Combinatorics, Probability and Computing (2013) 22, 517–526.
doi:10.1017/S0963548313000175

The Entropy of Random-Free Graphons and Properties

HAMED HATAMI[1†] and SERGUEI NORINE[2‡]

[1]School of Computer Science, McGill University, Montreal, Canada
(e-mail: hatami@cs.mcgill.ca)
[2]Department of Mathematics & Statistics, McGill University, Montreal, Canada
(e-mail: snorin@math.mcgill.ca)

Received 17 July 2012; revised 11 April 2013; first published online 16 May 2013

Every graphon defines a random graph on any given number n of vertices. It was known that the graphon is random-free if and only if the entropy of this random graph is subquadratic. We prove that for random-free graphons, this entropy can grow as fast as any subquadratic function. However, if the graphon belongs to the closure of a random-free hereditary graph property, then the entropy is $O(n \log n)$. We also give a simple construction of a non-step-function random-free graphon for which this entropy is linear, refuting a conjecture of Janson.

2010 *Mathematics subject classification*: 05C99

1. Introduction

In recent years a theory of convergent sequences of dense graphs has been developed. One can construct a limit object for such a sequence in the form of certain symmetric measurable functions called graphons. Every graphon defines a random graph on any given number of vertices. In [5] several facts about the asymptotics of the entropies of these random variables are established. These results provide good understanding of the situation when the graphon is not 'random-free', but they say essentially nothing about random-free graphons. The purpose of this article is to study these entropies in the case of random-free graphons.

1.1. Preliminaries

For every natural number n, denote $[n] := \{1, \ldots, n\}$. In this paper all graphs are simple and finite. For a graph G, let $V(G)$ and $E(G)$, respectively, denote the set of the vertices

† Research supported by an NSERC and an FQRNT grant.
‡ Research supported by an NSERC Discovery grant.

and the edges of G. Let $\mathcal{U}$ denote set of all graphs up to isomorphism. Moreover, for $n \geqslant 0$, let $\mathcal{U}_n \subset \mathcal{U}$ denote the set of all graphs in $\mathcal{U}$ with exactly n vertices. We will usually work with labelled graphs. For every $n \geqslant 1$, denote by $\mathcal{L}_n$ the set of all graphs with vertex set $[n]$.

The *homomorphism density* of a graph H in a graph G, denoted by $t(H;G)$, is the probability that a random mapping $\phi : V(H) \to V(G)$ preserves adjacencies, *i.e.*, $uv \in E(H)$ implies $\phi(u)\phi(v) \in E(G)$. The *induced density* of a graph H in a graph G, denoted by $p(H;G)$, is the probability that a random *embedding* of the vertices of H in the vertices of G is an embedding of H in G.

We call a sequence of finite graphs $(G_n)_{n=1}^{\infty}$ *convergent* if, for every finite graph H, the sequence $\{p(H;G_n)\}_{n=1}^{\infty}$ converges. It is not difficult to construct convergent sequences $(G_n)_{n=1}^{\infty}$ whose limits cannot be recognized as graphs, *i.e.*, there is no graph G with $\lim_{n\to\infty} p(H;G_n) = p(H;G)$ for every H. Thus naturally one considers $\overline{\mathcal{U}}$, the completion of $\mathcal{U}$ under this notion of convergence. It is not hard to see that $\overline{\mathcal{U}}$ is a compact metrizable space which contains $\mathcal{U}$ as a dense subset. The elements of the complement $\mathcal{U}^{\infty} := \overline{\mathcal{U}} \setminus \mathcal{U}$ are called *graph limits*. Trivially, a graph limit Γ is uniquely determined by the numbers $p(H;\Gamma)$ for all $H \in \mathcal{U}$.

Note that a sequence of graphs $(G_n)_{n=1}^{\infty}$ with $|V(G_n)| \to \infty$ cannot converge to a finite graph G, as $p(H;G) = 0$ for every graph H with $|V(H)| > |V(G)|$. Hence a sequence of graphs $(G_n)_{n=1}^{\infty}$ converges to a *graph limit* if and only if $|V(G_n)| \to \infty$ and $p(H;G_n)$ converges for every graph H.

It is shown in [7] that every graph limit Γ can be represented by a *graphon*, which is a symmetric measurable function $W : [0,1]^2 \to [0,1]$. The set of all graphons is denoted by $\mathcal{W}_0$. Given a graph G with vertex set $[n]$, we define the corresponding graphon $W_G : [0,1]^2 \to \{0,1\}$ as follows. Let A_G denote the adjacency matrix of G. Then $W_G(x,y) := A_G(\lceil xn \rceil, \lceil yn \rceil)$ if $x, y \in (0,1]$, and if $x = 0$ or $y = 0$, set W_G to 0. It is easy to see that if $(G_n)_{n=1}^{\infty}$ is a graph sequence that converges to a graph limit Γ, then for every graph H,

$$p(H;\Gamma) = \lim_{n\to\infty} \mathbb{E}\Bigg[\prod_{uv\in E(H)} W_{G_n}(x_u, x_v) \prod_{uv\in E(H^c)} (1 - W_{G_n}(x_u, x_v))\Bigg],$$

where $\{x_u\}_{u\in V(H)}$ are independent random variables taking values in $[0,1]$ uniformly, and $E(H^c) = \{uv : u \neq v, uv \notin E(H)\}$. Lovász and Szegedy [7] showed that for every graph limit Γ, there exists a graphon W such that, for every graph H, we have $p(H;\Gamma) = p(H;W)$, where

$$p(H;W) := \mathbb{E}\Bigg[\prod_{uv\in E(H)} W(x_u, x_v) \prod_{uv\in E(H^c)} (1 - W(x_u, x_v))\Bigg].$$

Furthermore, this graphon is unique in the following sense. For a measurable $\sigma : [0,1] \to [0,1]$, define $W \circ \sigma : [0,1]^2 \to [0,1]$ as $W \circ \sigma : (x,y) \mapsto W(\sigma(x), \sigma(y))$. Now if W_1 and W_2 are two different graphons representing the same graph limit, then there exists a graphon W and two measure-preserving maps $\sigma_1, \sigma_2 : [0,1] \to [0,1]$ such that

$$W_1 = W \circ \sigma_1 \quad \text{and} \quad W_2 = W \circ \sigma_2, \tag{1.1}$$

almost everywhere [4]. With these considerations, sometimes we shall not distinguish between the graph limits and their corresponding graphons. We define the δ_1 distance of two graphons W_1 and W_2 as

$$\delta_1(W_1, W_2) = \inf \| W_1 - W_2 \circ \sigma \|_1,$$

where the infimum is over all measure-preserving maps $\sigma : [0,1] \to [0,1]$.

A graphon W is called a *step-function* if there is a partition of $[0,1]$ into a finite number of measurable sets $S_1, \dots, S_n$ so that W is constant on every $S_i \times S_j$. The partition classes will be called the *steps* of W.

Let W be a graphon and $x_1, \dots, x_n \in [0,1]$. The random graph $G(x_1, \dots, x_n, W) \in \mathcal{L}_n$ is obtained by including the edge ij with probability $W(x_i, x_j)$, independently for all pairs (i,j) with $1 \leqslant i < j \leqslant n$. By picking $x_1, \dots, x_n$ independently and uniformly at random from $[0,1]$, we obtain the random graph $G(n, W) \in \mathcal{L}_n$. Note that for every $H \in \mathcal{L}_n$,

$$\mathbb{P}[G(n, W) = H] = p(H; W).$$

1.2. Graph properties and entropy

A subset of the set $\mathcal{U}$ is called a *graph class*. Similarly, a *graph property* is a property of graphs that is invariant under graph isomorphisms. There is an obvious one-to-one correspondence between graph classes and graph properties and we will not distinguish between a graph property and the corresponding class. A graph class or property $\mathcal{Q}$ is *hereditary* if, whenever a graph G has the property $\mathcal{Q}$, every induced subgraph of G also has $\mathcal{Q}$.

Let $\mathcal{Q} \subseteq \mathcal{U}$ be a graph class. For every $n > 1$, we denote by $\mathcal{Q}_n$ the set of graphs in $\mathcal{Q}$ with exactly n vertices. We let $\overline{\mathcal{Q}} \subseteq \overline{\mathcal{U}}$ be the closure of $\mathcal{Q}$ in $\overline{\mathcal{U}}$.

Define the *binary entropy* function $h : [0,1] \mapsto \mathbb{R}_+$ as $h(x) = -x \log x - (1-x) \log(1-x)$ for $x \in (0,1)$ and $h(0) = h(1) = 0$ so that h is continuous on $[0,1]$, where, here and throughout the paper, $\log(\cdot)$ denotes the logarithm to base 2. The *entropy* of a graphon W is defined by

$$\mathrm{Ent}(W) := \int_0^1 \int_0^1 h(W(x,y))\, dx\, dy.$$

Note that it follows from the uniqueness result (1.1) that entropy is a function of the underlying graph limit, and it does not depend on the choice of the graphon representing it. It is shown in [1] and Theorem D.5 of [6] that

$$\lim_{n \to \infty} \frac{\mathrm{Ent}(G(n, W))}{\binom{n}{2}} = \mathrm{Ent}(W), \tag{1.2}$$

where $\mathrm{Ent}(G(n, W))$ is the usual entropy of the random variable $G(n, W)$.

A graphon is called *random-free* if it is $\{0,1\}$-valued almost everywhere. Note that a graphon W is random-free if and only if $\mathrm{Ent}(W) = 0$, which by (1.2) is equivalent to $\mathrm{Ent}(G(n, W)) = o(n^2)$. Our first theorem shows that this is sharp in the sense that the growth of $\mathrm{Ent}(G(n, W))$ for random-free graphons W can be arbitrarily close to quadratic.

Theorem 1.1. *Let $\alpha : \mathbb{N} \to \mathbb{R}_+$ be a function with $\lim_{n\to\infty} \alpha(n) = 0$. Then there exists a random-free graphon W such that* $\mathrm{Ent}(G(n, W)) = \Omega(\alpha(n)n^2)$.

A graph property $\mathcal{Q}$ is called *random-free* if every $W \in \overline{\mathcal{Q}}$ is random-free. Our next theorem shows that in contrast to Theorem 1.1, when a graphon W is the limit of a sequence of graphs with a random-free *hereditary* property, then $\mathrm{Ent}(G(n, W))$ cannot grow faster than $O(n \log n)$.

Theorem 1.2. *Let $\mathcal{Q}$ be a random-free* hereditary *property, and let W be the limit of a sequence of graphs in $\mathcal{Q}$. Then* $\mathrm{Ent}(G(n, W)) = O(n \log n)$.

Remark. We defined $G(n, W)$ as a *labelled* graph in $\mathcal{L}_n$. Both Theorems 1.1 and 1.2 remain valid if we consider the random variable $G_u(n, W)$ taking values in $\mathcal{U}_n$ obtained from $G(n, W)$ by forgetting the labels. Indeed, $\mathrm{Ent}(G_u(n, W)) = \mathrm{Ent}(G(n, W)) - \mathrm{Ent}(G(n, W) \mid G_u(n, W))$ and $\mathrm{Ent}(G(n, W) \mid G_u(n, W) = H) = O(n \log n)$ for every $H \in \mathcal{U}_n$. It follows that

$$\mathrm{Ent}(G(n, W)) - O(n \log n) \leqslant \mathrm{Ent}(G_u(n, W)) \leqslant \mathrm{Ent}(G(n, W)).$$

2. Proof of Theorem 1.1

A *bigraph* is a bipartite graph with a specified bipartition. For every positive integer m, let F_m denote the unique bigraph $([m], [2^m], E)$ with the property that the vertices in $[2^m]$ all have different sets of neighbours. The *transversal-uniform graph* G_U is the unique graph (up to an isomorphism) with vertex set $\mathbb{N}$ which satisfies the following property. The vertices are partitioned into sets $\{A_i\}_{i=1}^{\infty}$ with $\log |A_i| = \sum_{j=1}^{i-1} |A_{i-1}|$. There are no edges inside the A_i, and for every i, the bigraph induced by $(\cup_{j=1}^{i-1} A_j, A_i)$ is isomorphic to $F_{\sum_{j=1}^{i-1} |A_j|}$.

Let $\mathcal{I} = \{I_i\}_{i\in\mathbb{N}}$ be a partition of $[0, 1]$ into intervals. We define its corresponding *transversal-uniform graphon* $W_{\mathcal{I}}$ by assigning weights $|I_i|/|A_i|$ to all the vertices in A_i in the transversal-uniform graph G_U described above. More precisely, we partition each I_i into $|A_i|$ equal size intervals (corresponding to the elements in A_i), and mapping all the points in each of these subintervals to its corresponding vertex in A_i. This measurable surjection $\pi_{\mathcal{I}} : [0, 1] \to \mathbb{N}$, together with the transversal-uniform graph described above, defines the transversal-uniform graphon $W_{\mathcal{I}}$ by setting

$$W_{\mathcal{I}}(x, y) = \begin{cases} 1 & \text{if } \pi_{\mathcal{I}}(x)\pi_{\mathcal{I}}(y) \in E(G_U), \\ 0 & \text{if } \pi_{\mathcal{I}}(x)\pi_{\mathcal{I}}(y) \notin E(G_U). \end{cases}$$

Note that by construction $W_{\mathcal{I}}$ has the following property. Let $s < k$ be positive integers, and $x_1, \ldots, x_s \in \cup_{i<k} I_i$ belong to pairwise distinct intervals in $\mathcal{I}$. For every $f : [s] \to \{0, 1\}$, we have

$$\mathbb{P}[\forall i,\ W_{\mathcal{I}}(x_i, y) = f(i) \mid y \in I_k] = \frac{1}{2^s},$$

https://doi.org/10.1017/S0963548313000175 Published online by Cambridge University Press

where y is a random variable taking values uniformly in $[0,1]$. It follows that for every graph H on s vertices,

$$\mathbb{P}[G(x_1,\dots,x_s,W_{\mathcal{I}}) = H \mid \forall i,\ x_i \in I_{k_i}] = \frac{1}{2^{\binom{s}{2}}}, \tag{2.3}$$

where $x_1,\dots,x_s$ are now i.i.d. random variables taking values uniformly in $[0,1]$, and $k_1,\dots,k_s$ are distinct natural numbers.

We translate (2.3) into a lower bound on (conditional) entropy of transversal-uniform graphons. First we need a simple lemma.

Lemma 2.1. *Let $W_{\mathcal{I}}$ be a transversal-uniform graphon, and let $\phi : [n] \to [0,1]$ be a uniformly random map. For every $\rho : [n] \to \mathbb{N}$, we have*

$$\mathrm{Ent}(G(\phi(1),\dots,\phi(n),W_{\mathcal{I}}) \mid \pi_{\mathcal{I}} \circ \phi = \rho) \geqslant \binom{|\mathrm{Im}(\rho)|}{2}.$$

Proof. Pick a set of representatives $K \subseteq [n]$ so that $\rho|_K : K \to \mathrm{Im}(\rho)$ is a bijection. Equation (2.3) implies that for every graph H with $V(H) = K$,

$$\mathbb{P}[G(\phi(1),\dots,\phi(n),W_{\mathcal{I}})[K] = H \mid \pi_{\mathcal{I}} \circ \phi = \rho] = \frac{1}{2^{\binom{|\mathrm{Im}(\rho)|}{2}}}.$$

Therefore,

$$\begin{aligned}\mathrm{Ent}(G(\phi(1),\dots,\phi(n),W_{\mathcal{I}}) \mid \pi_{\mathcal{I}} \circ \phi = \rho) &\geqslant \mathrm{Ent}(G(\phi(1),\dots,\phi(n),W_{\mathcal{I}})[K] \mid \pi_{\mathcal{I}} \circ \phi = \rho)\\ &= \binom{|\mathrm{Im}(\rho)|}{2}.\end{aligned}$$

□

In the proof of Theorem 1.1 below we will make use of the following well-known inequality about conditional entropy. For discrete random variables X and Y,

$$\mathrm{Ent}(X \mid Y) := \sum_{y \in \mathrm{supp}(Y)} \mathbb{P}[Y = y]\mathrm{Ent}(X \mid Y = y) \leqslant \mathrm{Ent}(X). \tag{2.4}$$

Proof of Theorem 1.1. For every positive integer k, define

$$g_k := \max\{\{2^{k+5}\} \cup \{n \mid \alpha(n) > 2^{-2k-9}\}\}.$$

The numbers g_k are well-defined, as the condition $\lim_{n\to\infty} \alpha(n) = 0$ implies that the set $\{n \mid \alpha(n) > 2^{-2k-9}\}$ is finite. Define the sums $G_k := \sum_{i=1}^{k} g_k$, and set $\beta_i = \frac{1}{g_k 2^k}$ for all the g_k indices $i \in (G_{k-1}, G_k]$. Let $\mathcal{I} = \{I_i\}_{i\in\mathbb{N}}$ be a partition of $[0,1]$ into intervals with $|I_i| = \beta_i$, and let $W_{\mathcal{I}}$ be the corresponding transversal-uniform graphon.

Consider a sufficiently large $n \in \mathbb{N}$, and let $k \in \mathbb{N}$ be chosen to be the maximal k such that $2^{k+4} \leqslant n$ and $\alpha(n) \leqslant 2^{-2k-7}$. We have $n < 2^{k+5}$ or $\alpha(n) > 2^{-2k-9}$. Therefore $n \leqslant g_k$ by the definition of g_k. Let $\phi : [n] \to [0,1]$ be random and uniform. By Lemma 2.1, for any fixed $\rho : [n] \to \mathbb{N}$, we have

$$\mathrm{Ent}(G(\phi(1),\dots,\phi(n),W_{\mathcal{I}}) | \pi_{\mathcal{I}} \circ \phi = \rho) \geqslant \binom{|\mathrm{Im}(\rho)|}{2}.$$

https://doi.org/10.1017/S0963548313000175 Published online by Cambridge University Press

Thus

$$\mathrm{Ent}(G(n,W_{\mathcal{I}})) \geqslant \mathrm{Ent}(G(n,W_{\mathcal{I}})|\pi_{\mathcal{I}}\circ\phi) \geqslant \mathbb{P}\big[|\mathrm{Im}(\pi_{\mathcal{I}}\circ\phi)| \geqslant n2^{-k-2}\big]\binom{n2^{-k-2}}{2}. \tag{2.5}$$

Define the random variable $X := |\mathrm{Im}(\pi_{\mathcal{I}}\circ\phi)\cap(G_{k-1},G_k]| \leqslant |\mathrm{Im}(\pi_{\mathcal{I}}\circ\phi)|$. We have

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i\in(G_{k-1},G_k]} \mathbb{P}[\phi^{-1}(I_i)\neq\emptyset] = \sum_{i\in(G_{k-1},G_k]} (1-(1-\beta_i)^n) \\ &= g_k\left(1-\left(1-\frac{1}{g_k 2^k}\right)^n\right) \geqslant n2^{-k-1},\end{aligned}$$

where we used the fact that $g_k 2^k \geqslant 2n$ and that $(1-x)^n \leqslant 1-nx+n^2x^2 \leqslant 1-nx/2$ for $x\in[0,1/2n]$. As the events $\phi^{-1}(I_i)\neq\emptyset$ and $\phi^{-1}(I_j)\neq\emptyset$ are negatively correlated for $i\neq j$, we have $\mathrm{Var}[X]\leqslant\mathbb{E}[X]$. Hence by Chebyshev's inequality

$$\begin{aligned}\mathbb{P}\big[|\mathrm{Im}(\pi_{\mathcal{I}}\circ\phi)| \geqslant n2^{-k-2}\big] &\geqslant \mathbb{P}\big[X \geqslant n2^{-k-2}\big] \geqslant 1-\mathbb{P}\left[|X-\mathbb{E}[X]| \geqslant \frac{\mathbb{E}[X]}{2}\right] \\ &\geqslant 1-\frac{4\mathrm{Var}[X]}{\mathbb{E}[X]^2} \geqslant 1-\frac{4}{n2^{-k-1}} \geqslant \frac{1}{2}.\end{aligned}$$

Substituting in (2.5), we obtain

$$\mathrm{Ent}(G(n,W_{\mathcal{I}})) \geqslant \frac{1}{2}\binom{n2^{-k-2}}{2} \geqslant n^2 2^{-2k-7} \geqslant \alpha(n)n^2,$$

as desired. □

3. Proof of Theorem 1.2

Lovász and Szegedy [8] obtained a combinatorial characterization of random-free hereditary graph properties. To state this result it is convenient to distinguish between bipartite graphs and bigraphs. A *bipartite* graph is a graph (V,E) whose node set has a partition into two classes such that all edges connect nodes in different classes. A *bigraph* is a triple (U_1,U_2,E) where U_1 and U_2 are finite sets and $E\subseteq U_1\times U_2$. So a bipartite graph becomes a bigraph if we fix a bipartition and specify which bipartition class is first and second. On the other hand, if $F=(V,E)$ is a graph, then (V,V,E') is an associated bigraph, where $E'=\{(x,y): xy\in E\}$.

If $G=(V,E)$ is a graph, then an induced sub-bigraph of G is determined by two (not necessarily disjoint) subsets $S,T\subseteq V$, and its edge set consists of those pairs $(x,y)\in S\times T$ for which $xy\in E$ (so this is an induced subgraph of the bigraph associated with G).

For a bigraph $H=(U_1,U_2,E)$ and a graphon W, analogous to the definition of the induced density of a graph in a graphon, we define

$$p^{\mathsf{b}}(H;W) = \mathbb{E}\left[\prod_{\substack{u\in U_1,\, v\in U_2\\ uv\in E}} W(x_u,y_v) \prod_{\substack{u\in U_1,\, v\in U_2\\ uv\in(U_1\times U_2)\setminus E}} (1-W(x_u,y_v))\right],$$

where $\{x_u\}_{u\in U_1}, \{y_v\}_{v\in U_2}$ are independent random variables taking values in $[0,1]$ uniformly. Now we are ready to state Lovász and Szegedy's characterization of random-free graph properties.

Theorem 3.1 ([8]). *A hereditary graph property $\mathcal{Q}$ is random-free if and only if there exists a bigraph H such that $p^{\mathrm{b}}(H;W)=0$ for all $W\in\overline{\mathcal{Q}}$.*

The following lemma is due to Alon, Fischer and Newman (see [2, Lemma 1.6]).

Lemma 3.2 ([2]). *There is an absolute constant C for which the following is true. Let k be a positive integer and let $\delta>0$ be a small real. For every graph G, either there exists a step-function graphon W' with $r\leqslant\left(\frac{k}{\delta}\right)^{Ck}$ steps such that $\delta_1(W_G,W')\leqslant\delta$, or for every bigraph H on k vertices $p^b(H;W)\geqslant\left(\frac{\delta}{k}\right)^{Ck^2}$.*

Every random-free graphon W can be approximated arbitrarily well in the δ_1 distance with W_G for some graph G, and furthermore, for every fixed H, the function $p^{\mathrm{b}}(H,\cdot)$ is continuous in the δ_1 distance. Thus Lemma 3.2 can be generalized to random-free graphons.

Corollary 3.3. *There is an absolute constant C for which the following is true. Let k be a positive integer and let $\delta>0$ be a small real. For every random-free graphon W, either there exists a step-function graphon W' with $r\leqslant\left(\frac{k}{\delta}\right)^{Ck}$ steps such that $\delta_1(W,W')\leqslant\delta$, or for every bigraph H on k vertices $p^b(H;W)\geqslant\left(\frac{\delta}{k}\right)^{Ck^2}$.*

Next we will prove two simple lemmas about entropy.

Lemma 3.4. *Let μ_1 and μ_2 be two discrete probability distributions on a finite set Ω. Then*

$$|\mathrm{Ent}(\mu_1)-\mathrm{Ent}(\mu_2)|\leqslant|\Omega|h\left(\frac{\|\mu_1-\mu_2\|_1}{|\Omega|}\right).$$

Proof. Define $0\log 0:=\lim_{x\to 0}x\log x=0$. By taking the derivative with respect to x, for fixed d we see that $(x+d)\log(x+d)-x\log x$ is monotone for $0\leqslant x\leqslant 1-d$. Therefore, for $x_1,x_2\in[0,1]$ we have

$$\begin{aligned}|x_2\log x_2-x_1\log x_1|&\leqslant\max\{-|x_2-x_1|\log|x_2-x_1|,-(1-|x_2-x_1|)\log(1-|x_2-x_1|)\}\\&\leqslant h(|x_2-x_1|).\end{aligned}$$

Thus

$$\begin{aligned}|\mathrm{Ent}(\mu_1)-\mathrm{Ent}(\mu_2)|&=\left|\sum_{x\in\Omega}\mu_1(x)\log\mu_1(x)-\mu_2(x)\log\mu_2(x)\right|\\&\leqslant\sum_{x\in\Omega}h(|\mu_1(x)-\mu_2(x)|)\leqslant|\Omega|h\left(\frac{\|\mu_1-\mu_2\|_1}{|\Omega|}\right),\end{aligned}$$

where the last inequality is by concavity of the binary entropy function h. □

https://doi.org/10.1017/S0963548313000175 Published online by Cambridge University Press

Lemma 3.5. *Let W_1 and W_2 be two graphons, and let μ_1 and μ_2 be the probability distributions on $\mathcal{L}_n$ induced by $G(n,W_1)$ and $G(n,W_2)$, respectively. Then*

$$\|\mu_1-\mu_2\|_1 \leqslant n^2\delta_1(W_1,W_2).$$

Proof. Note that without loss of generality W_1 and W_2 can be replaced by equivalent graphons so that $\delta_1(W_1,W_2)=\|W_1-W_2\|_1$. Let $x_1,\dots,x_n$ be i.i.d. uniform random variables with values in $[0,1]$. Let $\{\epsilon_{ij}:1\leqslant i<j\leqslant n\}$ be independent random variables taking values in $[0,1]$ uniformly. Let G_1,G_2 be random graphs on $[n]$ defined in the following way. There is an edge between two vertices $i<j$ in G_k if $W_k(x_i,x_j)\geqslant \epsilon_{ij}$ for $k=1,2$. Note that G_1 and G_2, respectively, have the same distributions as $G(n,W_1)$ and $G(n,W_2)$. Thus

$$\|\mu_1-\mu_2\|_1 \leqslant 2\mathbb{P}[G_1\neq G_2] \leqslant \mathbb{E}\Bigg[\sum_{i\neq j}|W_1(x_i,x_j)-W_2(x_i,x_j)|\Bigg] \leqslant n^2\|W_1-W_2\|_1. \qquad \square$$

Proof of Theorem 1.2. Since $\mathcal{Q}$ is random-free, by Theorem 3.1, there exists a bigraph H such that $p^{\mathrm{b}}(H;W)=0$ for all $W\in\overline{\mathcal{Q}}$. Applying Corollary 3.3 with $\delta=1/n^5$ shows that there exists a step-function graphon W' with $n^{O(1)}$ steps satisfying $\|W-W'\|_1\leqslant\delta$. Then, since $|\mathcal{L}_n|\leqslant 2^{n^2}$, Lemmas 3.4 and 3.5 imply

$$\begin{aligned}|\mathrm{Ent}(G(n,W'))-\mathrm{Ent}(G(n,W))| &\leqslant 2^{n^2}h\left(\frac{n^2\delta}{2^{n^2}}\right)\\ &= -2^{n^2}\left(\frac{n^2\delta}{2^{n^2}}\log\left(\frac{n^2\delta}{2^{n^2}}\right)+\left(1-\frac{n^2\delta}{2^{n^2}}\right)\log\left(1-\frac{n^2\delta}{2^{n^2}}\right)\right)\\ &\leqslant n^4\delta+n^2\delta(-2\log n-\log\delta)+2^{n^2}\cdot 2\frac{n^2\delta}{2^{n^2}}=o(1).\end{aligned}$$

Since W' is random-free and it has $n^{O(1)}$ steps, $|\mathrm{supp}(G(n,W'))|=n^{O(n)}$. Consequently $\mathrm{Ent}(G(n,W'))=O(n\log n)$. $\square$

4. Concluding remarks

(1) Note that if W is a random-free step-function, then $\mathrm{Ent}(G(n,W))=O(n)$. In [6] it is conjectured that the converse is also true. That is, $\mathrm{Ent}(G(n,W))=O(n)$ if and only if W is equivalent to a random-free step-function. The following simple example disproves this conjecture.

Let μ be the probability distribution on $\mathbb{N}$ defined by $\mu(\{i\})=2^{-i}$. Consider the random variable $X=(X_1,\dots,X_n)\in\mathbb{N}^n$, where X_i are i.i.d. random variables with distribution μ. We have $\mathrm{Ent}(X_i)=\sum_{i=1}^{\infty}2^{-i}i=2$. Hence $\mathrm{Ent}(X)=\sum\mathrm{Ent}(X_i)=2n$.

Partition $[0,1]$ into intervals $\{I_i\}_{i=1}^{\infty}$, where $|I_i|=2^{-i}$. Let W be the graphon that is constant 1 on $\cup_{i=1}^{\infty}I_i\times I_i$ and 0 everywhere else. Note that

$$\mathrm{Ent}(G(n,W))\leqslant\mathrm{Ent}(X)\leqslant 2n.$$

Therefore $G(n,W)$ has linear entropy.

It remains to verify that W is not equivalent to a step-function. This follows immediately from the fact that W has infinite rank as a kernel. It can also be verified in a more combinatorial way. A *homogenous set* of vertices in a graph H is a set of vertices which are either all pairwise adjacent to each other, or all pairwise non-adjacent. If W is equivalent to a step-function with k steps, then every $H \in \mathrm{supp}(G(n, W))$ clearly contains a homogenous set of size at least n/k. On the other hand, if $H \in \mathcal{L}_{n^2}$ is a disjoint union of n complete graphs on n vertices, then the largest homogenous set in H has size n, but $H \in \mathrm{supp}(G(n^2, W))$ by construction.

(2) Theorem 1.2 shows that when W is a limit of a random-free hereditary property, then the entropy of $G(n, W)$ is small. However, the support of $G(n, W)$ can be comparatively large. For every $\epsilon > 0$, we construct examples for which $\log(|\mathrm{supp}(G(n, W))|) = \Omega(n^{2-\epsilon})$. Note that Theorem 1.2 implies that $G(n, W)$ is far from being uniform on the support in these examples, as the entropy of a uniform random variable with support of size $2^{\Omega(n^{2-\epsilon})}$ is $\Omega(n^{2-\epsilon})$.

Let us now describe the construction. Fix a positive integer t, and let $\mathcal{Q}$ be the set of graphs that do not contain $K_{t,t}$ as a subgraph. Partition $[0, 1]$ into intervals $\{S_i\}_{i=1}^{\infty}$ with non-zero lengths, and let $\{H_i\}_{i=1}^{\infty}$ be an enumeration of the graphs in $\mathcal{Q}$. Define W to be the graphon that is 0 on $S_i \times S_j$ for $i \neq j$, and is equivalent to W_{H_i} (scaled properly) on $S_i \times S_i$. By construction $p(H; W) > 0$ if $H \in \mathcal{Q}$. Thus $|\mathrm{supp}(G(n, W))| \geqslant |\mathcal{Q}_n|$. Since there exist $K_{t,t}$-free graphs with $n^{2-2/t}$ edges (see, *e.g.*, [3, p. 316, Theorem VI.2.10]), we have $|\mathcal{Q}_n| \geqslant 2^{n^{2-2/t}}$.

It remains to show that W is a limit of graphs in some random-free property. Unfortunately, $W \notin \overline{\mathcal{Q}}$. We construct a larger random-free property $\mathcal{Q}'$ so that $W \in \overline{\mathcal{Q}'}$, as follows.

Fix a bigraph B, so that the corresponding graph contains $K_{t,t}$ as a subgraph and is connected. Suppose further that no two vertices of B have the same neighbourhood. Note that such a bigraph trivially exists. For example, one can take $B = (V_1 \cup U_1, V_2 \cup U_2, E)$ so that V_1, U_1, V_2, U_2 are disjoint sets of size t, every vertex of V_1 is joined to every vertex of V_2, and the edges between V_1 and U_2, as well as the edges between U_1 and V_2, form a matching of size t. Let $\mathcal{Q}' \supseteq \mathcal{Q}$ be the set of graphs not containing B as an induced sub-bigraph. Then $\mathcal{Q}'$ is random-free by Theorem 3.1, as $p^{\mathrm{b}}(B, W') = 0$ for every $W' \in \overline{\mathcal{Q}'}$.

Let $r = |V(B)|$ and suppose that $G = G(x_1, x_2, \ldots, x_r, W)$ contains B as an induced sub-bigraph. Then there exists i so that $x_1, x_2, \ldots, x_r \in S_i$, as G is connected. It follows further that G is an induced subgraph of H_i, as no two vertices of G have the same neighbourhood. Thus G contains no $K_{t,t}$ subgraph, contradicting our assumption that G contains B. We conclude that $\mathrm{supp}(G(n, W)) \subseteq \mathcal{Q}'$ for every positive integer n. By Lemma 2.6 of [7] the sequence $\{G(n, W)\}_{n=1}^{\infty}$ converges to W with probability one. Thus $W \in \overline{\mathcal{Q}'}$, as desired.

Acknowledgement

The authors wish to thank the anonymous referee whose detailed comments greatly improved the exposition.

References

[1] Aldous, D. J. (1985) Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour XIII, 1983*, Vol. 1117 of *Lecture Notes in Mathematics*, Springer, pp. 1–198.

[2] Alon, N., Fischer, E. and Newman, I. (2007) Efficient testing of bipartite graphs for forbidden induced subgraphs. *SIAM J. Comput*. **37** 959–976.

[3] Bollobás, B. (1978) *Extremal Graph Theory*, Vol. 11 of *London Mathematical Society Monographs*, Academic Press.

[4] Borgs, C., Chayes, J. and Lovász, L. (2010) Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal*. **19** 1597–1619.

[5] Hatami, H., Janson, S. and Szegedy, B. Graph properties, graph limits and entropy. In preparation.

[6] Janson, S. Graphons, cut norm and distance, couplings and rearrangements. `arXiv:1009.2376`

[7] Lovász, L. and Szegedy, B. (2006) Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957.

[8] Lovász, L. and Szegedy, B. (2010) Regularity partitions and the topology of graphons. In *An Irregular Mind*, Vol. 21 of *Bolyai Soc. Math. Stud*., pp. 415–446.