GRAPH PROPERTIES, GRAPH LIMITS AND ENTROPY

HAMED HATAMI, SVANTE JANSON, AND BALÁZS SZEGEDY

Abstract. We study the relation between the growth rate of a graph property and the entropy of the graph limits that arise from graphs with that property. In particular, for hereditary classes we obtain a new description of the colouring number, which by well-known results describes the rate of growth.

We study also random graphs and their entropies. We show, for example, that if a hereditary property has a unique limiting graphon with maximal entropy, then a random graph with this property, selected uniformly at random from all such graphs with a given order, converges to this maximizing graphon as the order tends to infinity.

1. Introduction and results

In recent years a theory of convergent sequences of dense graphs has been developed, see e.g. the book [Lov12]. One can construct a limit object for such a sequence in the form of certain symmetric measurable functions called graphons. The theory of graph limits not only provides a framework for addressing some previously unapproachable questions, but also leads to new interesting questions. For example one can ask: Which graphons arise as limits of sequences of graphs with a given property? Does a sequence of random graphs drawn from the set of graphs with a given property converge, and if so, what is the limit graphon? These types of questions has been studied for certain properties [CD11, LS06, DHJ08, Jan13b]. In this article we study the relation between these questions, the entropy of graphons, and the growth rate of graph properties.

The growth rate of graph properties has been studied extensively in the past, see e.g. [Ale92, BT97, BBW00, BBW01, BBS04, Bol07, BBSS09]. The standard method has been to use the Szemerédi regularity lemma, while we use graph limits; this should not be surprising, since it has been known since the introduction of graph limits that there is a strong connection with the Szemerédi regularity lemma. Some of our proofs reminisce the proofs from previous works, but in different formulations, cf. e.g. Bollobás and Thomason [BT97].

Date: 12 December, 2013.

HH supported by an NSERC, and an FQRNT grant. SJ supported by the Knut and Alice Wallenberg Foundation. This research was mainly done during the workshop *Graph limits, homomorphisms and structures II* at Hraniční Zámeček, Czech Republic, 2012.

1.1. **Preliminaries.** For every natural number n, denote $[n] := \{1, \dots, n\}$. In this paper all graphs are simple and finite. For a graph G, let $V(G)$ and $E(G)$, respectively, denote the vertex set and the edge set of G. We write for convenience $|G|$ for $|V(G)|$, the number of vertices. Let $\mathcal{U}$ denote set of all unlabelled graphs. (These are formally defined as equivalence classes of graphs up to isomorphisms.) Moreover for $n \geq 1$, let $\mathcal{U}_n \subset \mathcal{U}$ denote the set of all graphs in $\mathcal{U}$ with exactly n vertices. Sometimes we shall work with labelled graphs. For every $n \geq 1$, denote by $\mathcal{L}_n$ the set of all graphs with vertex set $[n]$.

We recall the basic notions of graph limits, see e.g. [LS06, BCL$^+$08, DJ08, Lov12] for further details. The *homomorphism density* of a graph H in a graph G, denoted by $t(H;G)$, is the probability that a uniformly 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 uniformly random *embedding* of the vertices of H in the vertices of G is an embedding of H in G, i.e. $uv \in E(H) \iff \phi(u)\phi(v) \in E(G)$. (This is often denoted $t_{\mathrm{ind}}(H;G)$. We assume $|H| \leq |G|$ so that embeddings exist.) We call a sequence of finite graphs $\{G_i\}_{i=1}^\infty$ with $|G_i| \to \infty$ *convergent* if for every finite graph H, the sequence $\{p(H;G_i)\}_{i=1}^\infty$ converges. (This is equivalent to $\{t(H;G_i)\}_{i=1}^\infty$ being convergent for every finite graph H.) One then may construct a completion $\overline{\mathcal{U}}$ of $\mathcal{U}$ under this notion of convergence. More precisely, $\overline{\mathcal{U}}$ is a compact metric space which contains $\mathcal{U}$ as a dense subset; the functionals $t(H;G)$ and $p(H;G)$ extend by continuity to $G \in \overline{\mathcal{U}}$, for each fixed graph H; elements of the complement $\widehat{\mathcal{U}} := \overline{\mathcal{U}} \setminus \mathcal{U}$ are called *graph limits*; a sequence of graphs (G_n) converges to a graph limit Γ if and only if $|G_n| \to \infty$ and $p(H;G_n) \to p(H;\Gamma)$ for every graph H. Moreover a graph limit is uniquely determined by the numbers $p(H;\Gamma)$ for all $H \in \mathcal{U}$.

It is shown in [LS06] 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 are denoted by $\mathcal{W}_0$. (We do not distinguish between graphons that are equal almost everywhere.) Given a graph G with vertex set $[n]$ and adjacency matrix A_G, we define the corresponding graphon $W_G : [0,1]^2 \to \{0,1\}$ as follows. Let $W_G(x,y) := A_G(\lceil xn \rceil, \lceil yn \rceil)$ 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) is a graph sequence that converges to a graph limit Γ, then for every graph H,

$$\begin{aligned} p(H;\Gamma) &= \lim_{n\to\infty} p(H;G_n) \\ &= \lim_{n\to\infty} \mathbb{E}\left[\prod_{uv\in E(H)} W_{G_n}(X_u,X_v) \prod_{uv\in E(H)^c} (1 - W_{G_n}(X_u,X_v))\right], \end{aligned}$$

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 [LS06] 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

$$p(H;W) := \mathbb{E}\left[\prod_{uv\in E(H)} W(X_u,X_v) \prod_{uv\in E(H)^c} (1-W(X_u,X_v))\right]. \tag{1.1}$$

Unfortunately, this graphon is not unique. We say that two graphons W and W' are *(weakly) equivalent* if they represent the same graph limit, i.e., if $p(H;W) = p(H;W')$ for all graphs H. For example, a graphon $W(x,y)$ is evidently equivalent to $W(\sigma(x),\sigma(y))$ for any measure-preserving map $\sigma : [0,1] \to [0,1]$. Not every pair of equivalent graphons is related in this way, but almost: Borgs, Chayes and Lovász [BCL10] proved that if W_1 and W_2 are two different graphons representing the same graph limit, then there exists a third graphon W and measure-preserving maps $\sigma_i : [0,1] \to [0,1]$, $i = 1,2$, such that

$$W_i(x,y) = W(\sigma_i(x),\sigma_i(y)), \qquad \text{for a.e. } x,y. \tag{1.2}$$

(For other characterizations of equivalent graphons, see e.g. [BR09] and [Jan13a].)

The set $\widehat{\mathcal{U}}$ of graph limits is thus a quotient space of the set $\mathcal{W}_0$ of graphons. Nevertheless, we shall not always distinguish between graph limits and their corresponding graphons; it is often convenient (and customary) to let a graphon W also denote the corresponding graph limit. For example, we may write $G_n \to W$ when a sequence of graphs $\{G_n\}$ converges to the graph limit determined by the graphon W; similarly we say that a sequence of graphons W_n converges to W in $\mathcal{W}_0$ if the corresponding sequence of graph limits converges in $\overline{\mathcal{U}}$. (This makes $\mathcal{W}_0$ into a topological space that is compact but not Hausdorff.)

For every $n \geq 1$, a graphon W defines a random graph $G(n,W) \in \mathcal{L}_n$: Let $X_1,\dots,X_n$ be an i.i.d. sequence of random variables taking values uniformly in $[0,1]$. Given $X_1,\dots,X_n$, let ij be an edge with probability $W(X_i,X_j)$, independently for all pairs (i,j) with $1 \leq i < j \leq n$. It follows that for every $H \in \mathcal{L}_n$,

$$\mathbb{P}[G(n,W) = H] = p(H;W). \tag{1.3}$$

The distribution of $G(n,W)$ is thus the same for two equivalent graphons, so we may define $G(n,\Gamma)$ for a graph limit Γ; this is a random graph that also can be defined by the analogous relation $\mathbb{P}[G(n,\Gamma) = H] = p(H;\Gamma)$ for $H \in \mathcal{L}_n$.

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. Let $\mathcal{Q} \subseteq \mathcal{U}$ be a graph class. For every $n \geq 1$, we denote by $\mathcal{Q}_n := \mathcal{Q} \cap \mathcal{U}_n$ the set of graphs in $\mathcal{Q}$ with exactly n vertices. We also consider the corresponding class of labelled graphs, and define $\mathcal{Q}_n^L$ to be the set of all graphs in $\mathcal{L}_n$ that belong to $\mathcal{Q}$ (when we ignore labels). Furthermore, we let $\overline{\mathcal{Q}} \subseteq \overline{\mathcal{U}}$ be the closure of $\mathcal{Q}$ in $\overline{\mathcal{U}}$ and $\widehat{\mathcal{Q}} := \overline{\mathcal{Q}} \cap \widehat{\mathcal{U}} = \overline{\mathcal{Q}} \setminus \mathcal{Q}$ the set of graph limits that are limits of sequences of graphs in $\mathcal{Q}$.

Define the *binary entropy* function $h : [0,1] \mapsto \mathbb{R}_+$ as

$$h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$$

for $x \in [0,1]$, with the interpretation $h(0) = h(1) = 0$ so that h is continuous on $[0,1]$, where here and throughout the paper $\log_2$ denotes the logarithm to the base 2. Note that $0 \le h(x) \le 1$, with $h(x) = 0$ attained at $x = 0, 1$ and $h(x) = 1$ at $x = 1/2$, only. The *entropy* of a graphon W is defined as

$$\mathrm{Ent}(W) := \int_0^1 \int_0^1 h(W(x,y))\,\mathrm{d}x\,\mathrm{d}y. \tag{1.4}$$

This is related to the entropy of random graphs, see [Ald85] and [Jan13a, Appendix D.2] and (4.9) below; it has also previously been used by Chatterjee and Varadhan [CV11] and Chatterjee and Diaconis [CD11] to study large deviations of random graphs and exponential models of random graphs. Note that it follows from the uniqueness result (1.2) that the entropy is a function of the underlying graph limit and it does not depend on the choice of the graphon representing it; we may thus define the entropy $\mathrm{Ent}(\Gamma)$ of a graph limit Γ as the entropy $\mathrm{Ent}(W)$ of any graphon representing it.

Our first theorem bounds the rate of growth of an arbitrary graph class in terms of the entropy of the limiting graph limits (or graphons).

Theorem 1.1. *Let $\mathcal{Q}$ be a class of graphs. Then*

$$\limsup_{n\to\infty} \frac{\log_2 |\mathcal{Q}_n|}{\binom{n}{2}} \le \max_{\Gamma\in\widehat{\mathcal{Q}}} \mathrm{Ent}(\Gamma). \tag{1.5}$$

We present the proofs of this and the following theorems in Section 5.

Remark 1.2. For any graph class $\mathcal{Q}$, and $n \ge 1$,

$$|\mathcal{Q}_n| \le |\mathcal{Q}_n^L| \le n!|\mathcal{Q}_n|. \tag{1.6}$$

The factor $n!$ is for our purposes small and can be ignored, since $\log_2 n! = o(n^2)$. Thus we may replace $|\mathcal{Q}_n|$ by $|\mathcal{Q}_n^L|$ in Theorem 1.1. The same holds for the theorems below.

Remark 1.3. $|\mathcal{Q}_n| \le |\mathcal{U}_n| \le |\mathcal{L}_n| = 2^{\binom{n}{2}}$, so the left-hand side of (1.5) is at most 1, and it equals 1 if $\mathcal{Q}$ is the class of all graphs, cf. (1.6). Furthermore, by (1.4), $\mathrm{Ent}(W) \in [0,1]$ for every graphon W. In the trivial case when $\mathcal{Q}$ is a finite class, $\mathcal{Q}_n = \emptyset$ for all large n and the left-hand side is $-\infty$; in this case $\widehat{\mathcal{Q}} = \emptyset$ and the right-hand side is also (interpreted as) $-\infty$. We exclude in the sequel this trivial case; thus both sides of (1.5) are in $[0,1]$. Note further that $\mathrm{Ent}(W) = 1$ only when $W = 1/2$ a.e.; thus the right-hand side of (1.5) equals 1 if and only if $\widehat{\mathcal{Q}}$ contains the graph limit defined by the constant graphon $W = 1/2$. (This graphon is the limit of sequences of quasi-random graphs, see [LS06].)

A graphon is called *random-free* if it is $\{0,1\}$-valued almost everywhere, see [LS10, Jan13a]. Note that a graphon W is random-free if and only if $\mathrm{Ent}(W) = 0$. This is preserved by equivalence of graphons, so we may define a graph limit to be random-free if some (or any) representing graphon is random-free; equivalently, if its entropy is 0. A property $\mathcal{Q}$ is called *random-free* if every $\Gamma \in \widehat{\mathcal{Q}}$ is random-free. Theorem 1.1 has the following immediate corollary:

Corollary 1.4. *If $\mathcal{Q}$ is a random-free class of graphs, then $|\mathcal{Q}_n| = 2^{o(n^2)}$.*

For further results on random-free graphons and random-free classes of graphs, see Hatami and Norine [HN12].

A graph class $\mathcal{P}$ is *hereditary* if whenever a graph G belongs to $\mathcal{Q}$, then every induced subgraph of G also belongs to $\mathcal{P}$.

Our second theorem says that when $\mathcal{Q}$ is a hereditary graph property, equality holds in (1.5). (See also Theorem 1.9 below.)

Theorem 1.5. *Let $\mathcal{Q}$ be a hereditary class of graphs. Then*

$$\lim_{n\to\infty} \frac{\log_2 |\mathcal{Q}_n|}{\binom{n}{2}} = \max_{\Gamma\in\widehat{\mathcal{Q}}} \mathrm{Ent}(\Gamma).$$

Our next theorem concerns the limit of the sequences of random graphs that are sampled from a graph class. There are two natural ways to sample a random graph sequence (G_n), with $|G_n| = n$, from a graph class $\mathcal{Q}$. The first is to pick an unlabelled graph G_n uniformly at random from $\mathcal{Q}_n$, for each $n \geq 1$ (assuming that $\mathcal{Q}_n \neq \emptyset$). The second is to pick a labelled graph G_n uniformly at random from $\mathcal{Q}_n^L$. We call the resulting random graph G_n a *uniformly random unlabelled element of $\mathcal{Q}_n$* and a *uniformly random labelled element of $\mathcal{Q}_n$*, respectively.

Theorem 1.6. *Suppose that $\max_{\Gamma\in\widehat{\mathcal{Q}}} \mathrm{Ent}(\Gamma)$ is attained by a unique graph limit $\Gamma_{\mathcal{Q}}$. Suppose further that equality holds in (1.5), i.e.*

$$\lim_{n\to\infty} \frac{\log_2 |\mathcal{Q}_n|}{\binom{n}{2}} = \mathrm{Ent}(\Gamma_{\mathcal{Q}}). \tag{1.7}$$

Then

(i) *If $G_n \in \mathcal{U}_n$ is a uniformly random unlabelled element of $\mathcal{Q}_n$, then G_n converges to $\Gamma_{\mathcal{Q}}$ in probability as $n \to \infty$.*
(ii) *The same holds if $G_n \in \mathcal{L}_n$ is a uniformly random labelled element of $\mathcal{Q}_n^L$.*

Remark 1.7. Note that for hereditary properties, it suffices to only assume that $\max_{\Gamma\in\widehat{\mathcal{Q}}} \mathrm{Ent}(\Gamma)$ is attained by a unique graph limit $\Gamma_{\mathcal{Q}}$ as then (1.7) follows from Theorem 1.5.

The next theorem concerns sequences of random graphs drawn from arbitrary distributions, not necessarily uniform. A random labelled [unlabelled] graph G_n on n vertices is thus any random variable with values in $\mathcal{U}_n$ [$\mathcal{L}_n$]. We consider convergence in distribution of G_n, regarding G_n as a random element of $\mathcal{U} \subset \overline{\mathcal{U}}$ (ignoring labels if there are any); the limit in distribution (if it exists) is thus a random element of $\overline{\mathcal{U}}$, which easily is seen to be concentrated on $\widehat{\mathcal{U}}$; in other words, the limit is a random graph limit.

Recall that the entropy $\mathrm{Ent}(X)$ of a random variable X taking values in some finite (or countable) set A is $\sum_{a\in A}(-p_a \log_2 p_a)$, where $p_a := \mathbb{P}(X = a)$.

Theorem 1.8. *Suppose that G_n is a (labelled or unlabelled) random graph on n vertices with some distribution μ_n. Suppose further that as $n \to \infty$, G_n converges in distribution to some random graph limit with distribution μ. Then*

$$\limsup_{n\to\infty} \frac{\mathrm{Ent}(G_n)}{\binom{n}{2}} \le \max_{W\in \mathrm{supp}(\mu)} \mathrm{Ent}(W),$$

where $\mathrm{supp}(\mu) \subseteq \widehat{\mathcal{U}}$ *is the support of the probability measure μ.*

1.3. **Maximal entropy graphons.** The results in Section 1.2 show that graphons with maximal entropy capture the growth rate and other asymptotic behaviors of graph classes. In this section we study the structure of those graphons for hereditary classes.

We define the *randomness support* of a graphon W as

$$\mathrm{rand}(W) := \left\{(x,y) \in [0,1]^2 : 0 < W(x,y) < 1\right\}, \tag{1.8}$$

and its *random part* as the restriction of W to $\mathrm{rand}(W)$. Finally the *randomness support graphon* of W is defined as $\mathbf{1}_{\mathrm{rand}(W)}$, the indicator of its randomness support.

A graphon W is called K_r-free (where $r \ge 1$) if $p(K_r, W) = 0$; by (1.1), this is equivalent to $\prod_{1\le i<j\le r} W(x_i, x_j) = 0$ for almost every $x_1, \dots, x_r$. (The case $r = 1$ is trivial: no graphon is K_1-free.) Recall that the Turán graph $T_{n,r}$ is the balanced complete r-partite graph with n vertices. For each $r \ge 1$, the graphs $T_{n,r}$ converge to the K_{r+1}-free graphon W_{K_r} as $n \to \infty$.

Let E_r denote the support of W_{K_r}, i.e., $E_r := \bigcup_{i\ne j} I_i \times I_j$ where $I_i := ((i-1)/r, i/r]$ for $i = 1, \dots, r$, and also define $E_\infty := [0,1]^2$. For $1 \le r \le \infty$, let R_r be the set of graphons W such that $W(x,y) = \frac{1}{2}$ on E_r and $W(x,y) \in \{0,1\}$ otherwise. In other words, W has randomness support E_r and its random part is $\frac{1}{2}$ everywhere. Note that $E_1 = \emptyset$ and thus R_1 is the set of random-free graphons, while R_∞ consists only of the constant graphon $\frac{1}{2}$. If $W \in R_r$, then

$$\mathrm{Ent}(W) = \int_{E_r} h(1/2) = |E_r| = 1 - \frac{1}{r}. \tag{1.9}$$

A simple example of a graphon in R_r is $\frac{1}{2}W_{K_r}$. (For $r < \infty$, this is the almost surely limit of a uniformly random subgraph of $T_{n,r}$ as $n \to \infty$.) More generally, if $r < \infty$, we can modify $\frac{1}{2}W_{K_r}$ by changing it on each square I_i^2

for $i = 1, \dots, r$ to a symmetric measurable $\{0,1\}$-valued function (i.e. to any random-free graphon, scaled in the natural way); this gives all graphons in R_r.

We let, for $1 \le r < \infty$ and $0 \le s \le r$, $W^*_{r,s}$ be the graphon in R_r that is 1 on $I_i \times I_i$ for $i \le s$ and 0 on $I_i \times I_i$ for $i > s$. (Thus $W^*_{r,0} = \frac{1}{2} W_{K_r}$.)

For a class $\mathcal{Q}$ of graphs, let

$$\widehat{\mathcal{Q}}^* := \left\{ \Gamma \in \widehat{\mathcal{Q}} : \operatorname{Ent}(\Gamma) = \max_{\Gamma \in \widehat{\mathcal{Q}}} \operatorname{Ent}(\Gamma) \right\}$$

denote the set of graph limits in $\widehat{\mathcal{Q}}$ with maximum entropy. It follows from Lemma 3.3 below that the maximum is attained and that $\widehat{\mathcal{Q}}^*$ is a non-empty closed subset of $\widehat{\mathcal{Q}}$, and thus a non-empty compact set.

After these preparations, we state the following result, improving Theorem 1.5.

Theorem 1.9. *Let $\mathcal{Q}$ be a hereditary class of graphs. Then there exists a number $r \in \{1, 2 \dots, \infty\}$ such that $\max_{\Gamma \in \widehat{\mathcal{Q}}} \operatorname{Ent}(\Gamma) = 1 - \frac{1}{r}$, every graph limit in $\widehat{\mathcal{Q}}^*$ can be represented by a graphon $W \in R_r$, and*

$$|\mathcal{Q}_n| = 2^{(1 - r^{-1} + o(1))\binom{n}{2}}. \tag{1.10}$$

Hence, $\widehat{\mathcal{Q}}^* = \widehat{\mathcal{Q}} \cap R_r$. Moreover, r has the further characterisations

$$r = \min \left\{ s \ge 1 : \mathbf{1}_{\mathrm{rand}(W)} \text{ is } K_{s+1}\text{-free for all graphons } W \in \widehat{\mathcal{Q}} \right\} \tag{1.11}$$

$$= \sup \left\{ t : W^*_{t,u} \in \widehat{\mathcal{Q}} \text{ for some } u \le t \right\}, \tag{1.12}$$

where the minimum in (1.11) is interpreted as ∞ when there is no such s. Furthermore $r = 1$ if and only if $\mathcal{Q}$ is random-free, and $r = \infty$ if and only if $\mathcal{Q}$ is the class of all graphs.

The result (1.10) is a fundamental result for hereditary classes of graphs, proved by Alekseev [Ale92] and Bollobás and Thomason [BT97], see also the survey [Bol07] and e.g. [BBW00, BBW01, BBS04, BBS09, BBSS09, ABBM11]. The number r is known as the *colouring number* of $\mathcal{Q}$.

Remark 1.10. Let, for $1 \le r < \infty$ and $0 \le s \le r$, $\mathcal{C}(r,s)$ be the hereditary class of all graphs such that the vertex set can be partitioned into r (possibly empty) sets V_i with the subgraph induced by V_i complete for $1 \le i \le s$ and empty for $s < i \le r$. Note that $G(n, W^*_{r,s}) \in \mathcal{C}(r,s)$ a.s., and that every graph in $\mathcal{C}(r,s)$ with n vertices appears with positive probability. (In fact, $G \in \mathcal{C}(r,s) \iff p(G, W^*_{r,s}) > 0$.) It follows from (1.3) and Lemma 3.2 below that, for any hereditary class $\mathcal{Q}$, $W^*_{r,s} \in \widehat{\mathcal{Q}}$ if and only if $\mathcal{C}(r,s) \subseteq \mathcal{Q}$. Hence, (1.12) shows that r (when finite) is the largest integer such that $\mathcal{C}(r,s) \subseteq \mathcal{Q}$ for some s; this is the traditional definition of the colouring number, see e.g. [Bol07] where further comments are given.

2. Examples

We give a few examples to illustrate the results. We begin with a simple case.

Example 2.1 (Bipartite graphs)**.** Let $\mathcal{Q}$ be the class of *bipartite graphs*; note that this equals the class $\mathcal{C}(2,0)$ in Remark 1.10. Suppose that a graph limit $\Gamma \in \widehat{\mathcal{Q}}$. Then there exists a sequence of graphs $G_n \to \Gamma$ with $G_n \in \mathcal{Q}$, where for simplicity we may assume $|G_n| = n$. Since G_n is bipartite, it has a bipartition that can be assumed to be $\{1,\dots,m_n\}$ and $\{m_n+1,\dots,n\}$. By selecting a subsequence, we may assume that $m_n/n \to a$ for some $a \in [0,1]$, and it is then easy to see (for example by using the bipartite limit theory in [DHJ08, Section 8]) that Γ can be represented by a graphon that vanishes on $[0,a]^2 \cup [a,1]^2$. Conversely, if W is such a graphon, then the random graph $G(n;W)$ is bipartite, and thus $W \in \widehat{\mathcal{Q}}$. Hence $\widehat{\mathcal{Q}}$ equals the set of graph limits represented (non-uniquely) by the graphons

$$\bigcup_{a\in[0,1]} \bigl\{W : W = 0 \text{ on } [0,a]^2 \cup [a,1]^2\bigr\}. \tag{2.1}$$

If W is a graphon in the set (2.1), with a given a, then the support of W has measure at most $2a(1-a)$, and thus

$$\operatorname{Ent}(W) \le 2a(1-a), \tag{2.2}$$

with equality if and only if $W = \frac12$ on $(0,a)\times(a,1) \cup (a,1)\times(0,a)$. The maximum entropy is obtained for $a = 1/2$, and thus

$$\max_{\Gamma\in\widehat{\mathcal{Q}}} \operatorname{Ent}(\Gamma) = \tfrac12, \tag{2.3}$$

and the maximum is attained by a unique graph limit, represented by the graphon $W^*_{2,0}$ defined above.

Theorem 1.5 thus says that $|\mathcal{Q}_n| = 2^{\frac12\binom{n}{2}+o(n^2)}$ (which can be easily proved directly). Theorem 1.6 says that if G_n is a uniformly random (labelled or unlabelled) bipartite graph, then $G_n \to W^*_{2,0}$ in probability. The colouring number r in Theorem 1.9 equals 2, and both (1.11) and (1.12) are easily verified directly.

Example 2.2 (Triangle-free graphs)**.** Let $\mathcal{Q}$ be the class of *triangle-free* graphs. It is easy to see that the corresponding class of graph limits $\widehat{\mathcal{Q}}$ is the class of triangle-free graph limits $\{\Gamma : p(K_3,\Gamma) = 0\}$ defined in Section 1.3, see [Jan13b, Example 4.3].

This class is strictly larger than the class of bipartite graphs; the set $\widehat{\mathcal{Q}}$ of triangle-free graph limits thus contains the set (2.1) of bipartite graph limits, and it is easily seen that it is strictly larger. (An example of a triangle-free graph limit that is not bipartite is W_{C_5}.)

We do not know any representation of all triangle-free graph limits similar to (2.1), but it is easy to find the ones of maximum entropy. If a graphon W is triangle-free, then so is its randomness support graphon, and Lemma 6.4

below shows that $\text{Ent}(W) \leq \frac{1}{2}$, with equality only if $W \in R_2$ (up to equivalence). Furthermore, it is easy to see that if $W \in R_r$ is triangle-free, then $W(x,y) \neq 1$ a.e., and thus $W = W^*_{2,0}$. (Use Theorem 6.1 below, or note that $\max\{W(x,y), \frac{1}{2}\}$ is another triangle-free graphon.) Thus, as in Example 2.1, $W^*_{2,0}$ represents the unique graph limit in $\widehat{\mathcal{Q}}$ with maximum entropy.

Theorem 1.5 and 1.9 thus say that $|\mathcal{Q}_n| = 2^{\frac{1}{2}\binom{n}{2}+o(n^2)}$, as shown by Erdős, Kleitman and Rothschild [EKR76]. (They also proved that almost all triangle-free graphs are bipartite; this seems related to the fact that the two graph classes have the same maximum entropy graph limit, although we do not know any direct implication.)

Theorem 1.6 says that if G_n is a uniformly random (labelled or unlabelled) triangle-free graph, then $G_n \to W^*_{2,0}$ in probability.

The same argument applies to K_t-free graphs, for any $t \geq 2$. The colouring number is $t-1$ and thus the number of such graphs of order n is $2^{\frac{r-2}{r-1}\binom{n}{2}+o(n^2)}$, as shown in [EKR76]. (See also [KPR85, KPR87].) The unique graph limit of maximum entropy is represented by $W^*_{t-1,0}$. Thus Theorem 1.6 applies and shows that, hardly surprising, a random K_t-free graph converges (in probability) to the graphon $W^*_{t-1,0}$.

Example 2.3 (Split graphs)**.** Another simple application of Theorem 1.6 is given in [Jan13b, Section 10], where it is shown that the class of *split graphs* has a unique graph limit with maximal entropy, represented by the graphon $W^*_{2,1}$; this is thus the limit (in probability) of a uniformly random split graph. Recall that the class of split graphs equals $\mathcal{C}(2,1)$ in Remark 1.10; in other words, a graph is a split graph if its vertex set can be partitioned into two sets, one of which is a clique and the other one is an isolated set. (Equivalently, G is a split graph if and only if $p(G; W^*_{2,1}) > 0$.)

Our final example is more complicated, and we have less complete results.

Example 2.4 (String graphs)**.** A *string graph* is the intersection graph of a family of curves in the plane. In other words, G is a string graph if there exists a collection $\{A_v : v \in V(G)\}$ of curves such that $ij \in E(G) \iff A_i \cap A_j \neq \emptyset$. It is easily seen that we obtain the same class of graphs if we allow the sets A_v to be arbitrary arcwise connected sets in the plane.

It is shown by Pach and Tóth [PT06] that the number of string graphs of order n is $2^{\frac{3}{4}\binom{n}{2}+o(n^2)}$. Thus, Theorems 1.5 and 1.9 hold with maximum entropy $\frac{3}{4}$ and colouring number 4.

We study this further by interpreting the proof of [PT06] in our graph limit context. To show a lower bound on the number of string graphs, [PT06] shows that every graph in the class $\mathcal{C}(4,4)$ is a string graph. (This was proved already in [KGK86, Corollary 2.7].) A minor modification of their construction is as follows: Let G be a graph with a partition $V(G) = \bigcup_{i=1}^4 V_i$ such that each V_i is a complete subgraph of G. Consider a drawing of the graph K_4 in the plane, with vertices $x_1, ..., x_4$ and non-crossing edges.

Replace each edge ij in K_4 by a number of parallel curves γ_{vw} from x_i to x_j, indexed by pairs $(v,w) \in V_i \times V_j$. (All curves still non-intersecting except at the end-points.) Choose a point x_{vw} on each curve γ_{vw}, and split γ_{vw} into the parts γ^*_{vw} from x_i to x_{vw} and γ^*_{wv} from x_{vw} to x_j, with x_{vw} included in both parts. If v is a vertex in G, and $v \in V_i$, let A_v be the (arcwise connected) set consisting of x_i and the curves γ^*_{vw} for all $w \notin V_i$ such that $vw \in E(G)$. Then G is the intersection graph defined by the collection $\{A_v\}$, and thus G is a string graph.

It follows, see Remark 1.10, that if $\mathcal{Q}$ is the class of string graphs, then $W^*_{4,4} \in \widehat{\mathcal{Q}}$.

To show an upper bound, Pach and Tóth [PT06] consider the graph G_5, which is the intersection graph of the family of the 15 subsets of order 1 or 2 of $\{1,\dots,5\}$. They show that G_5 is not a string graph, but that $G_5 \in \mathcal{C}(5,s)$ for every $0 \le s \le 5$. Thus $\mathcal{C}(5,s) \not\subseteq \mathcal{Q}$, and thus $W^*_{5,s} \notin \widehat{\mathcal{Q}}$, see Remark 1.10.

Consequently, we have $W^*_{4,4} \in \widehat{\mathcal{Q}}$ but $W^*_{5,s} \notin \widehat{\mathcal{Q}}$, for all s. Hence Theorem 1.9 shows that the colouring number $r = 4$, see (1.12), and that $W^*_{4,4}$ is one graphon in $\widehat{\mathcal{Q}}$ with maximal entropy.

However, in this case the graph limit of maximal entropy is *not* unique. Indeed, the construction above of string graphs works for any planar graph H instead of K_4, and G such that its vertex set can be partitioned into cliques V_i, $i \in V(H)$, with no edges in G between V_i and V_j unless $ij \in E(H)$. (See [KGK86, Theorem 2.3].) Taking H to be K_5 minus an edge, we thus see that if $G \in \mathcal{C}(4,4)$, and we replace the clique on V_1 by a disjoint union of two cliques (on the same vertex set V_1, leaving all other edges), then the new graph is also a string graph. It follows by taking the limit of a suitable sequence of such graphs, or by Lemma 3.2 below, that if $I_i := ((i-1)/4, i/4]$ and I_1 is split into $I_{11} := (0,a]$ and $I_{12} := (a, 1/4]$, where $0 \le a \le 1/8$, then the graphon $W^{**}_a \in R_4$ obtained from $W^*_{4,4}$ by replacing the value 1 by 0 on $(I_{11} \times I_{12}) \cup (I_{12} \times I_{11})$ satisfies $W^{**}_a \in \widehat{\mathcal{Q}} \cap R_4 = \widehat{\mathcal{Q}}^*$. Explicitly,

$$W^{**}_a(x,y) = \begin{cases} 1/2 & \text{on } \bigcup_{i \ne j}(I_i \times I_j); \\ 0 & \text{on } (I_{11} \times I_{12}) \cup (I_{12} \times I_{11}); \\ 1 & \text{on } (I_{11} \times I_{11}) \cup (I_{12} \times I_{12}) \cup \bigcup_{i=2}^4 (I_i \times I_i). \end{cases}$$

Thus $W^{**}_0 = W^*_{4,4}$, but the graphons W^{**}_a for $a \in [0, 1/8]$ are not equivalent, for example because they have different edge densities

$$\iint W^{**}_a = \frac{5}{8} - \frac{a}{2} + 2a^2 = \frac{19}{32} + 2\Big(\frac{1}{8} - a\Big)^2.$$

Thus there are infinitely many graph limits in $\widehat{\mathcal{Q}}^* = \widehat{\mathcal{Q}} \cap R_4$. (We do not know whether there are further such graph limits.)

Consequently, Theorem 1.6 does not apply to string graphs. We do not know whether a uniformly random string graph converges (in probability) to some graph limit as the size tends to infinity, and if so, what the limit is. We leave this as an open problem.

3. Some auxiliary facts

We start by recalling some basic facts about the binary entropy. First note that h is concave on $[0,1]$. In particular if $0 \le x_1 \le x_2 \le 1$, then

$$h(x_2) - h(x_1) \le h(x_2 - x_1) - h(0) = h(x_2 - x_1),$$

and

$$-(h(x_2) - h(x_1)) = h(x_1) - h(x_2) = h(1 - x_1) - h(1 - x_2) \le h(x_2 - x_1);$$

hence

$$|h(x_2) - h(x_1)| \le h(x_2 - x_1). \tag{3.1}$$

The following simple lemma relates $\binom{N}{m}$ to the binary entropy.

Lemma 3.1. *For integers $N \ge m \ge 0$, we have*

$$\binom{N}{m} \le \left(\frac{N}{m}\right)^m \left(\frac{N}{N-m}\right)^{N-m} = 2^{Nh(m/N)}.$$

Proof. Set $p = m/N$. If X has the binomial distribution $\mathrm{Bin}(N,p)$, then

$$1 \ge \mathbb{P}[X = m] = \binom{N}{m} p^m (1-p)^{N-m}$$

and thus

$$\binom{N}{m} \le p^{-m}(1-p)^{-(N-m)} = \left(\frac{N}{m}\right)^m \left(\frac{N}{N-m}\right)^{N-m} = 2^{Nh(p)}. \qquad \square$$

We will need the following simple lemma about hereditary classes of graphs [Jan13b]:

Lemma 3.2. *Let $\mathcal{Q}$ be a hereditary class of graphs and let W be a graphon. Then $W \in \widehat{\mathcal{Q}}$ if and only if $p(F;W) = 0$ when $F \notin \mathcal{Q}$.*

Proof. If $F \notin \mathcal{Q}$, then $p(F;G) = 0$ for every $G \in \mathcal{Q}$ since $\mathcal{Q}$ is hereditary, and thus $p(F;W) = 0$ for every $W \in \overline{\mathcal{Q}}$ by continuity.

For the converse, assume that $p(F;W) = 0$ when $F \notin \mathcal{Q}$. Thus $p(F;W) > 0 \implies F \in \mathcal{Q}$. By (1.3), if $\mathbb{P}(G(n,W) = H) > 0$, then $p(H;W) > 0$ and thus $H \in \mathcal{Q}$. Hence, $G(n,W) \in \mathcal{Q}$ almost surely. The claim follows from the fact [BCL$^+$08] that almost surely $G(n,W)$ converges to W as $n \to \infty$. $\square$

Next we recall that the *cut norm* of an $n \times n$ matrix $A = (A_{ij})$ is defined by

$$\|A\|_\square := \frac{1}{n^2} \max_{S,T \subseteq [n]} \left| \sum_{i \in S, j \in T} A_{ij} \right|.$$

Similarly, the *cut norm* of a measurable $W : [0,1]^2 \to \mathbb{R}$ is defined as

$$\|W\|_\square = \sup \left| \iint f(x) W(x,y) g(y) \,\mathrm{d}x \,\mathrm{d}y \right|,$$

where the supremum is over all measurable functions $f, g : [0,1] \to \{0,1\}$. (See [BCL+08] and [Jan13a] for other versions, equivalent within constant factors.) We use also the notation, for two graphons W_1 and W_2,

$$d_\square(W_1, W_2) := \|W_1 - W_2\|_\square. \tag{3.2}$$

The *cut distance* between two graphons W_1 and W_2 is defined as

$$\delta_\square(W_1, W_2) := \inf_{W_2'} d_\square(W_1, W_2') = \inf_{W_2'} \|W_1 - W_2'\|_\square, \tag{3.3}$$

where the infimum is over all graphons W_2' that are equivalent to W_2. (See [BCL+08] and [Jan13a] for other, equivalent, definitions.) The cut distance is a pseudometric on $\mathcal{W}_0$, with $\delta_\square(W_1, W_2) = 0$ if and only if W_1 and W_2 are equivalent.

The cut distance between two graphs F and G is defined as $\delta_\square(F, G) = \delta_\square(W_F, W_G)$. We similarly write $\delta_\square(F, W) = \delta_\square(W_F, W)$, $d_\square(F, W) = d_\square(W_F, W)$ and so on.

The cut distance is a central notion in the theory of graph limits. For example it is known (see [BCL+08] and [Lov12]) that a graph sequence (G_n) with $|G_n| \to \infty$ converges to a graphon W if and only if the sequence (W_{G_n}) converges to W in cut distance. Similarly, convergence of a sequence of graphons in $\mathcal{W}_0$ is the same as convergence in cut distance; hence, the cut distance induces a metric on $\widehat{\mathcal{U}}$ that defines its topology.

Let $\mathcal{P}$ be a partition of the interval $[0,1]$ into k measurable sets $I_1, \dots, I_k$. Then $I_1, \dots, I_k$ divide the unit square $[0,1]^2$ into k^2 measurable sets $I_i \times I_j$. We denote the corresponding σ-algebra by $\mathcal{B}_\mathcal{P}$; note that if W is a graphon, then $\mathbb{E}[W \mid \mathcal{B}_\mathcal{P}]$ is the graphon that is constant on each set $I_i \times I_j$ and obtained by averaging W over each such set. A partition of the interval $[0,1]$ into k sets is called an *equipartition* if all sets are of measure $1/k$. We let $\overline{\mathcal{P}}_k$ denote the equipartition of $[0,1]$ into k intervals of length $1/k$, and write, for any graphon W,

$$\overline{W}_k := \mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]. \tag{3.4}$$

Similarly, if $\mathcal{P}$ is a partition of $[n]$ into sets $V_1, \dots, V_k$, then we consider the corresponding partition $I_1, \dots, I_k$ of $[0,1]$ (that is $x \in (0,1]$ belongs to I_j if and only if $\lceil x_j \rceil \in V_j$) and again we denote the corresponding σ-algebra on $[0,1]^2$ by $\mathcal{B}_\mathcal{P}$. A partition $\mathcal{P}$ of $[n]$ into k sets is called an *equipartition* if each part is of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$.

The graphon version of the weak regularity lemma proved by Frieze and Kannan [FK99], see also [LS07] and [Lov12, Sections 9.1.2 and 9.2.2], says that for every every graphon W and every $k \geq 1$, there is an equipartition $\mathcal{P}$ of $[0,1]$ into k sets such that

$$\|W - \mathbb{E}[W \mid \mathcal{B}_\mathcal{P}]\|_\square \leq \frac{4}{\sqrt{\log_2 k}}. \tag{3.5}$$

Let us close this section with the following simple lemma. Part (ii) has been proved by Chatterjee and Varadhan [CV11], but we include a (different) proof for completeness.

Lemma 3.3. *The function* $\mathrm{Ent}(\cdot)$ *satisfies the following properties:*

(i) *If* W *is a graphon and* $\mathcal{P}$ *is a measurable partition of* $[0,1]$*, then*

$$\mathrm{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]) \geq \mathrm{Ent}(W).$$

(ii) *The function* $\mathrm{Ent}(\cdot)$ *is lower semicontinuous on* $\mathcal{W}_0$ *(and, equivalently, on* $\widehat{\mathcal{U}}$*). I.e., if* $W_m \to W$ *in* $\mathcal{W}_0$ *as* $m \to \infty$*, then*

$$\limsup_{m\to\infty} \mathrm{Ent}(W_m) \leq \mathrm{Ent}(W).$$

Remark 3.4. $\mathrm{Ent}(\cdot)$ is *not* continuous. For example, let G_n be a quasirandom sequence of graphs with $G_n \to W = \frac{1}{2}$ (a constant graphon); then $W_{G_n} \to W = \frac{1}{2}$ in $\mathcal{W}_0$, but $\mathrm{Ent}(W_{G_n}) = 0$ and $\mathrm{Ent}(W) = 1$.

Proof. Part (i) follows from Jensen's inequality and concavity of h.

To prove (ii), note that we can assume $\|W_m - W\|_\square \to 0$. For every $k \geq 1$, let $\overline{\mathcal{P}}_k$ be the partition of $[0,1]$ into k consecutive intervals of equal measure $1/k$. Consider the step graphons $\mathbb{E}[W_m \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]$ and $\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]$. For each k, $\mathbb{E}[W_m \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]$ converges to $\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]$ almost everywhere as $m \to \infty$, and thus by (1.4) and dominated convergence,

$$\lim_{m\to\infty} \mathrm{Ent}(\mathbb{E}[W_m \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]) = \mathrm{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]).$$

Consequently, using (i),

$$\limsup_{m\to\infty} \mathrm{Ent}(W_m) \leq \limsup_{m\to\infty} \mathrm{Ent}(\mathbb{E}[W_m \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]) = \mathrm{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]).$$

Finally, let $k \to \infty$. Then $\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}] \to W$ almost everywhere, and thus $\mathrm{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]) \to \mathrm{Ent}(W)$. □

4. Number of graphs and Szeméredi partitions

In this section we prove some of the key lemmas needed in this paper. These lemmas provide various estimates on the number of graphs on n vertices that are close to a graphon in cut distance. For an integer $n \geq 1$, a parameter $\delta > 0$, and a graphon W, define

$$\widehat{N}_\square(n, \delta; W) := |\{G \in \mathcal{L}_n : d_\square(W_G, W) \leq \delta\}| \tag{4.1}$$

and

$$N_\square(n, \delta; W) := |\{G \in \mathcal{L}_n : \delta_\square(G, W) \leq \delta\}|\,. \tag{4.2}$$

Since $\delta_\square(G, W) \leq d_\square(W_G, W)$, cf. (3.3), we have trivially

$$\widehat{N}_\square(n, \delta; W) \leq N_\square(n, \delta; W). \tag{4.3}$$

We will show an estimate in the opposite direction, showing that for our purposes, $\widehat{N}_\square(n, \delta; W)$ and $N_\square(n, \delta; W)$ are not too different. We begin with the following estimate. We recall that $\overline{W}_k := \mathbb{E}[W \mid \mathcal{B}_{\overline{\mathcal{P}}_k}]$ is obtained by averaging W over squares of side $1/k$, see (3.4).

Lemma 4.1. *Let W be a graphon. If $G \in \mathcal{L}_n$, then there is a graph $\widetilde{G} \in \mathcal{L}_n$ isomorphic to G such that*

$$d_\square(\widetilde{G}, W) \le \delta_\square(G, W) + 2d_\square(W, \overline{W}_n) + \frac{18}{\sqrt{\log_2 n}}. \tag{4.4}$$

Proof. Regard $\overline{W}_n$ as a weighted graph on n vertices, and consider the random graph $G(\overline{W}_n)$ on $[n]$, defined by connecting each pair $\{i, j\}$ of nodes by an edge ij with probability $\overline{W}_n(i/n, j/n)$, independently for different pairs. By [Lov12, Lemma 10.11], with positive probability (actually at least $1 - e^{-n}$),

$$d_\square(G(\overline{W}_n), \overline{W}_n) \le \frac{10}{\sqrt{n}}.$$

Let G' be one realization of $G(\overline{W}_n)$ with

$$d_\square(G', \overline{W}_n) \le \frac{10}{\sqrt{n}}. \tag{4.5}$$

Then, by the triangle inequality and (4.5),

$$\begin{aligned} \delta_\square(G, G') &\le \delta_\square(G, W) + \delta_\square(W, \overline{W}_n) + \delta_\square(\overline{W}_n, G') \\ &\le \delta_\square(G, W) + d_\square(W, \overline{W}_n) + \frac{10}{\sqrt{n}}. \end{aligned} \tag{4.6}$$

Since G and G' both are graphs on $[n]$, we can by [Lov12, Theorem 9.29] permute the labels of G and obtain a graph $\widetilde{G} \in \mathcal{L}_n$ such that

$$d_\square(\widetilde{G}, G') \le \delta_\square(G, G') + \frac{17}{\sqrt{\log_2 n}}. \tag{4.7}$$

Consequently, by the triangle inequality again and (4.5)–(4.7),

$$\begin{aligned} d_\square(\widetilde{G}, W) &\le d_\square(\widetilde{G}, G') + d_\square(G', \overline{W}_n) + d_\square(\overline{W}_n, W) \\ &\le \delta_\square(G, G') + \frac{17}{\sqrt{\log_2 n}} + \frac{10}{\sqrt{n}} + d_\square(W, \overline{W}_n) \\ &\le \delta_\square(G, W) + 2d_\square(W, \overline{W}_n) + \frac{17}{\sqrt{\log_2 n}} + \frac{20}{\sqrt{n}}. \end{aligned}$$

The claim follows for $n > 2^{20}$, say; for smaller n it is trivial since $d_\square(\widetilde{G}, W) \le 1$ for every $\widetilde{G}$. ∎

Lemma 4.2. *For any graphon W, $\delta > 0$ and $n \ge 1$,*

$$N_\square(n, \delta; W) \le n! \widehat{N}_\square(n, \delta + \varepsilon_n; W), \tag{4.8}$$

where $\varepsilon_n := 18/\sqrt{\log_2 n} + 2d_\square(\overline{W}_n, W) \to 0$ as $n \to \infty$.

Proof. By Lemma 4.1, if $G \in \mathcal{L}_n$ and $\delta_\square(G, W) \le \delta$, then $d_\square(\widetilde{G}, W) \le \delta + \varepsilon_n$ for some relabelling $\widetilde{G}$ of G. There are at most $\widehat{N}_\square(n, \delta + \varepsilon_n; W)$ such graphs $\widetilde{G}$ by (4.1), and each corresponds to at most $n!$ graphs G. Finally, note the well-known fact that $d_\square(\overline{W}_n, W) \le \|\overline{W}_n - W\|_{L^1} \to 0$ as $n \to \infty$. ∎

Remark 4.3. The bound (4.4) in Lemma 4.1 is not valid without the term $d_\square(W,\overline{W}_n)$. For a simple example, let n be even and let G be a balanced complete bipartite graph. Further, let $W := W_G(\{nx\},\{ny\})$, where $\{x\}$ denotes the fractional part. (Thus W is obtained by partitioning $[0,1]^2$ into n^2 squares and putting a copy of W_G in each. Furthermore, $W = W_{G'}$ for a blow-up G' of G with n^2 vertices.) Then W is equivalent to W_G, so $\delta_\square(G,W) = 0$. Furthermore, $\overline{W}_n = 1/2$ (the edge density), and it is easily seen that for any relabelling $\widetilde{G}$ of G, $d_\square(\widetilde{G},W) \geq d_\square(\widetilde{G},\overline{W}_n) \geq \frac{1}{8}$. Hence the left-hand side of (4.4) does not tend to 0 as $n\to\infty$; thus the term $d_\square(W,\overline{W}_n)$ is needed.

After these preliminaries, we turn to estimating $\widehat{N}_\square(n,\delta;W)$ and $N_\square(n,\delta;W)$ using $\mathrm{Ent}(W)$.

Lemma 4.4. *For every graphon W and for every $\delta > 0$,*

$$\liminf_{n\to\infty} \frac{\log_2 \widehat{N}_\square(n,\delta;W)}{\binom{n}{2}} \geq \mathrm{Ent}(W).$$

Proof. Consider the random graph $G(n,W) \in \mathcal{L}_n$. As shown in [LS06], $G(n,W) \to W$ almost surely, and thus in probability; in other words, the probabilities $p_n := \mathbb{P}[\delta_\square(G(n,W),W) \leq \delta]$ converge to 1 as $n\to\infty$. Moreover it is shown in [Ald85] and [Jan13a, Appendix D] that

$$\lim_{n\to\infty} \frac{\mathrm{Ent}(G(n,W))}{\binom{n}{2}} = \mathrm{Ent}(W), \tag{4.9}$$

where $\mathrm{Ent}(\cdot)$ denotes the usual entropy of a (discrete) random variable.

Let $I_n := \mathbf{1}_{[\delta_\square(G(n,W),W)\leq\delta]}$ so that $\mathbb{E}[I_n] = p_n$. We have, by simple standard results on entropy,

$$\begin{aligned}
\mathrm{Ent}(G(n,W)) &= \mathbb{E}[\mathrm{Ent}(G(n,W) \mid I_n)] + \mathrm{Ent}(I_n)\\
&= p_n \mathrm{Ent}(G(n,W) \mid I_n = 1) + (1-p_n)\mathrm{Ent}(G(n,W) \mid I_n = 0) + h(p_n)\\
&\leq p_n \log_2 N_\square(n,\delta;W) + (1-p_n)\binom{n}{2} + h(p_n)\\
&\leq \log_2 N_\square(n,\delta;W) + (1-p_n)\binom{n}{2} + 1\\
&= \log_2 N_\square(n,\delta;W) + o\big(n^2\big).
\end{aligned}$$

By Lemma 4.2, this yields

$$\mathrm{Ent}(G(n,W)) \leq \log_2 \widehat{N}_\square(n,\delta+\varepsilon_n;W) + o\big(n^2\big)$$

for some sequence $\varepsilon_n \to 0$. The result follows now from (4.9), if we replace δ by $\delta/2$. ∎

We define, for convenience, for $x \geq 0$,

$$h^*(x) := h\big(\min(x,\tfrac{1}{2})\big); \tag{4.10}$$

thus $h^*(x) = h(x)$ for $0 \le x \le \frac{1}{2}$, and $h^*(x) = 1$ for $x > \frac{1}{2}$. Note that h^* is non-decreasing.

Lemma 4.5. *Let W be a graphon, $n \ge k \ge 1$ be integers and $\delta > 0$. For any equipartition $\mathcal{P}$ of $[n]$ into k sets, we have*

$$\frac{\log_2 \widehat{N}_\square(n,\delta;W)}{n^2} \le \frac{1}{2}\mathrm{Ent}(\mathbb{E}[W \mid \mathcal{B}_\mathcal{P}]) + \frac{1}{2}h^*(4k^2\delta) + 2k^2\frac{\log_2 n}{n^2}.$$

Proof. Denote the sets in $\mathcal{P}$ by $V_1,\dots,V_k \subseteq [n]$ and their sizes by $n_1,\dots,n_k$, and let $I_1,\dots,I_k$ be the subsets in the corresponding partition of $[0,1]$.

Let w_{ij} denote the value of $\mathbb{E}[W \mid \mathcal{B}_\mathcal{P}]$ on $I_i \times I_j$. Suppose that $G \in \mathcal{L}_n$ and $\|W_G - W\|_\square \le \delta$. Let $e(V_i,V_j)$ be the number of edges in G from V_i to V_j when $i \ne j$, and twice the number of edges with both endpoints in V_i when $i = j$. Then

$$e(V_i,V_j) = n^2 \int_{I_i\times I_j} W_G(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

and thus

$$\bigl|e(V_i,V_j) - w_{ij}n_in_j\bigr| = n^2 \left|\int_{I_i\times I_j} (W_G(x,y) - W(x,y))\,\mathrm{d}x\,\mathrm{d}y\right| \le \delta n^2.$$

Hence

$$\left|\frac{e(V_i,V_j)}{n_in_j} - w_{ij}\right| \le \frac{\delta n^2}{n_in_j} \le \delta\left(\frac{n}{\lfloor n/k\rfloor}\right)^2 \le 4k^2\delta. \tag{4.11}$$

Fix numbers $e(V_i,V_j)$ satisfying (4.11), and let N_1 be the number of graphs on $[n]$ with these $e(V_i,V_j)$. By Lemma 3.1, for $i \ne j$, the edges in G between V_i and V_j can be chosen in

$$\binom{n_in_j}{e(V_i,V_j)} \le 2^{n_in_jh(e(V_i,V_j)/n_in_j)} \tag{4.12}$$

number of ways. For $i = j$, the edges in V_i may be chosen in

$$\binom{\binom{n_i}{2}}{\frac{1}{2}e(V_i,V_i)} \le 2^{\binom{n_i}{2}h(\frac{1}{2}e(V_i,V_i)/\binom{n_i}{2})} \le 2^{\frac{1}{2}n_i^2h(e(V_i,V_i)/n_i^2)} \tag{4.13}$$

number of ways, where the second inequality holds because h is concave with $h(0) = 0$ and thus $h(x)/x$ is decreasing.

Consequently, by (4.12) and (4.13),

$$\begin{aligned}\log_2 N_1 &\le \sum_{i<j} n_in_jh\bigl(e(V_i,V_j)/n_in_j\bigr) + \frac{1}{2}\sum_i n_i^2h\bigl(e(V_i,V_i)/n_i^2\bigr)\\ &= \frac{1}{2}\sum_{i,j=1}^k n_in_jh\bigl(e(V_i,V_j)/n_in_j\bigr).\end{aligned}$$

Using (4.11) and (3.1), we obtain

$$\log_2 N_1 \le \frac{1}{2}\sum_{i,j=1}^{k} n_i n_j \big(h(w_{ij}) + h^*(4k^2\delta)\big),$$

and thus

$$\begin{aligned} n^{-2}\log_2 N_1 &\le \frac{1}{2}\sum_{i,j} |I_i||I_j|\big(h(w_{ij}) + h^*(4k^2\delta)\big) \\ &= \frac{1}{2}\operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]) + \frac{1}{2}h^*(4k^2\delta). \end{aligned}$$

Each $e(V_i, V_j)$ may be chosen in at most n^2 ways, and thus the total number of choices is at most n^{2k^2}, and we obtain $\widehat{N}_\square(n,\delta;W) \le n^{2k^2}\max N_1$. Consequently,

$$n^{-2}\log_2 \widehat{N}_\square(n,\delta;W) \le \frac{1}{2}\operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]) + \frac{1}{2}h^*(4k^2\delta) + 2k^2\frac{\log_2 n}{n^2}. \quad \square$$

Lemma 4.6. *Let W be a graphon. Then for any $k \ge 1$, $\delta > 0$ and any equipartition $\mathcal{P}$ of $[0,1]$ into k sets,*

$$\limsup_{n\to\infty} \frac{\log_2 N_\square(n,\delta;W)}{\binom{n}{2}} \le \operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]) + h^*(4k^2\delta).$$

Consequently

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \frac{\log_2 N_\square(n,\delta;W)}{\binom{n}{2}} \le \operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]).$$

Proof. By a suitable measure preserving re-arrangement $\sigma : [0,1] \to [0,1]$, we may assume that $\mathcal{P}$ is the partition $\overline{\mathcal{P}}_k$ into k intervals $((j-1)/k, j/k]$ of length $1/k$.

For every $n > 1$, let $\mathcal{P}_n$ be the corresponding equipartition of $[n]$ into k sets $P_{n1}, \dots, P_{nk}$ where $P_{nj} := \{i : \lfloor (j-1)n/k \rfloor < i \le \lfloor jn/k \rfloor\}$. Note that $\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}_n}]$ converges to $\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]$ almost everywhere as $n \to \infty$, and hence

$$\lim_{n\to\infty} \operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}_n}]) = \operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}}]).$$

Then by Lemmas 4.2 and 4.5, we have, with $\varepsilon_n \to 0$,

$$\begin{aligned} \frac{\log_2 N_\square(n,\delta;W)}{n^2} \le \frac{\log_2(n!)}{n^2} &+ \frac{1}{2}\operatorname{Ent}(\mathbb{E}[W \mid \mathcal{B}_{\mathcal{P}_n}]) \\ &+ \frac{1}{2}h^*\left(4k^2(\delta + \varepsilon_n)\right) + 2k^2\frac{\log_2 n}{n^2} \end{aligned}$$

and the result follows by letting $n \to \infty$. $\square$

We can now show our main lemma. As usual, if A is a set of graph limits, we define $\delta_\square(G, A) := \inf_{W\in A} \delta_\square(G, W)$.

Lemma 4.7. *Let $A \subseteq \widehat{\mathcal{U}}$ be a closed set of graph limits and let*

$$N_\square(n,\delta;A) := \left|\{G \in \mathcal{L}_n : \delta_\square(G,A) \le \delta\}\right|.$$

Then

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\frac{\log_2 N_\square(n,\delta;A)}{\binom{n}{2}} = \lim_{\delta\to 0}\limsup_{n\to\infty}\frac{\log_2 N_\square(n,\delta;A)}{\binom{n}{2}} = \max_{W\in A}\operatorname{Ent}(W). \tag{4.14}$$

Proof. First note that the maximum in the right-hand side of (4.14) exists as a consequence of the semicontinuity of $\operatorname{Ent}(\cdot)$ in Lemma 3.3 (ii) and the compactness of A.

Let $\delta > 0$ and $k \ge 1$. Since A is a compact subset of $\widehat{\mathcal{U}}$, there exists a finite set of graphons $\{W_1,\dots,W_m\} \subseteq A$ such that $\min_i \delta_\square(W,W_i) \le \delta$ for each $W \in A$. Hence

$$N_\square(n,\delta;A) \le \sum_{i=1}^m N_\square(n,2\delta;W_i). \tag{4.15}$$

By (3.5), for each W_i, we can choose an equipartition $\mathcal{P}_i$ of $[0,1]$ into at most k sets such that

$$\|W_i - \mathbb{E}[W_i|\mathcal{B}_{\mathcal{P}_i}]\|_\square \le \frac{4}{\sqrt{\log_2 k}}. \tag{4.16}$$

By (4.15) and Lemma 4.6,

$$\begin{aligned}\limsup_{n\to\infty}\frac{\log_2 N_\square(n,\delta;A)}{\binom{n}{2}} &\le \max_{i\le m}\limsup_{n\to\infty}\frac{\log_2 N_\square(n,2\delta;W_i)}{\binom{n}{2}} \\ &\le \max_{i\le m}\operatorname{Ent}(\mathbb{E}[W_i \mid \mathcal{B}_{\mathcal{P}_i}]) + h^*(8k^2\delta).\end{aligned} \tag{4.17}$$

For each $k \ge 1$, take $\delta = 2^{-k}$ and let $i(k)$ denote the index maximizing $\operatorname{Ent}(\mathbb{E}[W_i|\mathcal{B}_{\mathcal{P}_i}])$ in (4.17); further let $W_k' := W_{i(k)}$ and $W_k'' := \mathbb{E}[W_{i(k)}|\mathcal{B}_{\mathcal{P}_{i(k)}}]$. Thus $W_k' \in A$, and by (4.16)–(4.17),

$$\|W_k' - W_k''\|_\square \le \frac{4}{\sqrt{\log_2 k}} \tag{4.18}$$

and

$$\limsup_{n\to\infty}\frac{\log_2 N_\square(n,2^{-k};A)}{\binom{n}{2}} \le \operatorname{Ent}(W_k'') + h^*(8k^2 2^{-k}). \tag{4.19}$$

Since A is compact, we can select a subsequence such that W_k' converges, and then $W_k' \to W'$ for some $W' \in A$. By (4.18), also $W_k'' \to W'$ in $\mathcal{W}_0$ and thus Lemma 3.3 shows that

$$\limsup_{k\to\infty}\operatorname{Ent}(W_k'') \le \operatorname{Ent}(W'). \tag{4.20}$$

Since $N_\square(n,\delta;A)$ is an increasing function of δ, letting $k\to\infty$, it follows from (4.19) and (4.20) that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{\log_2 N_\square(n,\delta;A)}{\binom{n}{2}} = \lim_{k\to\infty}\limsup_{n\to\infty}\frac{\log_2 N_\square(n,2^{-k};A)}{\binom{n}{2}} \le \operatorname{Ent}(W') \le \max_{W\in A}\operatorname{Ent}(W),$$

which shows that the right-hand side in (4.14) is an upper bound.

To see that the right-hand side in (4.14) also is a lower bound, note that (4.3) implies that for every $W\in A$,

$$N_\square(n,\delta;A)\ge N_\square(n,\delta;W)\ge \widehat{N}_\square(n,\delta;W).$$

The sought lower bound thus follows from Lemma 4.4, which completes the proof. □

5. Proofs of Theorems 1.1–1.8

Proof of Theorem 1.1. Let $\delta>0$. First observe that for sufficiently large n, if $G\in\mathcal{Q}_n$, then $\delta_\square(G,\widehat{\mathcal{Q}})<\delta$. Indeed, if not, then we could find a sequence G_n with $|G_n|\to\infty$ and $\delta_\square(G,\widehat{\mathcal{Q}})\ge\delta$. Then, by compactness, G_n would have a convergent subsequence, but the limit cannot be in $\widehat{\mathcal{Q}}$ which is a contradiction. Consequently for sufficently large n, we have $|\mathcal{Q}_n|\le|\mathcal{Q}_n^L|\le N_\square(n,\delta;\widehat{\mathcal{Q}})$. Thus

$$\limsup_{n\to\infty}\frac{\log_2|\mathcal{Q}_n|}{\binom{n}{2}}\le\limsup_{n\to\infty}\frac{\log_2 N_\square(n,\delta;\widehat{\mathcal{Q}})}{\binom{n}{2}}.$$

The result now follows from the Lemma 4.7. □

Proof of Theorem 1.5. Let W be a graphon representing some $\Gamma\in\widehat{\mathcal{Q}}$ and consider the random graph $G(n,W)\in\mathcal{L}_n$. Since $\mathcal{Q}$ is hereditary, it is easy to see that almost surely $G(n,W)\in\mathcal{Q}_n^L$, see Lemma 3.2 and (1.3) or [Jan13b]. Consequently, letting $\operatorname{Ent}(G(n,W))$ denote the entropy of the random graph $G(n,W)$ (as a random variable in the finite set $\mathcal{Q}_n^L$),

$$\operatorname{Ent}(G(n,W))\le\log_2|\mathcal{Q}_n^L|.$$

Hence, (4.9) and (1.6) show that, for every $W\in\widehat{\mathcal{Q}}$,

$$\liminf_{n\to\infty}\frac{\log_2|\mathcal{Q}_n|}{\binom{n}{2}}=\liminf_{n\to\infty}\frac{\log_2|\mathcal{Q}_n^L|}{\binom{n}{2}}\ge\operatorname{Ent}(W).$$

The result now follows from Theorem 1.1. □

Proof of Theorem 1.6. (i). Let $\delta>0$ and let $B(\delta)=\{G:\delta_\square(G,\Gamma_{\mathcal{Q}})<\delta\}$. The conclusion means that, for any δ, $\mathbb{P}[G_n\in B(\delta)]\to 1$ as $n\to\infty$, i.e.

$$\frac{|\mathcal{Q}_n\cap B(\delta)|}{|\mathcal{Q}_n|}\to 1.$$

If this is not true, then for some $c > 0$ there are infinitely many n with

$$|\mathcal{Q}_n \setminus B(\delta)| \geq c|\mathcal{Q}_n|. \tag{5.1}$$

Consider the graph property $\mathcal{Q}^* := \mathcal{Q} \setminus B(\delta)$. By (5.1) and the assumption (1.7)

$$\limsup_{n\to\infty} \frac{\log_2 |\mathcal{Q}_n^*|}{\binom{n}{2}} = \mathrm{Ent}(\Gamma_{\mathcal{Q}}).$$

Hence Theorem 1.1 shows that $\mathrm{Ent}(\Gamma_{\mathcal{Q}}) \leq \max_{\Gamma\in\widehat{\mathcal{Q}^*}} \mathrm{Ent}(\Gamma)$. So there exists $\Gamma^* \in \widehat{\mathcal{Q}^*}$ such that $\mathrm{Ent}(\Gamma_{\mathcal{Q}}) \leq \mathrm{Ent}(\Gamma^*)$.

On the other hand, $\overline{\mathcal{Q}^*} \subseteq \overline{\mathcal{Q}}$ and $\Gamma_{\mathcal{Q}} \notin \overline{\mathcal{Q}^*}$ so $\widehat{\mathcal{Q}^*} = \overline{\mathcal{Q}^*} \cap \widehat{\mathcal{U}} \subseteq \widehat{\mathcal{Q}} \setminus \{\Gamma_{\mathcal{Q}}\}$, but by assumption $\mathrm{Ent}(\Gamma) < \mathrm{Ent}(\Gamma_{\mathcal{Q}})$ for $\Gamma \in \widehat{\mathcal{Q}} \setminus \{\Gamma_{\mathcal{Q}}\}$. This yields a contradiction which completes the proof of (i).

(ii). The labelled case follows in the same way, now using $\mathcal{Q}_n^L$ and (1.6). □

Proof of Theorem 1.8. Let $\delta > 0$ and let $B_\delta = \{G : \delta_\square(G, \mathrm{supp}(\mu)) < \delta\}$. Then B_δ is an open neighborhood of $\mathrm{supp}(\mu)$ in $\overline{\mathcal{U}}$ and thus the assumption that G_n converges in distribution to μ implies that $\lim_{n\to\infty} \mathbb{P}[G_n \in B_\delta] = 1$. We have, similarly to the proof of Lemma 4.4,

$$\begin{aligned}
\mathrm{Ent}(G_n) &= \mathbb{E}[\mathrm{Ent}(G_n \mid \mathbf{1}_{[G_n\in B_\delta]})] + \mathrm{Ent}(\mathbf{1}_{[G_n\in B_\delta]}) \\
&= \mathbb{P}[G_n \in B_\delta]\mathrm{Ent}(G_n \mid G_n \in B_\delta) + \mathbb{P}[G_n \notin B_\delta]\mathrm{Ent}(G_n|G_n \notin B_\delta) \\
&\qquad + h^*(\mathbb{P}[G_n \in B_\delta]) \\
&\leq \mathrm{Ent}(G_n \mid G_n \in B_\delta) + \mathbb{P}[G_n \notin B_\delta]\binom{n}{2} + 1 \\
&\leq \log_2 N_\square(n, \delta; \mathrm{supp}(\mu)) + \mathbb{P}[G_n \notin B_\delta]\binom{n}{2} + 1.
\end{aligned}$$

Hence using $\lim_{n\to\infty} \mathbb{P}[G_n \notin B_\delta] = 0$,

$$\limsup_{n\to\infty} \frac{\mathrm{Ent}(G_n)}{\binom{n}{2}} \leq \limsup_{n\to\infty} \frac{\log_2 N_\square(n, \delta, \mathrm{supp}(\mu))}{\binom{n}{2}}.$$

The result follows from Lemma 4.7 by letting $\delta \to 0$. □

6. Proof of Theorem 1.9

The stability version of Turán's theorem, due to Erdős and Simonovits [Erd67, Sim68], is equivalent to the following statement for graphons, see [Pik10, Lemma 23] for a detailed proof and further explanations of the connection.

Theorem 6.1 ([Pik10]). *If a graphon W is K_{r+1}-free, then $\iint W \leq 1 - \frac{1}{r}$ with equality if and only if W is equivalent to the graphon W_{K_r}.*

Recall the definition of randomness support and randomness support graphon, see (1.8).

Lemma 6.2. *Let $1 \le r < \infty$. If $\mathcal{Q}$ is a hereditary graph class and $W \in \widehat{\mathcal{Q}}$ has a randomness support graphon that is not K_r-free, then there exists $s \in \{0, 1, \dots, r\}$ such that $W^*_{r,s} \in \widehat{\mathcal{Q}}$. In particular, $\widehat{\mathcal{Q}} \cap R_r \neq \emptyset$.*

Proof. First define $W'(x,y) := W(x,y)$ if $W(x,y) \in \{0,1\}$ and $W'(x,y) := \frac{1}{2}$ if $0 < W(x,y) < 1$. Then $W'(x,y) \in \{0, \frac{1}{2}, 1\}$ for all (x,y). Moreover, W' has the same randomness support as W and it is easily seen that for any graph F, $p(F;W) = 0$ if and only if $p(F;W') = 0$. It follows from Lemma 3.2 that $W' \in \widehat{\mathcal{Q}}$.

Let $x_1, \dots, x_r \in (0,1)$ be chosen at random, uniformly and independently. By assumption, with positive probability, we have $W'(x_i, x_j) = 1/2$ for all pairs $i \neq j$. Choose one such sequence $x_1, \dots, x_r$ such that furthermore $x_1, \dots, x_r$ are distinct and (x_i, x_j) is a Lebesgue point of W' when $i \neq j$; this is possible since the additional conditions hold almost surely. Let m be a positive integer, and set $J_{i,\epsilon} = (x_i - \epsilon, x_i + \epsilon)$ for $\epsilon > 0$ and $i = 1, \dots, r$. If ϵ is sufficiently small, then these intervals are disjoint subintervals of $(0,1)$, and further, if $i \neq j$, then

$$\lambda\left(\left\{(x,y) \in J_{i,\epsilon} \times J_{j,\epsilon} : W'(x,y) = \tfrac{1}{2}\right\}\right) > \left(1 - \tfrac{1}{m}\right) |J_{i,\epsilon}| \cdot |J_{j,\epsilon}|, \qquad (6.1)$$

where λ denotes the Lebesgue measure.

Take such an ϵ and let W'_m be the graphon obtained by scaling the restriction of W' to $(\bigcup_{i=1}^r J_{i,\epsilon}) \times (\bigcup_{i=1}^r J_{i,\epsilon})$ to a graphon in the natural way, by mapping $I_i := (\frac{i-1}{r}, \frac{i}{r}]$ linearly to $J_{i,\epsilon}$ for every $i = 1, \dots, r$. Then

$$p(F;W) = 0 \implies p(F;W') = 0 \implies p(F;W'_m) = 0,$$

for every graph F, and it follows from Lemma 3.2 that $W'_m \in \widehat{\mathcal{Q}}$.

By construction and (6.1), if $i \neq j$,

$$\lambda\left(\left\{(x,y) \in I_i \times I_j : W'_m(x,y) = \tfrac{1}{2}\right\}\right) \ge \left(1 - \tfrac{1}{m}\right) |I_i| \cdot |I_j|. \qquad (6.2)$$

Regard $W'_m|_{I_i \times I_i}$ as a graphon (rescaling I_i to $[0,1]$), and choose a subsequence of W'_m such that $W'_m|_{I_i \times I_i}$ converges for each i to some limit U_i. It then follows from (6.2) that $W'_m \to W^*$ along the subsequence, where $W^* = \frac{1}{2}$ on $I_i \times I_j$ when $i \neq j$ and $W^* = U_i$ on $I_i \times I_i$. Thus $W^* \in \widehat{\mathcal{Q}}$.

Furthermore, for any n, if N is large enough, then by Ramsey's theorem, the random graph $G(N, U_i)$ contains a copy of K_n or its complement $\overline{K_n}$, and hence $G(n, U_i)$ equals K_n or $\overline{K_n}$ with positive probability. It follows easily, using (1.3), that either $p(K_n; U_i) > 0$ for all n or $p(\overline{K_n}; U_i) > 0$ for all n (or both). In the first case we may modify W^* by replacing U_i by the constant 1 on $I_i \times I_i$, and in the second case we may instead replace U_i by 0; using Lemma 3.2 and (1.3), it is easily seen that the modification still belongs to $\widehat{\mathcal{Q}}$. Doing such a modification for each i, we obtain (possibly after a rearrangement of the intervals I_i) a graphon $W^*_{r,s} \in \widehat{\mathcal{Q}}$ for some s. □

Remark 6.3. We allow $r = 1$ in Lemma 6.2. Since no graphon is K_1-free, it then says that if $\mathcal{Q}$ is any (infinite) hereditary class of graphs, then $\widehat{\mathcal{Q}}$ contains some graphon in R_1, i.e., some random-free graphon, and more precisely, at least one of the constant graphons $W^*_{1,0} = 0$ and $W^*_{1,1} = 1$. (The proof above is valid, but may be much simplified in this case. See also [Jan13b].)

Lemma 6.4. *Let $1 \le r < \infty$. If the randomness support graphon of W is K_{r+1}-free, then $\operatorname{Ent}(W) \le 1 - \frac{1}{r}$ with equality if and only if $W \in R_r$ up to equivalence.*

Proof. Let $W'(x,y)$ by the randomness support graphon of W. By assumption, W' is K_{r+1}-free, so by Theorem 6.1, we have $\iint W' \le 1 - \frac{1}{r}$. Moreover since $h(x) \le 1$ always, and $h(x) = 0$ when $x \in \{0,1\}$, we have $h(W(x,y)) \le W'(x,y)$ for all x, y and thus

$$\operatorname{Ent}(W) = \iint h(W) \le \iint W' \le 1 - \frac{1}{r},$$

with equality holding only if $W = \frac{1}{2}$ almost everywhere on its randomness support and W' is equivalent to W_{K_r}, which implies that W is equivalent to a graphon in R_r. (For a rigorous proof of the latter fact, we may use [Pik10, Lemma 23], which implies that W' a.e. equals W_{K_r} up to a measure-preserving bijection of $[0,1]$.)

Conversely, it is obvious that $\operatorname{Ent}(W) = 1 - \frac{1}{r}$ for every $W \in R_r$, see (1.9). □

Proof of Theorem 1.9. Let $r \le \infty$ be defined by (1.11). If $r < \infty$, then every $W \in \widehat{\mathcal{Q}}$ has a randomness support graphon that is K_{r+1}-free and thus $\operatorname{Ent}(W) \le 1 - \frac{1}{r}$ by Lemma 6.4. Moreover, there exists $W \in \widehat{\mathcal{Q}}$ with a randomness support graphon that is not K_r-free and thus by Lemma 6.2 there exists $W' \in \widehat{\mathcal{Q}}$ with $W' \in R_r$, which by (1.9) implies $\operatorname{Ent}(W') = 1 - 1/r$. Hence,

$$\max_{W \in \widehat{\mathcal{Q}}} \operatorname{Ent}(W) = 1 - \frac{1}{r}, \tag{6.3}$$

and by Lemma 6.4 the maximum is attained only for $W \in \widehat{\mathcal{Q}} \cap R_r$. Thus $\widehat{\mathcal{Q}}^* = \widehat{\mathcal{Q}} \cap R_r$. Furthermore, by Lemma 6.2, some $W^*_{r,s} \in \widehat{\mathcal{Q}}$, and $W^*_{t,s} \notin \widehat{\mathcal{Q}}$ for $t > r$ since $\operatorname{Ent}(W^*_{t,s}) = 1 - 1/t > 1 - 1/r$; hence (1.12) holds.

If $r = \infty$, there is for every $s < \infty$ a graphon in $\widehat{\mathcal{Q}}$ whose randomness support graphon is not K_s-free and thus Lemma 6.2 shows that there exists a graphon $W_s \in \widehat{\mathcal{Q}} \cap R_s$. But then W_s converges to the constant graphon $\frac{1}{2}$ in cut norm (and even in L^1) as $s \to \infty$. Thus the constant graphon $\frac{1}{2}$ belongs to $\widehat{\mathcal{Q}}$. Since $\operatorname{Ent}(\frac{1}{2}) = 1$, it follows that $\max_{W \in \widehat{\mathcal{Q}}} \operatorname{Ent}(W) = 1$, i.e., (6.3) holds in the case $r = \infty$ too. Moreover, $\frac{1}{2}$ is the only graphon with entropy 1, see Remark 1.3, and thus $\widehat{\mathcal{Q}}^* = \{\frac{1}{2}\} = R_\infty$.

We have shown (6.3) for any $r \le \infty$, and thus (1.10) follows by Theorem 1.5.

By (6.3), $r = 1$ if and only if $\mathrm{Ent}(W) = 0$ for every $W \in \widehat{\mathcal{Q}}$, i.e., if and only if every $W \in \widehat{\mathcal{Q}}$ is random-free, which by definition means that $\mathcal{Q}$ is random-free.

Finally, if $r = \infty$, then we have established that $\frac{1}{2} \in \widehat{\mathcal{Q}}$. Since $p(F; \frac{1}{2}) > 0$ for every F by (1.1), Lemma 3.2 shows that $\mathcal{Q}$ is the class of all graphs. Hence, every graphon belongs to $\widehat{\mathcal{Q}}$, so (1.12) holds trivially in this case too. ∎

References

[ABBM11] Noga Alon, József Balogh, Béla Bollobás, and Robert Morris. The structure of almost all graphs in a hereditary property. *J. Combin. Theory Ser. B*, 101(2):85–110, 2011.

[Ald85] David J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

[Ale92] V. E. Alekseev. Range of values of entropy of hereditary classes of graphs. *Diskret. Mat.*, 4(2):148–157, 1992.

[BBS04] József Balogh, Béla Bollobás, and Miklós Simonovits. The number of graphs without forbidden subgraphs. *J. Combin. Theory Ser. B*, 91(1):1–24, 2004.

[BBS09] József Balogh, Béla Bollobás, and Miklós Simonovits. The typical structure of graphs without given excluded subgraphs. *Random Structures Algorithms*, 34(3):305–318, 2009.

[BBSS09] József Balogh, Béla Bollobás, Michael Saks, and Vera T. Sós. The unlabelled speed of a hereditary graph property. *J. Combin. Theory Ser. B*, 99(1):9–19, 2009.

[BBW00] József Balogh, Béla Bollobás, and David Weinreich. The speed of hereditary properties of graphs. *J. Combin. Theory Ser. B*, 79(2):131–156, 2000.

[BBW01] József Balogh, Béla Bollobás, and David Weinreich. The penultimate rate of growth for graph properties. *European J. Combin.*, 22(3):277–289, 2001.

[BCL+08] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[BCL10] Christian Borgs, Jennifer Chayes, and László Lovász. Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.*, 19(6):1597–1619, 2010.

[Bol07] Béla Bollobás. Hereditary and monotone properties of combinatorial structures. In *Surveys in combinatorics 2007*, volume 346 of *London Math. Soc. Lecture Note Ser.*, pages 1–39. Cambridge Univ. Press, Cambridge, 2007.

[BR09] Béla Bollobás and Oliver Riordan. Metrics for sparse graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Math. Soc. Lecture Note Ser.*, pages 211–287. Cambridge Univ. Press, Cambridge, 2009.

[BT97] Béla Bollobás and Andrew Thomason. Hereditary and monotone properties of graphs. In *The mathematics of Paul Erdős, II*, volume 14 of *Algorithms Combin.*, pages 70–78. Springer, Berlin, 1997.

[CD11] Sourav Chatterjee and Persi Diaconis. Estimating and understanding exponential random graph models. Preprint, 2011. arXiv:1102.2650.

[CV11] Sourav Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.*, 32(7):1000–1017, 2011.

[DHJ08] Persi Diaconis, Susan Holmes, and Svante Janson. Threshold graph limits and random threshold graphs. *Internet Math.*, 5(3):267–320 (2009), 2008.

[DJ08] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)*, 28(1):33–61, 2008.

[EKR76] P. Erdős, D. J. Kleitman, and B. L. Rothschild. Asymptotic enumeration of K_n-free graphs. In *Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II*, pages 19–27. Atti dei Convegni Lincei, No. 17. Accad. Naz. Lincei, Rome, 1976.

[Erd67] P. Erdős. Some recent results on extremal problems in graph theory. Results. In *Theory of Graphs (Internat. Sympos., Rome, 1966)*, pages 117–123 (English); 124–130 (French). Gordon and Breach, New York, 1967.

[FK99] Alan Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[HN12] Hamed Hatami and Serguei Norine. The entropy of random-free graphons and properties. Preprint, 2012. arXiv:1205.3529.

[Jan13a] Svante Janson. Graphons, cut norm and distance, rearrangements and coupling. *New York J. Math. Monographs*, 24:1–76, 2013.

[Jan13b] Svante Janson. Graph limits and hereditary properties. Preprint, 2013. arXiv:1102.3571v2.

[KGK86] Jan Kratochvíl, Miroslav Goljan, and Petr Kučera. String graphs. *Rozpravy Československé Akad. Věd Řada Mat. Přírod. Věd*, 96(3):96, 1986.

[KPR85] Ph. G. Kolaitis, H. J. Prömel, and B. L. Rothschild. Asymptotic enumeration and a 0-1 law for m-clique free graphs. *Bull. Amer. Math. Soc. (N.S.)*, 13(2):160–162, 1985.

[KPR87] Ph. G. Kolaitis, H. J. Prömel, and B. L. Rothschild. K_{l+1}-free graphs: asymptotic structure and a 0-1 law. *Trans. Amer. Math. Soc.*, 303(2):637–671, 1987.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[LS07] László Lovász and Balázs Szegedy. Szemerédi's lemma for the analyst. *Geom. Funct. Anal.*, 17(1):252–270, 2007.

[LS10] László Lovász and Balázs Szegedy. Regularity partitions and the topology of graphons. In *An irregular mind*, volume 21 of *Bolyai Soc. Math. Stud.*, pages 415–446. János Bolyai Math. Soc., Budapest, 2010.

[Pik10] Oleg Pikhurko. An analytic approach to stability. *Discrete Math.*, 310(21):2951–2964, 2010.

[PT06] János Pach and Géza Tóth. How many ways can one draw a graph? *Combinatorica*, 26(5):559–576, 2006.

[Sim68] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.

School of Computer Science, McGill University, Montreal, Canada.
E-mail address: hatami@cs.mcgill.ca

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.
E-mail address: svante.janson@math.uu.se

University of Toronto, Department of Mathematics, St George St. 40, Toronto, ON, M5R 2E4, Canada.
E-mail address: szegedy@math.toronto.edu