DENSE RANDOM GRAPHS: INTRODUCTION

Frank den Hollander
Leiden University, The Netherlands

Spring School in Probability: Complex Networks,
Technische Universität Darmstadt, Germany, 2-6 March 2020.

§ GRAPHONS

For $n \in \mathbb{N}$, let $\mathcal{G}_n$ denote the set of all $2^{\binom{n}{2}}$ undirected labelled simple graphs with n vertices. Any graph $G \in \mathcal{G}_n$ can be represented by a symmetric $n \times n$ matrix with elements

$$A^G(i,j) := \begin{cases} 1, & i \sim j, \\ 0, & i \not\sim j, \end{cases}$$

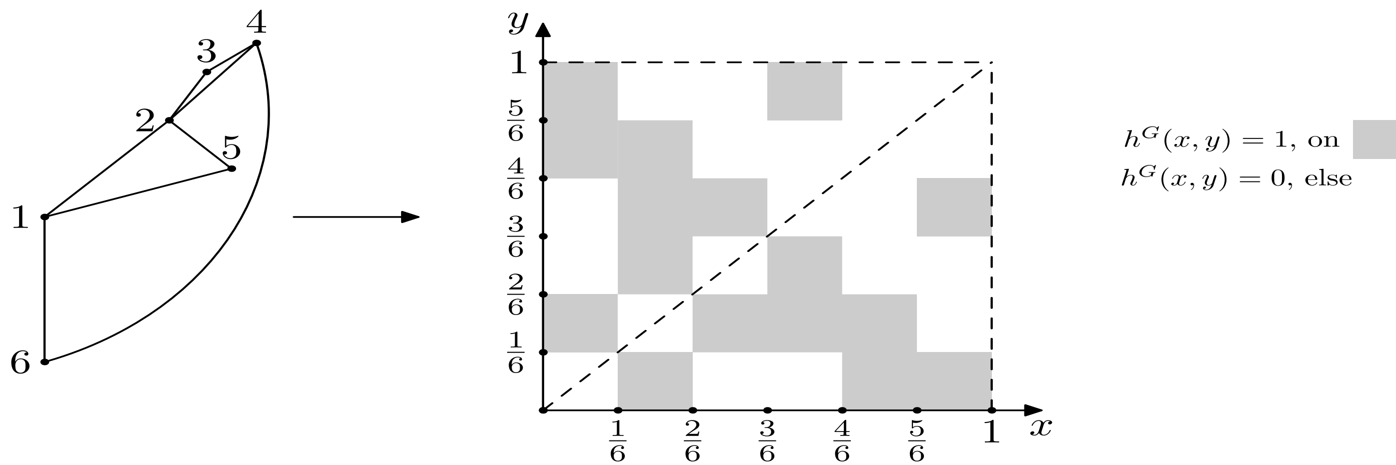
called the adjacency matrix.

There is a natural way to embed A^G into a set of functions called graphons. Namely, let

$$W = \Big\{ h\colon\ [0,1]^2 \to [0,1]\colon\ h(x,y) = h(y,x)\,\forall\, x,y \Big\}$$

and represent G as a graphon $h^G \in W$ by setting

$$h^G(x,y) := \begin{cases} 1, & \lceil nx \rceil \sim \lceil ny \rceil, \\ 0, & \lceil nx \rceil \not\sim \lceil ny \rceil. \end{cases}$$



An example of a graph G with $n = 6$

and its associated graphon h^G

The space of graphons W is endowed with the cut distance

$$d_\square(h_1, h_2) := \sup_{S,T\subset[0,1]} \left| \int_{S\times T} \mathrm{d}x\,\mathrm{d}y\,[h_1(x,y) - h_2(x,y)] \right|.$$

On W there is a natural equivalence relation. Indeed, let Σ be the set of measure-preserving bijections $\sigma\colon [0,1] \to [0,1]$. We say that

$$h_1 \equiv h_2 \text{ if and only if } h_1^\sigma = h_2 \text{ for some } \sigma \in \Sigma,$$

where $h_1^\sigma(x,y) = h_1(\sigma x, \sigma y)$. This equivalence relation yields the quotient space $(\tilde{W}, \delta_\square)$, where $\delta_\square$ is called the cut metric, defined by

$$\delta_\square(\tilde{h}_1, \tilde{h}_2) := \inf_{\sigma_1,\sigma_2\in\Sigma} d_\square(h_1^{\sigma_1}, h_2^{\sigma_2}).$$

The equivalence classes arise naturally from a relabeling of the vertices of the graph.

Let F and G be two finite simple graphs with vertex sets $V(F)$ and $V(G)$, and let $\hom(F,G)$ denote the number of homomorphisms from F to G. The homomorphism density is defined as

$$t(F,G) := \frac{1}{|V(G)|^{|V(F)|}} \hom(F,G).$$

Two graphs are said to be similar when they have similar homomorphism densities.

DEFINITION:

A sequence of labelled simple graphs $(G_n)_{n\in\mathbb{N}}$ is said to be convergent when $(t(F,G_n))_{n\in\mathbb{N}}$ converges for any finite simple graph F.

Consider a finite simple graph F, with $|V(F)| = k$ and edge set $E(F)$, and let $h \in W$. Define the density

$$t(F,h) := \int_{[0,1]^k} \mathrm{d}x_1 \cdots \mathrm{d}x_k \prod_{(i,j)\in E(F)} h(x_i, x_j).$$

Then

$$t(F,h^G) = \int_{[0,1]^k} \mathrm{d}x_1 \cdots \mathrm{d}x_k \prod_{(i,j)\in E(F)} h^G(x_i, x_j)$$
$$= \frac{1}{|V(G)|^{|V(F)|}} \mathrm{hom}(F,G) = t(F,G).$$

Hence, a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is convergent to $h \in W$ when

$$\lim_{n\to\infty} t(F,G_n) = t(F,h) \qquad \forall\, F \text{ finite simple.}$$

Think of h as the $n \to \infty$ limit of the graphon h^{G_n} of a dense graph $G_n \in \mathcal{G}_n$, i.e., whose degrees are of order n.

§ THREE BASIC PROPERTIES

PROPOSITION 1: Borgs et al. 2008

The following are equivalent:

(i) $(G_n)_{n\in\mathbb{N}}$ *is convergent.*

(ii) $(\tilde{h}^{G_n})_{n\in\mathbb{N}}$ *is Cauchy in the* $\delta_\square$*-metric.*

(iii) $(t(F,h^{G_n}))_{n\in\mathbb{N}}$ *converges* $\forall\, F$ *finite simple.*

(iv) $\exists\, h \in W$*:* $\lim_{n\to\infty} t(F,h^{G_n}) = t(F,h)\ \forall\, F$ *finite simple.*

PROPOSITION 2: Lovász & Szegedy 2006

$(\tilde{W},\delta_\square)$ *is compact.*

PROPOSITION 3: Borgs et al. 2008

Let G_1, G_2 *be two labelled finite simple graphs. Then*

$|t(F,G_1) - t(F,G_2)| \leq 4|E(F)|\delta_\square(G_1,G_2) \quad \forall\, F$ finite simple.

§ GRAPHON OPERATORS

With $h \in W$ we associate a graphon operator acting on $L^2([0,1])$:

$$(hu)(x) = \int_{[0,1]} \mathrm{d}y\, h(x,y)u(y), \qquad x \in [0,1].$$

We use the symbol h both for the graphon and the graphon operator. The operator norm of the graphon operator h is defined as

$$\|h\| = \sup_{\substack{u \in L^2([0,1]) \\ \|u\|_2 = 1}} \|hu\|_2,$$

where $\|\cdot\|_2$ denotes the L^2-norm.

Given a graphon $h \in \mathcal{W}$, we have that $\|h\| \leq \|h\|_2$. Hence, a graphon sequence converging in the L^2-norm also converges in the operator norm.

The product of two graphons $h_1, h_2 \in W$ is defined as

$$(h_1 h_2)(x,y) = \int_{[0,1]} \mathrm{d}z \, h_1(x,z) h_2(z,y), \quad x,y \in [0,1],$$

and the m-th power of a graphon $h \in W$ as

$$h^m(x,y) = \int_{[0,1]^{m-1}} \mathrm{d}z_1 \cdots \mathrm{d}z_{m-1} \, h(x,z_1) \times \cdots \times h(z_{m-1},y), \quad x,y \in [0,1],\ m \in \mathbb{N}.$$

DEFINITION $\lambda \in \mathbb{R}$ is said to be an eigenvalue of the graphon operator h if there exists a nonzero function $u \in L^2([0,1])$ such that

$$(hu)(x) = \lambda u(x), \quad x \in [0,1].$$

The function u is said to be an eigenfunction associated with λ.

PROPOSITION 4: Borgs et al. 2008

The following hold for h as operator:

(i) h is self-adjoint, bounded and continuous.

(ii) h is diagonalisable and has countably many eigenvalues, all of which are real and can be ordered as

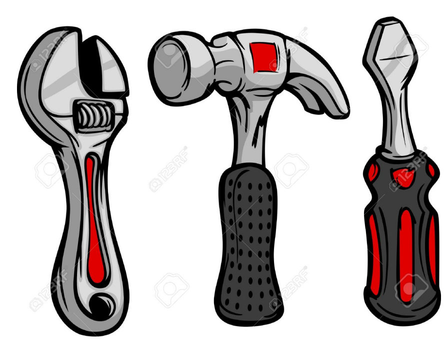
$$\lambda_1 \geq \lambda_2 \geq \dots$$

with an orthonormal basis of eigenfunctions for $L^2([0,1])$.

(iii) $\lambda_1 > 0$, with an associated eigenfunction $u_1 > 0$.

(iv) $\lambda_1 = \|h\|$.

Monograph on graphons: Lovász 2012



§ LDP FOR ERDŐS-RÉNYI

For $p \in (0,1)$ and $u \in [0,1]$, let

$$I_p(u) := u \log\left(\frac{u}{p}\right) + (1-u)\log\left(\frac{1-u}{1-p}\right),$$

with the convention that $0 \log 0 = 0$. For $h \in W$ write

$$I_p(h) := \int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\; I_p(h(x,y)).$$

On the quotient space, $(\tilde{W}, \delta_\square)$ define $I_p(\tilde{h}) = I_p(h)$, where h is any element of the equivalence class $\tilde{h}$.

THEOREM 1: Chatterjee & Varadhan 2011
I_p is a good rate function, i.e., $I_p \not\equiv \infty$ and I_p is lower semi-continuous and has compact level sets with respect to the $\delta_\square$-metric.

Consider the set $\mathcal{G}_n$ of all graphs on n vertices and the Erdős-Rényi probability distribution $\mathbb{P}_{n,p}$ on $\mathcal{G}_n$. Through the mappings $G \to h^G \to \tilde{h}^G$, we obtain a probability distribution on W, also denoted by $\mathbb{P}_{n,p}$, and a probability distribution $\tilde{\mathbb{P}}_{n,p}$ on $\tilde{W}$.

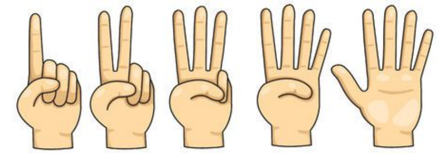
THEOREM 2: Chatterjee & Varadhan 2011 **LDP**

$(\tilde{\mathbb{P}}_{n,p})_{n\in\mathbb{N}}$ *satisfies the large deviation principle on* $(\tilde{W}, \delta_\square)$ *with rate* $\binom{n}{2}$ *and with rate function* I_p*, i.e.,*

$$\limsup_{n\to\infty} \binom{n}{2}^{-1} \log \tilde{\mathbb{P}}_{n,p}(\tilde{C}) \leq - \inf_{\tilde{h}\in\tilde{C}} I_p(\tilde{h}) \quad \forall\, \tilde{C} \subset \tilde{W} \ \textit{closed},$$

$$\liminf_{n\to\infty} \binom{n}{2}^{-1} \log \tilde{\mathbb{P}}_{n,p}(\tilde{O}) \geq - \inf_{\tilde{h}\in\tilde{O}} I_p(\tilde{h}) \quad \forall\, \tilde{O} \subset \tilde{W} \ \textit{open}.$$

§ COUNTING GRAPHS



Using the LDP, we can find asymptotic expressions for the number of simple graphs on n vertices with a given property. In what follows, a property of a graph is defined through an operator $\vec{T}\colon W \to \mathbb{R}^m$ for some $m \in \mathbb{N}$, which we may view as a constraint.

We assume that the operator $\vec{T}$ is continuous with respect to the $\delta_\square$-metric, and for some $\vec{T}^* \in \mathbb{R}^m$ we consider the sets

$$\tilde{W}^* := \left\{\tilde{h} \in \tilde{W} :\ \vec{T}(\tilde{h}) = \vec{T}^*\right\},$$
$$\tilde{W}^*_n := \left\{\tilde{h} \in \tilde{W}^* :\ \tilde{h} = \tilde{h}^G \text{ for some } G \text{ on } n \text{ vertices}\right\}.$$

By the continuity of the operator $\vec{T}$, the set $\tilde{W}^*$ is closed.

COROLLARY: Chatterjee 2016

For any measurable set $\tilde{W}^* \subset \tilde{W}$*,*

$$-\inf_{\tilde{h}\in \mathrm{int}(\tilde{W}^*)} I(\tilde{h}) \le \liminf_{n\to\infty} \binom{n}{2}^{-1} \log|\tilde{W}_n^*| \le \limsup_{n\to\infty} \binom{n}{2}^{-1} \log|\tilde{W}_n^*| \le -\inf_{\tilde{h}\in\tilde{W}^*} I(\tilde{h}),$$

where $\mathrm{int}(\tilde{W}^*)$ *is the interior of* $\tilde{W}^*$*, and*

$$I(\tilde{h}) = I(h) := \int_{[0,1]^2} \mathrm{d}x\, \mathrm{d}y\; I(h(x,y)), \qquad h \in W,$$

with entropy function

$$I(u) := u \log u + (1-u)\log(1-u), \qquad u \in [0,1].$$

§ LDP FOR INHOMOGENEOUS ERDŐS-RÉNYI

It is possible to extend the LDP for graphons to the setting where the Erdős-Rényi random graph is inhomogeneous.

Let $r \in \mathcal{W}$ be a reference graphon satisfying

$$0 < r(x,y) < 1 \qquad \text{a.e.}$$

Consider the random graph G on n vertices where $i \sim j$ with probability $r(\frac{i}{n}, \frac{j}{n})$, independently of other pairs of vertices. Write $\mathbb{P}_n$ to denote the law of G, and $\tilde{\mathbb{P}}_n$ to denote the law of $\tilde{h}^G$.

THEOREM Dhara and Sen 2020

LDP

$(\widetilde{\mathbb{P}}_n)_{n\in\mathbb{N}}$ *satisfies the large deviation principle on* $(\tilde{W}, \delta_\square)$ *with rate* $\binom{n}{2}$ *and with rate function* $J_r^*\colon \tilde{W} \to \mathbb{R}$ *equal to the lower semi-continuous envelope of the function* J_r *given by*

$$J_r(h) = \inf_{\sigma\in\Sigma} I_r(h^\sigma),$$

$$I_r(h) = \int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, \mathcal{R}\Big(h(x,y) \mid r(x,y)\Big),$$

where

$$\mathcal{R}(a \mid b) = a \log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b}$$

is the relative entropy of two Bernoulli distributions with success probabilities $a \in [0,1]$ *and* $b \in (0,1)$.

THEOREM 3: Dhara and Sen 2020

J_r^* is a good rate function, i.e., $J_r^* \not\equiv \infty$ and J_r^* is lower semi-continuous and has compact level sets with respect to the $\delta_\square$-metric.

THEOREM 4: dH & Markering 2020

If

$$r \text{ is a.e. continuous,}$$
$$\log r,\ \log(1-r) \in L^1([0,1]^2),$$

then $J_r^* = J_r$, i.e., J_r is lower semi-continuous.