Limits of dense graph sequences

Jan Hladký
Mathematics Institute,
Academy of Sciences of the Czech Republic

JH's research is supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme.

Random graphs

introduced by Erdős and Rényi in 1959
arguably the most studied random discrete structure

$n \in \mathbb{N}$, $p \in (0,1)$
$\mathbb{G}(n,p)$ is a distribution on graphs on the vertex set $\{1,\ldots,n\}$;
connect i and j with probability p, independently for each pair ij

One can ask various questions such as:

- Does $\mathbb{G}(n,0.3)$ contain a triangle?

- What is the size of the largest clique in $\mathbb{G}(n,0.3)$?

Random graphs

introduced by Erdős and Rényi in 1959
arguably the most studied random discrete structure

$n \in \mathbb{N}$, $p \in (0, 1)$
$\mathbb{G}(n, p)$ is a distribution on graphs on the vertex set $\{1, \ldots, n\}$;
connect i and j with probability p, independently for each pair ij

One can ask various questions such as:

- Does $\mathbb{G}(n, 0.3)$ contain a triangle?
  Asymptotically almost surely, as $n \to \infty$, $\mathbb{G}(n, 0.3)$ contains a triangle.
- What is the size of the largest clique in $\mathbb{G}(n, 0.3)$?
  Asymptotically almost surely, as $n \to \infty$, the largest clique in $\mathbb{G}(n, 0.3)$ is of size $\left(\frac{2}{\log(\frac{1}{0.3})} \pm \epsilon\right) \cdot \log n$.

Limits of dense graph sequences

Lovász, Szegedy *JCTB'06* (Fulkerson Prize'12)
Borgs, Chayes, Lovász, Sós, Vesztergombi *Adv.Math.'06*
Borgs, Chayes, Lovász, Sós, Vesztergombi *Ann.Math.'12*

Limits of dense graph sequences

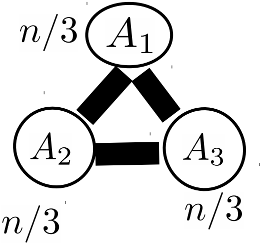
Lovász, Szegedy *JCTB'06* (Fulkerson Prize'12)
Borgs, Chayes, Lovász, Sós, Vesztergombi *Adv.Math.'06*
Borgs, Chayes, Lovász, Sós, Vesztergombi *Ann.Math.'12*

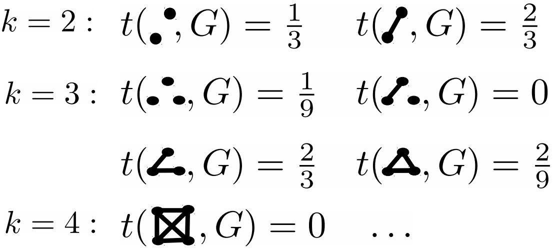
idea: convergence notion for sequences of finite graphs
compactification of the space of finite graphs $\Rightarrow$
... *graphons* symmetric Lebesgue-m. functions $\Omega^2 \to [0,1]$
Why? same story as with $\mathbb{Q}$ vs $\mathbb{R}$: only the latter allows
reasonable e.g. variational and integral calculus
for example *argmin*$(x^3 - 2x)$

Limits of dense graph sequences: an abstract approach

F is a "fixed graph" of order k, G is "large" of order n
We define **subgraph density** $t(F, G)$:

$$t(F,G) := \frac{\#\text{ copies of } F \text{ in } G}{\binom{n}{k}} = \mathbf{P}\big[G[\text{random } k\text{-set}] \cong F\big]$$





Limits of dense graph sequences: an abstract approach

F is a "fixed graph" of order k, G is "large" of order n
We define **subgraph density** $t(F, G)$:

$$t(F, G) := \frac{\#\text{ copies of } F \text{ in } G}{\binom{n}{k}} = \mathbf{P}\big[G[\text{random } k\text{-set}] \cong F\big]$$

A sequence of graphs $G_1, G_2, \dots$ **converges** if for each F, the sequence $t(F, G_1), t(F, G_2), \dots$ converges.
We get a **limit object** Ψ, $t(F, \Psi) = \lim_n t(F, G_n)$.

Why dense graph sequences?

If the proportion of edges $\searrow 0$ ($\lim \frac{e(G_n)}{n^2} = 0$) we get a trivial limit.
That is, the theory is void for trees, planar graphs, ...

Limits of dense graph sequences: an abstract approach

F is a "fixed graph" of order k, G is "large" of order n
We define **subgraph density** $t(F,G)$:

$$t(F,G) := \frac{\#\text{ copies of } F \text{ in } G}{\binom{n}{k}} = \mathbf{P}\big[G[\text{random } k\text{-set}] \cong F\big]$$

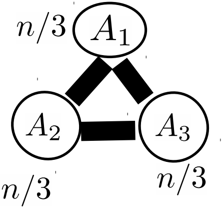
A sequence of graphs $G_1, G_2, \ldots$ **converges** if for each F, the sequence $t(F,G_1), t(F,G_2), \ldots$ converges.
We get a **limit object** Ψ, $t(F,\Psi) = \lim_n t(F,G_n)$.

Why dense graph sequences?

If the proportion of edges $\searrow 0$ ($\lim \frac{e(G_n)}{n^2} = 0$) we get a trivial limit.
That is, the theory is void for trees, planar graphs, ...

Razborov'07 flag algebras (next slide)

Extremal graph theory and Razborov's flag algebras I



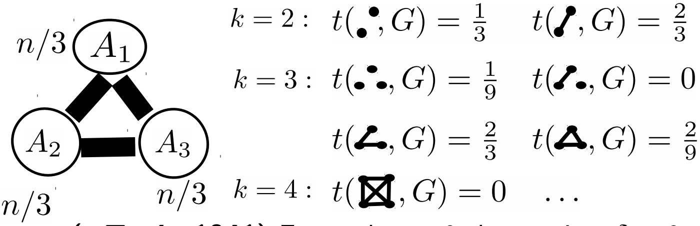
$k=2:$ $t($$, G) = \frac{1}{3}$ $t($$, G) = \frac{2}{3}$

$k=3:$ $t($$, G) = \frac{1}{9}$ $t($$, G) = 0$

$t($$, G) = \frac{2}{3}$ $t($$, G) = \frac{2}{9}$

$k=4:$ $t($$, G) = 0$ $\ldots$

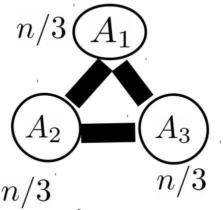
Extremal graph theory and Razborov's flag algebras I



Theorem (≈Turán 1941) For each $\epsilon > 0$ there exists $\delta > 0$: If an n-vertex graph has more than $(\frac{2}{3} + \epsilon)\binom{n}{2}$ edges then it contains $> \delta$-proportion of $\boxtimes$'s.

ExGrTh studies relations between $t(F_1, G)$, $t(F_2, G)$, $\dots$.

Extremal graph theory and Razborov's flag algebras I



$k = 2:$ $t($$, G) = \frac{1}{3}$ $t($$, G) = \frac{2}{3}$

$k = 3:$ $t($$, G) = \frac{1}{9}$ $t($$, G) = 0$

$t($$, G) = \frac{2}{3}$ $t($$, G) = \frac{2}{9}$

$k = 4:$ $t(\boxtimes, G) = 0$ $\ldots$

Theorem (≈Turán 1941) For each $\epsilon > 0$ there exists $\delta > 0$: If an n-vertex graph has more than $(\frac{2}{3} + \epsilon)\binom{n}{2}$ edges then it contains $> \delta$-proportion of $\boxtimes$'s.

ExGrTh studies relations between $t(F_1, G)$, $t(F_2, G)$, ...

Razborov: But lets rather study these relations on the limit space!
Approach to proving Turán: Suppose the theorem is false. $G_1, G_2, \ldots$ all contain $(\frac{2}{3} + \epsilon)$-proportion of edges but proportion of $\boxtimes$'s tends to 0. Pass to a subsequential limit Ψ. $t(|, \Psi) \geq \frac{2}{3} + \epsilon$ and $t(\boxtimes, \Psi) = 0$. Derive a contradiction.

Extremal graph theory and Razborov's flag algebras II

Razborov provides tools for deriving relations that hold on the limit object (Cauchy-Schwarz calculus, variational calculus, . . .)

First application: Razborov'08 (AMS Robbins Prize'12) solves the triangle density problem of Lovász and Simonovits 1983:
Suppose a graph has a given proportion of edges. What proportion of triangles can it have?

Extremal graph theory and Razborov's flag algebras II

Razborov provides tools for deriving relations that hold on the limit object (Cauchy-Schwarz calculus, variational calculus, . . .)

First application: Razborov'08 (AMS Robbins Prize'12) solves the triangle density problem of Lovász and Simonovits 1983:
Suppose a graph has a given proportion of edges. What proportion of triangles can it have?
Allows computer to aid for searching for the right inequalities (i.e., systematization of extremal graph theory)

Extremal graph theory and Razborov's flag algebras II

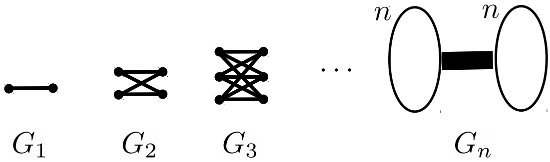
Razborov provides tools for deriving relations that hold on the limit object (Cauchy-Schwarz calculus, variational calculus, . . .)

First application: Razborov'08 (AMS Robbins Prize'12) solves the triangle density problem of Lovász and Simonovits 1983:
Suppose a graph has a given proportion of edges. What proportion of triangles can it have?
Allows computer to aid for searching for the right inequalities (i.e., systematization of extremal graph theory)

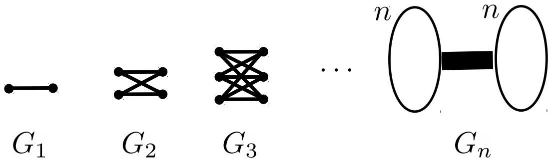
- H.-Král'-Norine'09: Caccetta-Häggkvist conj. (progress)
- H.-Hatami-Král'-Norine-Razborov'11 conjecture of Erdős 1984
- HHKNR'11 conjecture of Jagger-Štovíček-Thomason 1996
- . . . and many more

Hatami-Norine *J.AMS'11* deciding whether an inequality between subgraph densities holds for all graph limits is undecidable

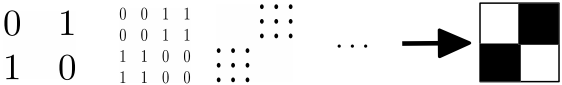
Graphons I



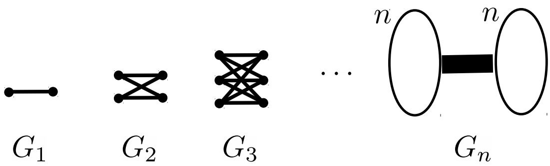
Graphons I



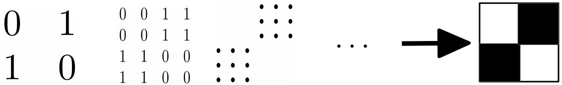
Represent these graphs by their adjacency matrices:



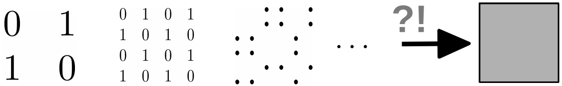
Graphons I



Represent these graphs by their adjacency matrices:



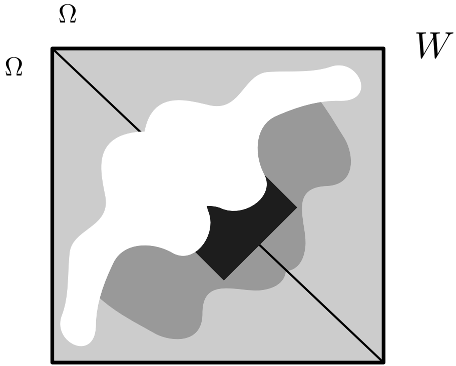
. . . works if you do things the right way. But, . . .



In general Szemerédi's Regularity Lemma can be used to determine "the right way" of ordering the vertices.

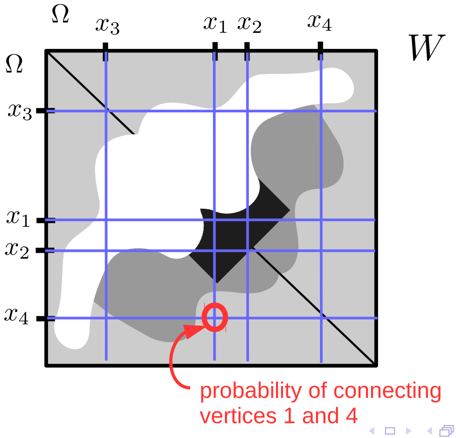
Graphons II

A **graphon** is a symmetric Lebesgue-m. function. $W : \Omega^2 \to [0,1]$.
Theorem (Lovász–Szegedy) sampling conv.$\Leftrightarrow$graphical conv.
Theorem (L–Sz.) Every graphon W can be achieved in the limit.
Proof:



Graphons II

A **graphon** is a symmetric Lebesgue-m. function. $W : \Omega^2 \to [0,1]$.
Theorem (Lovász–Szegedy) sampling conv.$\Leftrightarrow$graphical conv.
Theorem (L–Sz.) Every graphon W can be achieved in the limit.
Proof: Random graphs $G_1, G_2, \ldots$; $V(G_n) = \{1, \ldots, n\}$; sample $x_1, \ldots x_n \in \Omega$ and connect i with j with probability $W(x_i, x_j)$.



Graphons II

A **graphon** is a symmetric Lebesgue-m. function. $W : \Omega^2 \to [0,1]$.
Theorem (Lovász–Szegedy) sampling conv.$\Leftrightarrow$graphical conv.
Theorem (L–Sz.) Every graphon W can be achieved in the limit.
Proof: Random graphs $G_1, G_2, \ldots$; $V(G_n) = \{1, \ldots, n\}$; sample $x_1, \ldots x_n \in \Omega$ and connect i with j with probability $W(x_i, x_j)$.

It can be shown that almost surely, $G_1, G_2, \ldots \to W$. $\square$

Proof: Random graphs $G_1, G_2, \ldots$; $V(G_n) = \{1, \ldots, n\}$; sample $x_1, \ldots x_n \in \Omega$ and connect i with j with probability $W(x_i, x_j)$.

It can be shown that almost surely, $G_1, G_2, \ldots \to W$. $\square$

$\mathbb{G}(n, W)$ as a generalization of the Erdős–Rényi model $\mathbb{G}(n, p)$.

Interesting model *per se*!
Bollobás–Janson–Riordan'07, Doležal–H.–Máthé'15

Cliques in inhomogeneous random graphs: motivation

(joint work with Doležal and Máthé, arXiv:1510.02335)
notation $\omega(G)$ =size of the largest clique
motivation: some parameters are continuous, e.g. triangle density

$$G_1, G_2, \ldots \quad \to \quad W$$

$$\sum_{i,j,k} A_{G_n}(i,j) A_{G_n}(j,k) A_{G_n}(k,i)$$

Cliques in inhomogeneous random graphs: motivation

(joint work with Doležal and Máthé, arXiv:1510.02335)
notation $\omega(G)$ =size of the largest clique
motivation: some parameters are continuous, e.g. triangle density

$$G_1, G_2, \ldots \quad \rightarrow \quad W$$

$$\frac{1}{n^3} \cdot \sum_{i,j,k} A_{G_n}(i,j) A_{G_n}(j,k) A_{G_n}(k,i) \rightarrow \quad \int_{x,y,z} W(x,y)W(y,z)W(z,x)$$

Cliques in inhomogeneous random graphs: motivation

(joint work with Doležal and Máthé, arXiv:1510.02335)
notation $\omega(G)$ =size of the largest clique
motivation: some parameters are continuous, e.g. triangle density

$$\begin{array}{ccc} G_1, G_2, \ldots & \to & W \\ \frac{1}{n^3} \cdot \sum_{i,j,k} A_{G_n}(i,j) A_{G_n}(j,k) A_{G_n}(k,i) & \to & \int_{x,y,z} W(x,y)W(y,z)W(z,x) \end{array}$$

but the clique number is not. For example, we can find sequences

- $H_1, H_2, \ldots \to \mathbf{0}$ with $\omega(H_n) > n^{0.99}$
- $F_1, F_2, \ldots \to \mathbf{1}$ with $\omega(F_n) < \log \log n$

Cliques in inhomogeneous random graphs: motivation

(joint work with Doležal and Máthé, arXiv:1510.02335)
notation $\omega(G)$ =size of the largest clique
motivation: some parameters are continuous, e.g. triangle density

$$\begin{array}{ccc} G_1, G_2, \dots & \to & W \\ \frac{1}{n^3} \cdot \sum_{i,j,k} A_{G_n}(i,j) A_{G_n}(j,k) A_{G_n}(k,i) & \to & \int_{x,y,z} W(x,y)W(y,z)W(z,x) \end{array}$$

but the clique number is not. For example, we can find sequences

- $H_1, H_2, \dots \to \mathbf{0}$ with $\omega(H_n) > n^{0.99}$
- $F_1, F_2, \dots \to \mathbf{1}$ with $\omega(F_n) < \log \log n$

In order to get rid of these pathological examples, we study samples from $\mathbb{G}(n, W)$

Cliques in inhomogeneous random graphs: the result

Matula 1976, Grimmett–McDiarmid 1975:
For any fixed $p \in (0,1)$, asymptotically almost surely,

$$\omega(\mathbb{G}(n,p)) = \left(\frac{2}{\log(1/p)} \pm \epsilon\right) \cdot \log n \ .$$

Our result:
For any "reasonable" fixed graphon W, asymptotically almost surely,

$$\omega(\mathbb{G}(n,p)) = \kappa(W) \cdot \log n \ ,$$

where

$$\kappa(W) = \sup\left\{ \frac{2\|h\|_1^2}{\int_x \int_y h(x)h(y)\log\left(1/W(x,y)\right)} \ : \ h \in L^1(\Omega), h \geq 0 \right\} \ .$$