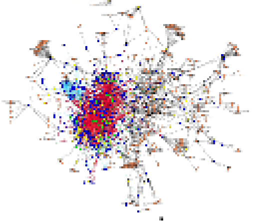
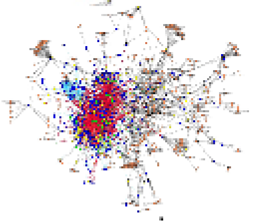


Graphons and Graph Limits Tutorial

Parts 1 & 2 (dense graphs)
Christian Borgs, UC Berkeley

Part 3 (sparse graphs)
Jennifer Chayes, UC Berkeley

Simons Institute for the Theory of Computing
Fall 2022

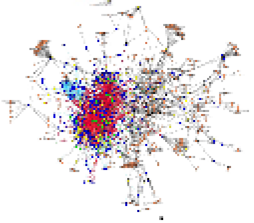


Motivation: Three Related Problems

Questions:

- When should we consider two large graphs to be similar?
- What is the "correct notion" of a limit of graphs (preserving "essential" properties of the finite graphs in the sequence)?
- How do I non-parametrically model massive real-world networks, and how do I estimate (learn) this non-parametric representation from data?

Lecture 1 + 2: Dense Graphs, summarizing our works with Lovasz, Sos and Vesztergombi [BCLSV '06-'12] and a few others



Motivation: when are two graphs similar?

<u>Combinatorialists/Social Scientists:</u>

- If they have similar local properties, in particular, subgraph counts

<u>Statistics:</u>

- If sampled subgraphs have similar distributions

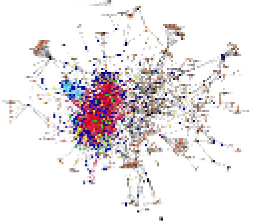
<u>Computer Science:</u>

- If they have similar global properties, in particular max cut, min-bisection, etc.

<u>Physicists:</u>

- If statistical physics models on them have similar free energies or ground state energies

<u>Thm1:</u> [BCLSV’06,’08,’12,Diaconis-Janson’07]: For dense graphs, these are all equivalent, and they are also equivalent to similarity in the so-called cut-metric

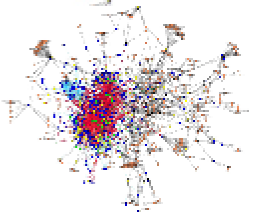


Motivation: what is the right notion of a limit?

- A collection of limiting subgraph counts?
- A collection of distributions on finite graphs?
- A collection of (suitable generalizations) of max-cut, min-bisections, …?
- A collection of free energies or ground state energies?

Thm2 [Lovasz-Szegedy’06, BCLSV’06 -’12]: For dense graphs, all the limiting quantities above can be described in terms of a graphon over $[0,1]$.

Def: A graphon over $[0,1]$ is a function W: $[0,1] \times [0,1] \to [0,1]$ s.th. $W(x, y)$=$W(y, x)$ for all $x, y \in \Omega$



Motivation: How to model large graphs?

Simple Models:

- $G(n,p)$
- Stochastic Block Model

Q: What is the "right" generalization

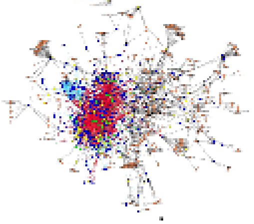
Thm3 [Aldous '81, Hoover '79]:

All "natural" dense random models can be generated by a (possibly random) graphons

Def: Given a probability space (Ω, μ), a graphon is a symmetric* 2-variable function

$$W: \Omega \times \Omega \to [0,1]: (x,y) \mapsto W(x,y)$$

*Here W is symmetric if $W(x,y)=W(y,x)$ for all $x,y \in \Omega$



Summary: Graphs and Graphons

Graphs

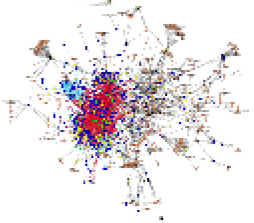
- Vertex set V
- Adjacency matrix $A: V \times V \to \{0,1\}$

Graph Limits ↓

↑ Non-Parametric Random Graph Models

Graphons

- Probability space $(\Omega, \mathcal{F}, \mu)$
- Symmetric, measurable function $W: \Omega \times \Omega \to [0,1]$



1) Modelling Random Graphs

Random Graph $G(n, p)$:

- vertex set $[n] =\{1, \ldots, n\}$
- each of the possible $\binom{n}{2}$ edges is present i.i.d. with probability p

Stochastic Block Model $SBM(n, B)$, where $B \in [0,1]^{k\times k}$ is a symmetric matrix:

- vertex set $[n]$
- each vertex has a label $x_i \in [k]$ chosen i.i.d. uniformly at random
- $i < j$ are connected independently with probability $P_{ij} = B_{x\_ix\_j}$

Inhomogneous Random Graph: $G_n(W)$

Start with a graphon, i.e., a symmetric function W over some probability space (Ω, μ)

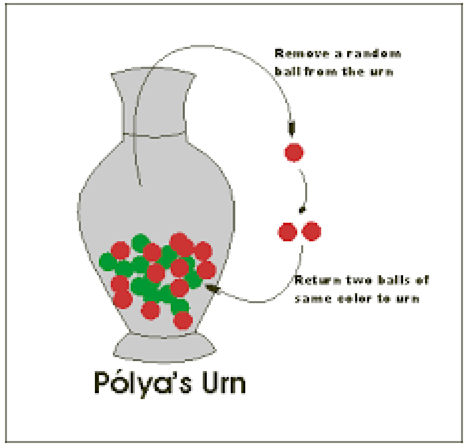
- vertex set $[n]$
- each vertex has a feature $x_i \in \Omega$ chosen i.i.d. according to μ
- $i < j$ are connected independently with probability $P_{ij} = W(x_i, x_j)$

Q: How general is this?

1a) De Finetti

Def: An infinite sequence of random variables $X_1, X_2, \ldots \in \{0,1\}$ is called exchangeable if for all n and all permutations $\pi: [n] \to [n]$,

$X_{\pi(1)}, \ldots, X_{\pi(n)}$ has the same distribution as $X_1, \ldots, X_n$



Ex: Polya-Urn

- Start with R red and G green balls
- Pull out a ball, and replace it with two of the same color
- Iterate

$$\Pr(rrrgg) = \frac{R(R+1)(R+2)G(G+1)}{(R+G)(R+G+1)\ldots(R+G+4)}$$

$$\Pr(rgrgr) = \frac{RG(R+1)(G+1)(R+2)}{(R+G)(R+G+1)\ldots(R+G+4)}$$

1a) De Finetti

Thm [De Finetti]: Assume $X_1, X_2, \dots \in \{0,1\}$ is exchangeable

Then there exists a distribution μ on $[0,1]$ s.th. $X_1, X_2, \dots$ can be obtained by

- first drawing $p \sim \mu$, and then
- choosing $X_1, X_2, \dots$ i.i.d. with distribution $Be(p)$.

Ex. Polya Urn: μ is the beta-distribution $\beta(R, G)$

1b) Aldous-Hoover Theorem

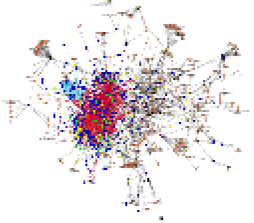
Exchangeable random graphs: an infinite random graph whose distribution is invariant under finite vertex relabeling is called *exchangeable*

Formal Definition in terms of adjaceny matrix:

An infinite random array $(X_{ij})_{ij\in\mathbb{N}}$ with entries in $\{0,1\}$ is called *exchangeable* if for all n and all permutations $\pi: [n] \to [n]$,

$(X_{\pi(i)\pi(j)})_{i,j\leq n}$ has the same distribution as $(X_{ij})_{i,j\leq n}$

Q: What is the analogue of De Finetti? $G(n, p)$ for a random p?



1b) Aldous-Hoover Theorem

Thm [Aldous-Hoover]: Let $(X_{ij})_{ij\in\mathbb{N}}$ be an exchangeable array with entries in $\{0,1\}$ and $X_{ii} = 0$.

Then there exists a measurable function $(x, y, \alpha) \mapsto W_\alpha(x, y)$ from $[0,1]^3 \to [0,1]$ s.th. $(X_{ij})_{ij\in\mathbb{N}}$ can be generated by

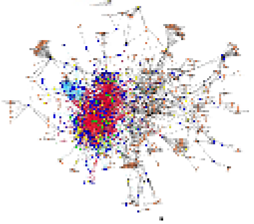
- first choosing $\alpha \in [0,1]$ uniformly at random,
- then choosing $x_1, x_2, \ldots$ i.i.d. uniformly at random in $[0,1]$,
- and then choosing $X_{ij} = X_{ji} \sim Be(W_\alpha(x_i, x_j))$, independently for all $i < j$

Rephrased: If G_n is a finite subgraph of an exchangeable infinite random graph G_∞ then the distribution of G_n can be generated by a random graphon W

1b) Aldous-Hoover Theorem

<u>Summary:</u>

- A graphon is a symmetric 2-variable function over a probability space (Ω, μ), $W: \Omega \times \Omega \to [0,1]: (x, y) \mapsto W(x, y)$
- It generates inhomogeneous random graph $G_n(W)$ on by
  - assigning i.i.d. features $x_i \in \Omega$ according to μ to the vertices
  - connected $i < j$ independently with probability $P_{ij} = W(x_i, x_j)$
- By Aldous- Hoover, any exchangeable family of random graphs $(G_n)_{n\geq 1}$ can be generated by a (possibly random) graphon W



Graphs and Graphons

Graphs

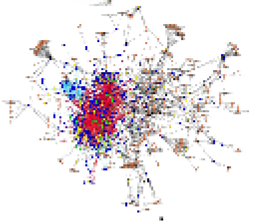
- Vertex set V
- Adjacency matrix $A: V \times V \rightarrow \{0,1\}$

↓ Graph Limits

↑ Non-Parametric Random Graph Models

Graphons

- Probability space $(\Omega, \mathcal{F}, \mu)$
- Symmetric, measurable function $W: \Omega \times \Omega \rightarrow [0,1]$



2) Different Notions of Similarity

Combinatorialists/Social Scientists:

- Similar local properties, in particular, subgraph counts

Statistics:

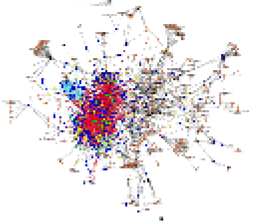
- Similar distributions for sampled subgraphs

Computer Science:

- Similar global properties, in particular max cut, min-bisection, etc.

Physicists:

- Similar free energies or ground state energies



2a) Subgraph counts

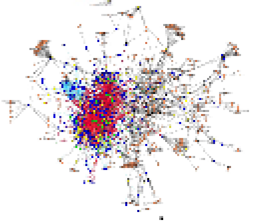
Idea: Test a large graph $G = (V, E)$ *"from the left"* by mapping a small graph H into G

Def: Subgraph frequencies: Given a graph $G = (V, E)$ with adjacency matrix A and a graph H on k nodes, define

$$t_0(H, G) = \frac{1}{|V|^k} \sum_{v_1,\dots,v_k \in V} \prod_{ij \in E(H)} A_{v_i v_j} \prod_{ij \notin E(H)} (1 - A_{v_i v_j})$$

Def: Subgraph Count Convergence:

- For all finite graphs H, $t_0(H, G_n)$ converges to some $t_0(H) \in [0,1]$



2b) Sampling

Given a graph $G = (V, E)$ and an integer $k \geq 1$, choose $x_1, \ldots, x_k \in V$, uniformly at random with replacement

- $Smpl_k(G)$ is the k-node graph with edge set $\{ij : x(i)x(j) \in E\}$

Def: A sequence of dense graphs G_n is called sampling convergent if the distribution of $Smpl_k(G_n)$ converges for all k

Rem: Sampling convergence is clearly equivalent to subgraph count convergence. We call this notion left-convergence

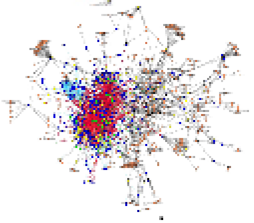
2c) Multiway Cuts

Notation: Given a graph $G = (V, E)$ on n nodes and $S, T \subset V$, set

$$e_G(S,T) = \frac{1}{n^2} \sum_{i \in S, j \in T} 1_{ij \in E}$$

$$MaxCut(G) = \max_{S \subset V} e_G(S, S^c),\ MinBisec(G) = \min_{S:|S|=\frac{n}{2}} e_G(S, S^c), \ldots$$

Q: How to generalize this for cuts into more than two groups?



2c) Multiway Cuts

Multiway-cuts: Given $J \in \mathbb{R}^{k\times k}$ and $\sigma: V \to [k]$ define

$$E_{G,J}(\sigma) = \frac{1}{n^2} \sum_{x,y:\{x,y\}\in E} J_{\sigma(x)\sigma(y)}$$

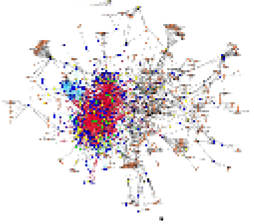
and for $\alpha \in \Delta_k$, set

$$MinCut_{J,\alpha}(G) = \min_{\sigma} E_{G,J}(\sigma)$$

where the minimum goes over all maps $\sigma: V \to [k]$ such that

$$\left|\, \left|\sigma^{-1}(\{i\})\right| - n\alpha_i \right| \leq 1 \text{ for all } i \in [k]$$

Rem: We call convergence of these multi-way cuts right convergence



2d) Statistical Physics

In statistical physics, $\sigma: V \to [k]$ is called a spin-configuration, $E_{G,J}(\sigma)$ is called its energy, and $MinCut_{J,\alpha}(G)$ is called the micro-canonical ground state energy.

Def: Micro-canonical free energy

$$F_{J,\alpha}(G) = -\frac{1}{n} \log Z_{J,\alpha}(G)$$

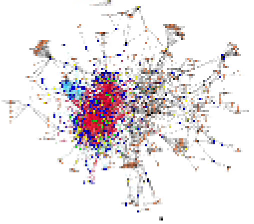
where $Z_{J,\alpha}(G)$ is the partition function

$$Z_{J,\alpha}(G) = \sum_{\sigma: V \to [k]} e^{-nE_{G,J}(\sigma)}$$

and the sum is over all $\sigma: V \to [k]$ such that

$$\left| \left|\sigma^{-1}(\{i\})\right| - n\alpha_i \right| \leq 1 \text{ for all } i \in [k]$$

Rem: This is another version of right convergence

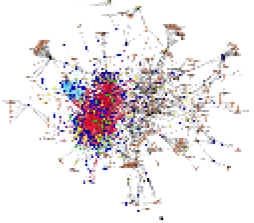


2e) All these notions are equivalent!

Thm: Let G_n be a sequence of graphs. Then the following are equivalent

1) For all finite graphs H, the subgraph frequencies $t_0(H, G_n)$ converge
2) For all $k \geq 1$, the distributions of $Smpl_k(G_n)$ converge
3) For all $k \geq 1$, $J \in \mathbb{R}^{k \times k}$ and $\alpha \in \Delta_k$, the multi-way cuts $MinCut_{J,\alpha}(G_n)$ converge
4) For all $k \geq 1$, $J \in \mathbb{R}^{k \times k}$ and $\alpha \in \Delta_k$, the micro-canonical free energies $F_{J,\alpha}(G_n)$ converge

Proof Idea: prove equivalence to being a Cauchy sequence in the cut-metric



3) Cut-Metric

Q: How do we compare to graphs on different numbers of nodes.

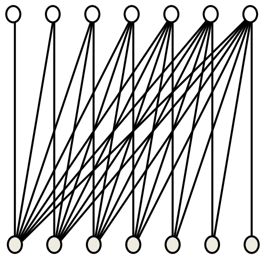
Step 1: Embed graphs into the space of graphons:

Empirical Graphon of a Graph G on n nodes

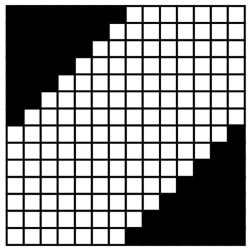
- Replace $[n]$ by n disjoint intervals $I_1, \dots, I_n$ of width $1/n$ and divide $[0,1]^2$ into n^2 squares $I_i \times I_j$ of side length $1/n$
- Set W_G to 1 on the square ij if ij is an edge in G and to 0 otherwise

Example:

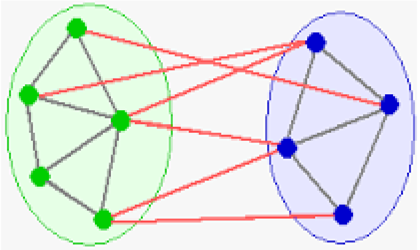
Half graph



$\sim$



3) Cut-Metric



Step 2: Cut norm* of a function $W: [0,1]^2 \to \mathbb{R}$

$$\|W\|_\Box = \max_{S,T \subset [0,1]} \left| \int_{S\times T} W(x,y)\,dx\,dy \right|$$

Problem: In general, isomorphic graphs have a non-zero distance

Step 3: For two graphons $W_1, W_2: [0,1]^2 \to [0,1]$ define the cut metric

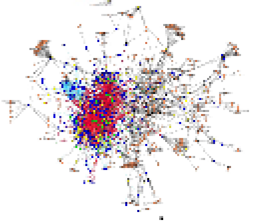
$$\delta_\Box(W_1, W_2) = \inf_\phi \left\| W_1^\phi - W_2 \right\|_\Box$$

where the infimum goes over measure preserving bijections and

$$W_1^\phi(x,y) = W_1(\phi(x), \phi(y))$$

*)Equivalently, we can define $\|W\|_\Box$ by

$$\|W\|_\Box = \max_{f,g:[0,1]\to[0,1]} \left| \int f(x) W(x,y) g(y)\,dx\,dy \right|$$



3) Cut-Metric

<u>Def</u>: For two finite graphs G_1, G_2 we set

$$\delta_\square(G_1, G_2) := \delta_\square\left(W_{G_1}, W_{G_2}\right)$$

$$= \inf_{\phi} \max_{S,T\subset[0,1]} \left| \int_{S\times T} \left(W_{G_1}(\phi(x), \phi(y)) - W_{G_2}(x, y) \right) dxdy \right|$$

3a) Comments on Proof Structure

Thm: Let G_n be a sequence of graphs. Then the following are equivalent

1) For all finite graphs H, the subgraph frequencies $t_0(H, G_n)$ converge
2) For all $k \geq 1$, the distributions of $Smpl_k(G_n)$ converge
3) For all $k \geq 1$, $J \in \mathbb{R}^{k \times k}$ and $\alpha \in \Delta_k$, the multi-way cuts $MinCut_{J,\alpha}(G_n)$ converge
4) For all $k \geq 1$, $J \in \mathbb{R}^{k \times k}$ and $\alpha \in \Delta_k$, the micro-canonical free energies $F_{J,\alpha}(G_n)$ converge
5) G_n is a Cauchy sequence in the cut metric $\delta_\square$

Proof Idea:

I) Prove that if $\delta_\square(G, G') \leq \epsilon$, the other properties differ by at most a constant times ϵ (the constant you will get will be moderate, roughly proportional to k^2, and maybe the norm of J). These proof are relatively elementary

II) The other direction is more difficult, and often will require k to be exponentially large in $1/\epsilon^2$

I will show this for some of the above quantities, to give you an idea of the flavor of the proofs.