

Large Graphs and Graph Limits

Dominik L. Forkert

December 15, 2015

We consider sequences of large graphs which have certain convergent graph parameters. Many important graph parameters like the edge density may be represented asymptotically as homomorphism densities. It turns out that convergence of homomorphism densities of a graph sequence gives rise to distance between graphs. However, the limit with respect to this distance need not be a finite graph any more but is instead represented by a symmetric measurable function on the unit square. These measurable functions are called graphons and induce a general class of random graphs whose numerical parameters, as long as they satisfy a Lipschitz condition, are concentrated. Moreover, we investigate the relation between spectra of large graphs and graphons.

1 Large Graphs and Their Limits

We recall that a *graph* G consists of a set of *nodes* $V(G)$ together a set of *edges* $E(G)$. Usually we will denote the number of nodes by n if there arises no ambiguity of the underlying graph. For simplicity, we will always consider non-empty and finite *simple graphs*, unless stated otherwise. In a simple graph the edges form a set, i.e. there are no multiple edges. Moreover, each edge is a pair of two distinct nodes.

A simple graph is called *complete* if every pair of distinct nodes is joined by an edge. We will make use symbols like \bullet , $\bullet\bullet$, or \blacktriangle to denote such graphs.

Now let us turn to the important notion of graph homomorphisms and their densities.

1.1 Definition Let G and H be graphs. A mapping $\varphi : V(G) \rightarrow V(H)$ is called **homomorphism** if φ preserves edge adjacency, i.e. for every edge $\{v, w\} \in E(G)$, the edge $\{\varphi(v), \varphi(w)\}$ belongs to $E(H)$.

Denote by $\text{hom}(G, H)$ the cardinality of all homomorphisms from G to H . Then the **homomorphism density** from G to H is defined by

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

Clearly, homomorphism density takes values in $[0, 1]$. Note that $t(\bullet\bullet, H)$ asymptotically represents the density of edges in H . Indeed, $\text{hom}(\bullet\bullet, H) = 2|E(H)|$. Since H may have at most $\frac{1}{2}n(n-1)$ edges, we obtain

$$t(\bullet\bullet, H) = \frac{2|E(H)|}{n^2} \sim \frac{2|E(H)|}{n(n-1)}.$$

Now let K be a complete graph. Then $\text{hom}(G, K)$ is just the number of colourings of the vertices of G with $|V(K)|$ colours such that the colours of adjacent nodes are different.

The homomorphism density as well as the edge density are two examples of *graph parameters*. A graph parameter assigns each equivalence class of isomorphic graphs a real number.

We say that $A_G \in \{0, 1\}^{V(G) \times V(G)}$ is the **adjacency matrix** of a graph G if $(A_G)_{ij}$ is the number of edges $\{i, j\}$ in $E(G)$. In the following example we consider the relation between the spectrum of the adjacency matrix and certain homomorphism densities. Since the adjacency matrix of an undirected graph is symmetric, its spectrum consists only of real eigenvalues.

1.2 Example (The spectrum of adjacency matrices) Let G be a graph and let C_k be the cycle on k nodes, i.e. each node of C_k is adjacent to exactly two other nodes. Then the eigenvalues $(\lambda_i)_{i=1}^n$ of the adjacency matrix A_G satisfy the following identity:

$$\text{hom}(C_k, G) = \text{tr}(A_G^k) = \sum_{i=1}^n \lambda_i^k.$$

In particular, $\text{hom}(C_{2k}, G)^{-2k}$ converges to the spectral radius of A_G as $(k \rightarrow \infty)$. ◀

Intuitively, above examples suggest that graphs G and H , which share equal homomorphism densities $\text{hom}(F, \cdot)$ for various graphs F , are more or less alike. This motivates the following notion of convergence on the class of finite graphs:

Let $(G_n)_{n \in \mathbb{N}}$ be a sequence of finite graphs G_n such that $\lim_{n \rightarrow \infty} V(G_n) = \infty$. Then $(G_n)_{n \in \mathbb{N}}$ is said to be **convergent** if $t(F, G_n)$ converges for every finite graph F as $(n \rightarrow \infty)$.

Convergence in above sense can be expressed with respect to a certain pseudo metric.

1.3 Definition Let two graphs H and H' be given where $n = |V(H)|$ and $n' = |V(H')|$. A **fractional overlay** of H and H' is a matrix $X \in [0, 1]^{V(H) \times V(H')}$ such that

$$\sum_{i \in V(H)} X_{ij} = \frac{1}{n'} \quad \text{and} \quad \sum_{j \in V(H')} X_{ij} = \frac{1}{n}.$$

Hence, X represents probability measure on $V(H) \times V(H')$ whose marginals are uniformly distributed. We denote the set of all fractional overlay of H and H' by $\mathcal{X}(H, H')$.

Set $\beta := A_H$ and $\beta' := A_{H'}$. Then the **cut-distance** between H and H' is defined as

$$\delta_{\square}(H, H') := \min_{X \in \mathcal{X}(H, H')} \max_{S, T \subseteq V(H) \times V(H')} \left| \sum_{\substack{(i,k) \in S \\ (j,l) \in T}} X_{ik} X_{jl} (\beta_{ij} - \beta'_{kl}) \right|.$$

Note that H and H' need not be defined on the same node set.

1.4 Proposition Let $(G_n)_{n \in \mathbb{N}}$ be a sequence of finite graphs G_n such that $\lim_{n \rightarrow \infty} V(G_n) = \infty$. Then $(G_n)_{n \in \mathbb{N}}$ is convergent precisely when $(G_n)_{n \in \mathbb{N}}$ is a Cauchy sequence with respect to δ_{\square} .

To identify a possible limit object of Cauchy sequences with respect to δ_{\square} , it is appropriate to transfer above definitions to a measure theoretic setting.

1.5 Definition We call a symmetric, Lebesgue measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ a **graphon**. From a graph-theoretic point of view, a graphon W is just certain graph on the infinite node set $[0, 1]$ with edge weights $W(x, y)$ between nodes $x, y \in [0, 1]$.

The **homomorphism density** of a graphon W with respect to a finite graph G is defined as

$$t(G, W) := \int_{[0, 1]^n} \prod_{i, j \in E(G)} W(x_i, x_j) d(x_1, \dots, x_n),$$

where as usual $n = V(G)$.

The **cut-distance** between two graphons W and U is given by

$$\delta_{\square}(W, U) := \inf_{\varphi, \psi} \sup_{S, T} \left| \int_{S \times T} W(\varphi(x), \varphi(y)) - U(\psi(x), \psi(y)) d(x, y) \right|,$$

where the infimum is taken over all measure preserving mappings $\varphi, \psi : [0, 1] \rightarrow [0, 1]$, and the supremum is taken over all Lebesgue measurable sets $S, T \subseteq [0, 1]$.

Finite graphs may be identified with graphons in the following way: The **step function** $W_G : [0, 1]^2 \rightarrow [0, 1]$ of a weighted graph G with edge weights $\beta_{ij} \geq 0$ is defined by

$$W_H(x, y) = \beta_{ij} \quad \forall (x, y) \in J_i \times J_j, 1 \leq i, j \leq n,$$

where $(J_i)_{i=1}^n$ is a equidistant partition of $[0, 1]$.

The following result justifies the choice of the definitions given above.

1.6 Proposition *Let G and H be finite graphs and let W_G and W_H be their respective step functions. Then the following statements hold:*

- (i) *The cut-distance satisfies $\delta_{\square}(W_G, W_H) = \delta_{\square}(G, H)$.*
- (ii) *The homomorphism density satisfies $t(F, W_G) = t(F, G)$ for all finite graphs F .*

In above measure theoretic framework it is possible to identify the limit object of Cauchy sequences of graphs with respect to δ_{\square} . To this end, it is reasonable to identify graphons with cut-distance 0, and denote the resulting space of equivalence classes by $\widetilde{\mathcal{W}}_0$. Now we are in the position to state the following topological result.

1.7 Theorem *$(\widetilde{\mathcal{W}}_0, \delta_{\square})$ is a compact metric space.*

In particular, this means that $\widetilde{\mathcal{W}}_0$ is complete.

2 Random Graphs and Sample Concentration

In this section we consider general random subgraphs which are induced by a finite graph or – more generally - a graphon.

2.1 Definition Let W be a graphon and fix a number $n \in \mathbb{N}$. Let $(X_i)_{i=1}^n$ be a family of independent and uniformly distributed random variables on $[0, 1]$. Denote by $\mathbb{G}(n, W)$ the random graph on the node set $(X_i)_{i=1}^n$, where $\{X_i, X_j\}$ is an edge with probability $W(X_i, X_j)$. $\mathbb{G}(n, W)$ is called a **random graph** induced by W .

For reasonably large n we expect the random graph $\mathbb{G}(n, W)$ to resemble the inducing graphon W . To this end, let us start with considering large *Erdős–Rényi graphs* which depict the main random graph model in classic graph theory.

2.2 Example (Limits of Erdős–Rényi graphs) Consider a constant graphon $W \equiv p \in [0, 1]$. Then the resulting random graph $\mathbb{G}(n, W)$ is called an *Erdős–Rényi graph (in the sense of Gilbert)* if we identify the (random) node set $V(\mathbb{G}(n, W))$ with $[n]$. Indeed, in this random graph each edge occurs independently with probability p . Thus, the probability to obtain any particular graph with m edges on the node set $[n]$ is given by

$$p^m (1 - p)^{\binom{n}{2} - m}.$$



Figure 1 As the the size of the Erdős–Rényi graph increases, its step function resembles a constant graphon if one takes some averaging into account. This intuitive *pixel picture convergence* corresponds precisely to convergence in the cut–distance.

As the number of nodes n increases, we expect $\mathbb{G}(n, W)$ to be convergent to the underlying graphon $W \equiv p$. Intuitively, this convergence may be visualized by *pixel picture convergence* as shown in Figure 1. Indeed, the following lemma shows that $\mathbb{G}(n, W)$ converges to the underlying graphon W as $(n \rightarrow \infty)$. ◀

2.3 Lemma (Sampling) For every graphon W we have

$$\mathbb{P}\left[\delta_{\square}(W, \mathbb{G}(k, W)) > \frac{22}{\sqrt{\ln k}}\right] \leq \exp\left(-\frac{k}{2 \ln k}\right) \quad \forall k \in \mathbb{N}.$$

In particular, $\mathbb{G}(k, W)$ converges in probability to W as $(k \rightarrow \infty)$.

As an alternative to convergence in the cut –distance, one may also consider convergence with respect to the homomorphism densities. More precisely, for every finite graph F the homomorphism density $t(F, \mathbb{G}(n, W))$ is close to $t(F, W)$ in high probability.

We will see that this holds in the more general context of graph parameters.

2.4 Definition A *simple graph parameter* is a real–valued mapping f on a set of simple graphs such that $f(G) = f(G')$ for every pair of isomorphic graphs G and G' .

A simple graph parameter f is called **reasonably smooth** if for every pair of graphs G and H on the same node set such that the symmetric difference $E(G) \Delta E(H)$ is incident with one common node or empty, we have

$$|f(G) - f(H)| \leq 1. \tag{1}$$

2.5 Theorem (Sample concentration) Let f be a reasonably smooth graph parameter, let W be a graphon and fix $k \in \mathbb{N}$. Then the expectation $f_0 := \mathbb{E}(f(\mathbb{G}(k, W)))$ satisfies

$$\mathbb{P}\left[f(\mathbb{G}(k, W)) \geq f_0 + \sqrt{2tk}\right] \leq e^{-t} \quad \forall t \geq 0. \tag{2}$$

Amongst several approaches (cf. section 6.1 and 8.1 in [3]) we will prove this theorem by means of martingale theory. To this end, recall that a *discrete–time martingale* is a sequence of random variables $(X_k)_{k \in \mathbb{N}_0}$ on a probability space such that all expectations $\mathbb{E}(X_k)$ are finite and the conditional expectations satisfy

$$\mathbb{E}(X_{k+1} | X_1, \dots, X_k) = X_k \quad \forall k \in \mathbb{N}_0.$$

For the proof of Theorem 2.5 we first need the following fundamental inequality for martingales.

2.6 Lemma (Azuma's inequality) *Let $(X_k)_{k \in \mathbb{N}_0}$ be a discrete-time martingale. Supposed that the increments satisfy $|X_k - X_{k-1}| \leq c_k$ almost surely for some constants $c_k > 0$, we have*

$$\mathbb{P}[X_N - X_0 \geq t] \leq \exp\left(-\frac{t^2}{2 \sum_{k=1}^N c_k^2}\right) \quad \forall N \in \mathbb{N}. \quad (3)$$

Proof of Theorem 2.5 Denote by $(\mathbb{G}_i)_{i=0}^k$ a sequence of subgraphs of $\mathbb{G}(k, W)$ with nodes $V(\mathbb{G}_i) = V(\mathbb{G}(k, W))$ and edges $E(\mathbb{G}_i)$ concentrated on at most i nodes in such a way that

$$\emptyset = E(\mathbb{G}_0) \subseteq E(\mathbb{G}_1) \subseteq E(\mathbb{G}_2) \subseteq \dots \subseteq E(\mathbb{G}_k) = E(\mathbb{G}(k, W)).$$

For a given graph parameter f , we define random variables

$$Z_i := \mathbb{E}(f(\mathbb{G}(k, W)) | \mathbb{G}_i) \quad \forall i \leq k.$$

To check that $(Z_i)_{i=0}^k$ is a martingale, we use the tower property of the conditional expectation to compute

$$\mathbb{E}(Z_i | G_{i-1}) = \mathbb{E}(\mathbb{E}(f(\mathbb{G}(k, W)) | \mathbb{G}_i) | G_{i-1}) = \mathbb{E}(f(\mathbb{G}(k, W)) | G_{i-1}) = Z_{i-1}.$$

Now the martingale property follows by induction.

Since f is reasonably smooth, we have the following bound for the increments:

$$|Z_i - Z_{i-1}| \leq 1 \quad \forall i \leq k.$$

Thus, we can apply Azuma's inequality to obtain the estimate

$$\mathbb{P}[Z_k - Z_0 \geq t] \leq \exp\left(-\frac{t^2}{2k}\right).$$

Finally it remains to note that $Z_0 = \mathbb{E}(f(\mathbb{G}(k, W)))$ and $Z_k = f(\mathbb{G}(k, W))$. ■

In particular, applying above theorem to the reasonably smooth graph parameters

$$f_{\pm}(G) := \pm \frac{V(G)}{V(F)} t(F, G)$$

implies the following concentration bound for the homomorphism density.

2.7 Corollary *For every finite simple graph F we have*

$$\mathbb{P}[|t(F, \mathbb{G}(k, W)) - t(F, W)| \geq \varepsilon] \leq 2 \exp\left(-\frac{\varepsilon^2 k}{8|V(F)|^2}\right) \quad \forall \varepsilon \in (0, 1), \forall k \in \mathbb{N}. \quad (4)$$

3 Spectral Graph Theory of Random Graphs

In Example 1.2 we considered the spectrum of the adjacency matrix A_G with respect to a simple graph G . For graphons the analogue of the adjacency matrix is given by the corresponding integral operator.

3.1 Definition Let a graphon W be given. Then the corresponding **Hilbert-Schmidt integral operator** $T_W : L^2([0, 1]) \rightarrow L^2([0, 1])$ is defined by

$$(T_W f)(x) := \int_0^1 W(x, y) f(y) dy \quad \forall x \in [0, 1].$$

Note that the operator T_W is compact as well as self-adjoint. Therefore, the spectrum $\sigma(T_W)$ consists of an at most countable subset of \mathbb{R} , where 0 either belongs to $\sigma(T_W)$ or is the only cluster point of the spectrum.

The following result shows that graph/graphon convergence with respect to the cut-distance preserves convergence of the eigenvalues.

3.2 Proposition *Let $(W_i)_{i \in \mathbb{N}}$ be a sequence of graphons, converging with respect to δ_{\square} to a limit graphon W . Denote by $(\lambda_k^i)_{k \in \mathbb{N}}$ the spectrum of W_i , taking multiplicity of the eigenvalues into account and ordered with respect to \geq for all $i \in \mathbb{N}$. In a similar way, denote by $(\lambda_k)_{k \in \mathbb{N}}$ the ordered spectrum of W . Then we have $\lim_{i \rightarrow \infty} \lambda_k^i = \lambda_k$ for all $k \in \mathbb{N}$.*

The result holds as well when all sequences are ordered with respect to \leq instead.

Note that for a simple graph G , the rescaled spectrum $\frac{1}{n}\sigma(A_G)$ of the adjacency matrix A_G together with 0 forms the spectrum of the corresponding Hilbert–Schmidt integral operator T_W . As a result, we obtain the following corollary of Proposition 3.2.

3.3 Corollary *Let $(G_i)_{i \in \mathbb{N}}$ be a sequence of finite simple graphs, converging with respect to δ_{\square} to a limit graphon W . Denote by $\lambda_1^i \geq \lambda_2^i \geq \lambda_3^i \geq \dots$ the spectrum of $\frac{1}{|V(G_i)|} A_{G_i}$, taking multiplicity of the eigenvalues into account for all $i \in \mathbb{N}$. In a similar way, denote by $(\lambda_k)_{k \in \mathbb{N}}$ the ordered spectrum of W . Then we have $\lim_{i \rightarrow \infty} \lambda_k^i = \lambda_k$ for all $k \in \mathbb{N}$.*

The result holds as well when all sequences are ordered with respect to \leq instead.

4 Bibliography

- [1] C. Borgs *et al.*, *Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing*, *Advances in Mathematics* **219** (2008), no. 6, 1801–1851
- [2] C. Borgs *et al.*, *Convergent sequences of dense graphs ii. multiway cuts and statistical physics*, *Annals of Mathematics* **176** (2012), no. 1, 151–219
- [3] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities: A nonasymptotic theory of independence*. (OUP Oxford, 2013).
- [4] A. E. Brouwer and W. H. Haemers, *Spectra of graphs*. (Springer Science & Business Media, 2011).
- [5] D. Glasscock, *A graphon?*, *Notices of the AMS* **62** (2015), no. 1
- [6] L. Lovász, *Very large graphs*, arXiv preprint arXiv:0902.0132 (2009)
- [7] L. Lovász, *Large networks and graph limits*, volume 60. (American Mathematical Soc., 2012).
- [8] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, *Journal of Combinatorial Theory, Series B* **96** (2006), no. 6, 933–957
- [9] L. Lovász and B. Szegedy, *Testing properties of graphs and functions*, *Israel Journal of Mathematics* **178** (2010), no. 1, 113–156