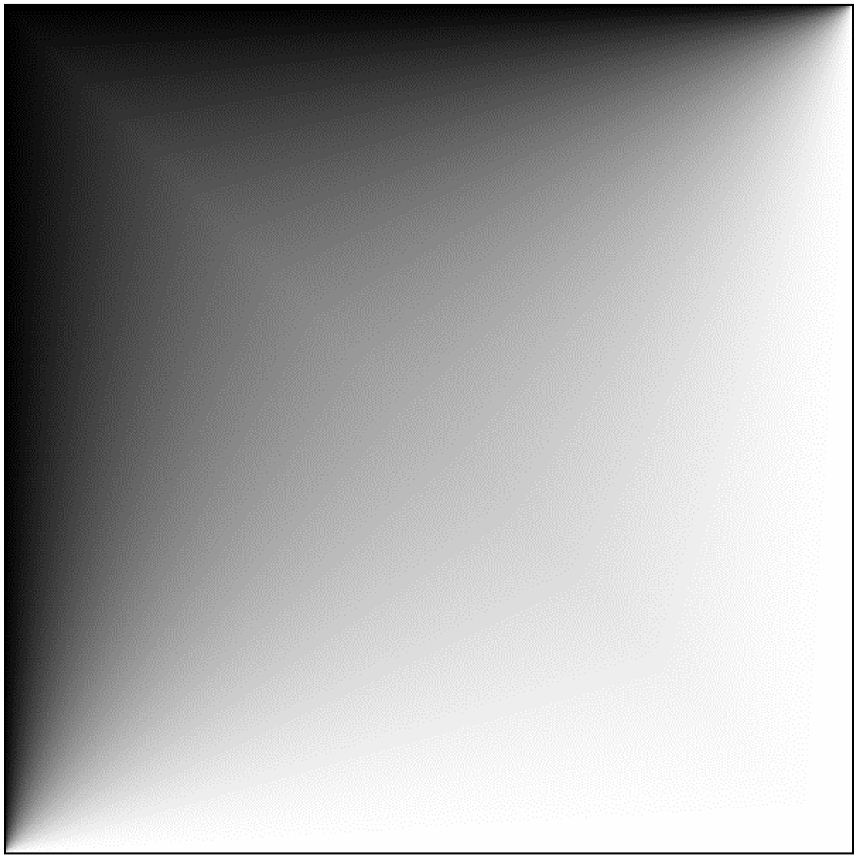
$1 - \max(x, y)$



A randomly grown prefix attachment graph

with 200 nodes

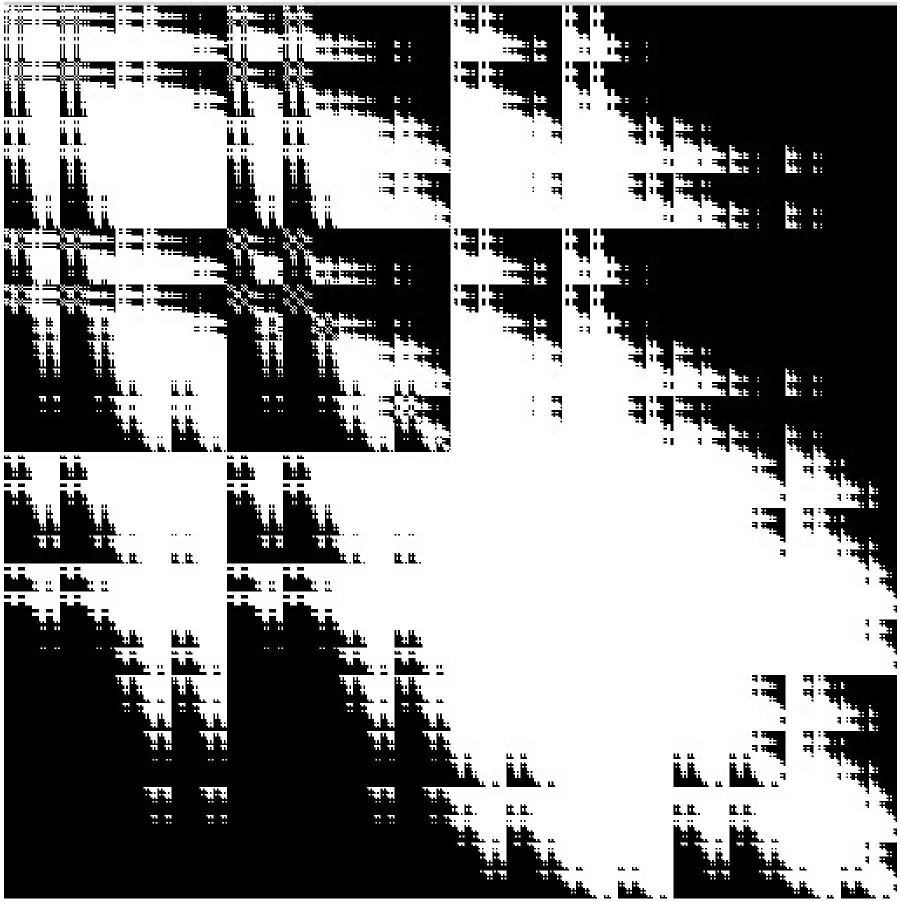




A randomly grown prefix attachment graph

with 200 nodes (ordered by degrees)

?



The limit of randomly grown prefix attachment graphs

(as a function on $[0,1]^2$)

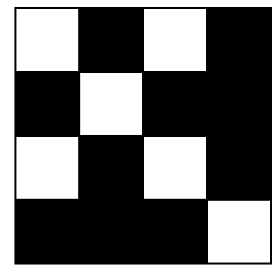
The limit object as a function

$$\mathcal{W}_0 = \left\{ W : [0,1]^2 \to [0,1] \ \text{symmetric, measurable} \right\}$$

$$t(F,W) = \int_{[0,1]^{V(F)}} \prod_{ij \in E(F)} W(x_i, x_j)\, dx$$

Example 1: Adjacency matrix of graph G: Associated function W_G:

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

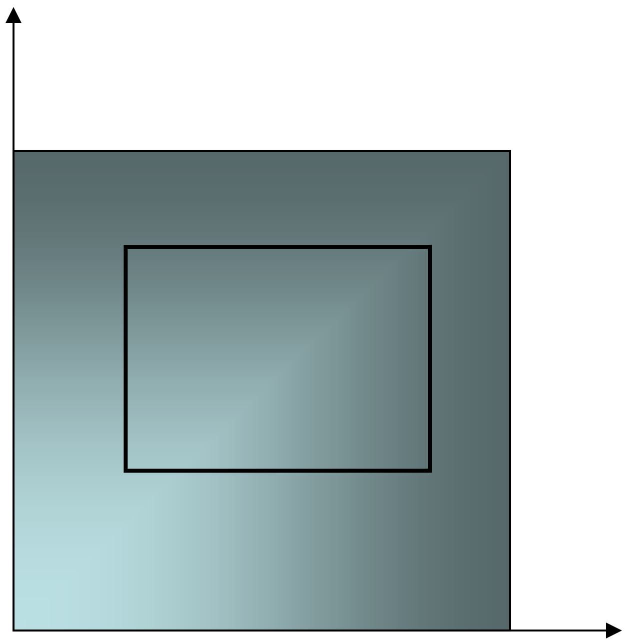


$$\boxed{t(F,G) = t(F,W_G)}$$

Example 2: $W(x,y) = \cos(2\pi(x-y))$

$t(F,W)=$ $2^{-|E(F)|}$ # of eulerian orientations of F

Distance of functions



$$\delta_{\square}(W,W') = \inf \sup_{S,T \subseteq [0,1]} \left| \int_{S\times T} (W - W') \right|$$

$$\delta_{\square}(G,G') = \delta_{\square}(W_G, W_{G'})$$

$$G_n \to W \iff \delta(W_{G_n}, W) \to 0$$

$$\mathcal{W}_0 = \{W : [0,1]^2 \to [0,1] \ \text{ symmetric, measurable}\}$$

$$t(F,W) = \int_{[0,1]^{V(F)}} \prod_{ij \in E(F)} W(x_i, x_j)\, dx$$

$$G_n \to W \iff (\forall F)\ t(F,G_n) \to t(F,W)$$

$$\mathbb{G}(n, \tfrac{1}{2}) \to \tfrac{1}{2}$$

Summary of main results

For every convergent graph sequence (G_n)
there is a W Î $\mathcal{W}_0$ such that G_n ® W

Conversely, $\forall W \, \exists (G_n)$ such that G_n ® W

W is essentially unique (up to measure-preserving transform).

$(\mathcal{W}_0, d_X)$ is compact.

So which function is the limit of the internet?

The internet is a sparse graph...

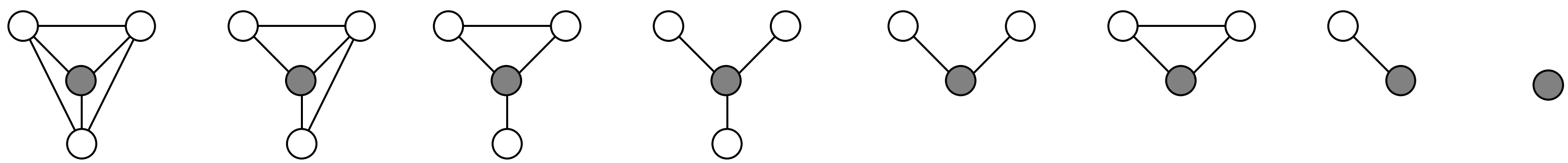
$\boxed{G_n\text{: sequence of graphs with degrees } \leq D}$

G_n is *left-convergent* if $\dfrac{\hom(F, G_n)}{|V(G_n)|}$ converges

$\forall$ connected F

Equivalent definition:

All possible neighborhoods with radius r



$$G \to (\ .1 \quad .2 \quad .13 \quad .27 \quad .2 \quad 0 \quad .1 \quad 0\ )$$

G_n left-convergent: this statistic converges for all r

G_n is right-convergent if

$$\frac{\log \hom(G_n, H)}{|V(G_n)|}$$

is convergent $\forall q \geq 1 \ \forall H$ in a neighborhood of J_q

J_q: complete graph K_q with loops, all weight 1

Right-convergent $\Leftrightarrow$ left-convergent

Borgs-Chayes-Kahn-L

Left-convergent $\Rightarrow \forall q \geq 2D \ \dfrac{\log \hom(G_n, K_q)}{|V(G_n)|}$ is convergent.

Number of q-colorings

H simple graph, $|V(H)| = q, \ \min\deg(H) \geq \left(1 - \dfrac{1}{2D}\right) q$

$\Rightarrow$

$\dfrac{\log \hom(G_n, H)}{|V(G_n)|}$ is convergent

G_n: $n \times n$ discrete torus

$$\hom(G_n, K_2) = \begin{cases} 2, \text{ if } n \text{ is even,} \\ 0, \text{ otherwise.} \end{cases}$$

Long-range interaction between colors

$$\frac{\log \hom(G_n, H)}{|V(G_n)|}$$ is convergent if H is connected nonbipartite.

Limits of graph sequences

Limits of sequences of graphs with bounded degree:

Aldous, Benjamini-Schramm, Lyons, Elek

Limits of sequences of dense graphs:

Borgs, Chayes, L, Sós, B.Szegedy, Vesztergombi

Limits loosing less information?

Distance?

Regularity Lemma?