



ON A QUESTION OF VERA T. SÓS ABOUT SIZE FORCING OF GRAPHONS

O. COOLEY^{1,*,\dagger}, M. KANG^{2,\dagger} and O. PIKHURKO^{3,\ddagger}

¹Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

e-mail: oliver.cooley@ist.ac.at

²Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria

e-mail: kang@math.tugraz.at

³Mathematics Institute and DIMAP, University of Warwick, Coventry CV4 7AL, UK

e-mail: O.Pikhurko@warwick.ac.uk

(Received May 26, 2021; accepted March 25, 2022)

Abstract. The k -sample $\mathbb{G}(k, W)$ from a graphon $W: [0, 1]^2 \rightarrow [0, 1]$ is the random graph on $\{1, \dots, k\}$, where we sample $x_1, \dots, x_k \in [0, 1]$ uniformly at random and make each pair $\{i, j\} \subseteq \{1, \dots, k\}$ an edge with probability $W(x_i, x_j)$, with all these choices being mutually independent. Let the random variable $X_k(W)$ be the number of edges in $\mathbb{G}(k, W)$.

Vera T. Sós asked in 2012 whether two graphons U, W are necessarily weakly isomorphic if the random variables $X_k(U)$ and $X_k(W)$ have the same distribution for every integer $k \geq 2$. This question when one of the graphons W is a constant function was answered positively by Endre Csóka and independently by Jacob Fox, Tomasz Łuczak and Vera T. Sós. Here we investigate the question when W is a 2-step graphon and prove that the answer is positive for a 3-dimensional family of such graphons.

We also present some related results.

1. Introduction

Graphons (that is, measurable symmetric functions $[0, 1]^2 \rightarrow [0, 1]$) have recently found many important applications in other areas, such as the limit

* Corresponding author.

\dagger Supported by Austrian Science Fund (FWF) Grant I3747.

\ddagger Supported by ERC Advanced Grant 101020255 and Leverhulme Research Project Grant RPG-2018-424.

An extended abstract of this paper appeared in the Proceedings of the European Conference on Combinatorics, Graph Theory and Applications (EuroComb 2021), CRM Research Perspectives, Springer.

Key words and phrases: graphon, k -sample, graphon forcing, graph container.

Mathematics Subject Classification: 05C99, 05C80.

theory of dense graphs (Lovász et al. [4,20,21]), large deviation principles for random graphs (Chatterjee and Varadhan [5]), property testing in computer science (Lovász and Szegedy [22]), etc. We refer the reader to the monograph by Lovász [18] for an introduction.

The k -sample $\mathbb{G}(k, W)$ from a graphon W is the random graph on $[k] := \{1, \dots, k\}$ obtained by sampling $x_1, \dots, x_k \in [0, 1]$ uniformly at random and making each pair $\{i, j\} \subseteq [k]$ an edge with probability $W(x_i, x_j)$, with all these choices being mutually independent. The (*homomorphism*) *density* $t(F, W)$ of a graph F on $[k]$ in W is the probability that $E(F) \subseteq E(\mathbb{G}(k, W))$, that is, every adjacent pair in F is also adjacent in $\mathbb{G}(k, W)$. Equivalently, we can define

$$(1) \quad t(F, W) := \int_{[0,1]^k} \prod_{\{i,j\} \in E(F)} W(x_i, x_j) dx_1 \cdots dx_k.$$

Let us call two graphons U and W *weakly isomorphic* if the random graphs $\mathbb{G}(k, U)$ and $\mathbb{G}(k, W)$ have the same distribution for every $k \in \mathbb{N}$. This is equivalent to $t(H, U) = t(H, W)$ for every connected graph H , because we can recover the distribution of $\mathbb{G}(k, W)$ using the following inclusion-exclusion formula

$$(2) \quad \mathbb{P}(\mathbb{G}(k, W) = G) = \sum_{\substack{F \supseteq G \\ V(F)=[k]}} (-1)^{|E(F) \setminus E(G)|} t(F, W),$$

for every graph G on $[k]$, and replacing each $t(F, W)$ by the product of the densities of the components of F (see Lemma 10 later).

Borgs, Chayes and Lovász [3] showed that all graphons in the weak isomorphism class of W can, roughly speaking, be obtained from W by applying measure-preserving transformations of the variables. (See also Diaconis and Janson [8] who derived this result from the Aldous–Hoover Theorem [1,14] by noting a connection to exchangeable arrays.) This gives an analogue of the classical moment problem, where each $t(F, W)$ can be thought of as the “ F -th moment” of W .

A *graphon parameter* f is a function that assigns to each graphon W a real number or a real vector $f(W)$ such that $f(W) = f(U)$ whenever U and W are weakly isomorphic. We say that a family $(f_i)_{i \in I}$ of graphon parameters *forces* a graphon W if every graphon U with $f_i(U) = f_i(W)$ for every $i \in I$ is weakly isomorphic to W . For example, the famous result of Chung, Graham and Wilson [6] on p -quasirandom graphs can be stated in this language as follows.

THEOREM 1 (Chung, Graham and Wilson [6]). *The constant- p graphon is forced by $t(K_2, \cdot)$ and $t(C_4, \cdot)$, that is, by the edge and 4-cycle densities.*

Call a family $(f_i)_{i \in I}$ of graphon parameters *forcing* if it forces every graphon W . For example, the densities $t(F, \cdot)$, where F ranges over all connected graphs, form a forcing family (by (2) and Lemma 10).

A lot of effort has gone into investigating whether a graphon W is forced by much less information than the densities of all graphs. Graphons that are forced by a finite set of graph densities are called *finitely forcible* and their systematic study was initiated by Lovász, Sós and Szegedy [19,23], motivated by quasirandom graphs and extremal graph theory. As one would expect, finitely forcible graphons are “rare”: they form a meagre subset of the space of all graphons ([23, Theorem 7.12]).

The authors are not aware of any results where a substantially smaller set of parameters than the densities of all connected graphs is shown to be forcing. Vera T. Sós [29] posed some questions in this direction, and in particular considered the following problem. For a graphon W and an integer $k \in \mathbb{N}$, let $X_k(W) := |E(\mathbb{G}(k, W))|$ be the size of, i.e. number of edges in, the k -sample $\mathbb{G}(k, W)$ from W . We identify the random variable $X_k(W)$ with the vector of probabilities $\mathbb{P}(X_k(W) = i)$ for $0 \leq i \leq \binom{k}{2}$, viewing it as a graphon parameter. Let \mathcal{W}_S be the family of graphons W that are forced by the sequence $(X_k(W))_{k \in \mathbb{N}}$, i.e. by the distributions of sizes of samples from W .

QUESTION 2 (Size Forcing Question (Sós [29])). *Is every graphon in \mathcal{W}_S ?*

Noga Alon (unpublished, see [7]) and independently Jakub Sliacan [28] proved that the constant $\frac{1}{2}$ graphon is in the family \mathcal{W}_S . Then Endre Csóka [7] and independently Jacob Fox, Tomasz Łuczak and Vera T. Sós [10] proved that the constant- p graphon is in the family \mathcal{W}_S for any $p \in (0, 1)$. A natural next step would be to try to determine whether $W \in \mathcal{W}_S$ when W is a *2-step graphon*, that is, we have a partition of $[0, 1]$ into two measurable sets A and B such that W is constant on each of the sets A^2 , B^2 and $(A \times B) \cup (B \times A)$. By replacing W by a weakly isomorphic graphon, we can assume that $A = [0, a)$ and $B = [a, 1]$ are intervals. Thus we need four parameters to describe a 2-step graphon: the measure of A as well as the three possible values of W .

Unfortunately, we were not able to prove that $W \in \mathcal{W}_S$ for every 2-step graphon W . However, we could prove this for the following 3-dimensional set of graphons.

THEOREM 3. *Let W be the 2-step graphon with parts $A := [0, a)$ and $B := [a, 1]$ such that its values on A^2 , $(A \times B) \cup (B \times A)$ and B^2 are respectively 0 , $p \in (0, 1]$ and $q \in (0, 1]$. If $(1 - a)q \leq (1 - 2a)p$, then $W \in \mathcal{W}_S$.*

Let us mention here that since $X_k(1 - W)$ has the same distribution as $\binom{k}{2} - X_k(W)$ (by taking complements), if W is forced by some sub-family of $(X_k)_{k \in \mathbb{N}}$, then so is $1 - W$. Thus if W lies in \mathcal{W}_S , then so does $1 - W$.

We can answer Question 2 for some other families of 2-step graphons W . Here we present two further examples (Theorems 4 and 5) where a *finite* set of some natural real-valued parameters suffices.

The first is motivated by the result of Csóka [7] who in fact proved that the constant- p graphon is forced by X_4 alone. The following theorem proves a similar claim to that of Csóka [7] for the limit of balanced quasirandom bipartite graphs, namely that it is forced by the edge distribution of its 5-sample.

THEOREM 4. *Let $p \in [0, 1]$ and let W be the graphon which is 0 on $[0, 1/2]^2 \cup [1/2, 1]^2$ and p everywhere else. Then W is forced by X_5 alone.*

Let the *independence ratio* $\alpha(W)$ of a graphon W be the supremum of the measure of $A \subseteq [0, 1]$ such that $W(x, y) = 0$ for a.e. $(x, y) \in A^2$. As easily follows from the results of Hladký, Hu and Piguet [12, Lemma 2.4], the supremum is in fact a maximum (that is, it is attained by some A). Also, the *clique ratio* $\omega(W) := \alpha(1 - W)$ is defined as the maximum measure of $A \subseteq [0, 1]$ with W being 1 a.e. on A^2 .

THEOREM 5. *Given $a, p \in [0, 1]$, set $A := [0, a]$ and $B := [a, 1]$, and let W be the graphon which is 0 on A^2 , 1 on B^2 , and p everywhere else. Then W is forced by (α, ω, X_4) .*

By using a basic version of the container method, we show that the value of α (and thus of ω) is determined by any infinite subsequence of $(X_k)_{k \in \mathbb{N}}$. More precisely, the following holds.

THEOREM 6. *For every graphon W , it holds that*

$$\alpha(W) = \lim_{k \rightarrow \infty} (\mathbb{P}(X_k(W) = 0))^{1/k}.$$

Hladký and Rocha [13] defined and studied graphon versions of various graph parameters, including the independence ratio $\alpha(W)$. In particular, they investigated how these parameters can be related to graph densities. Our Theorem 6, by relating $\alpha(W)$ to graph densities, fills one missing entry in [13, Table 1].

By combining Theorems 5 and 6, we directly obtain the following result.

COROLLARY 7. *Let W be a graphon as in Theorem 5 (that is, W is 0 on $[0, a]^2$, 1 on $[a, 1]^2$, and p everywhere else). Then $W \in \mathcal{W}_S$.*

Call a family \mathcal{F} of graphs *forcing* if the corresponding family of parameters $(t(F, \cdot))_{F \in \mathcal{F}}$ is forcing. Sós [29] also asked if one can find substantially smaller forcing families than taking all connected graphs. We show that two natural examples, namely the family of all cycles and the family of all complete bipartite graphs, do not suffice.

PROPOSITION 8. *The family of all connected graphs with at most one cycle is not forcing. In particular, the family of all cycles is not forcing.*

PROPOSITION 9. *For every integer d , the family of all graphs of diameter at most d is not forcing. In particular, the family of all complete bipartite graphs is not forcing.*

Also, let us mention here the somewhat related result of Shapira and Tyomkin [27] that the constant- p graphon is not forced by $(t(K_k, \cdot))_{k \in \mathbb{N}}$, that is, by clique densities.

Paper overview. The paper is arranged as follows. In Section 2 we recall various standard notation that we will use, in particular notation related to graphons. In Section 3 we collect some easy preliminary results which we will apply later. Theorem 4 is then proved in Section 4.

In Section 5, we first present an auxiliary graph result (Theorem 15) which relates small independent sets to large almost independent sets, and prove this result using the container method. We subsequently use this result to prove Theorem 6.

Sections 6 and 7 contain the proofs of Theorems 3 and 5, respectively. Finally, Propositions 8 and 9 are proved in Section 8. Section 9 contains concluding remarks.

2. Notation

Here we present some notation that is used in this paper.

Let k be a non-negative integer. We denote the k -th falling factorial of a real number r by

$$(r)_k := r(r-1) \cdots (r-k+1).$$

For a set X , let

$$\binom{X}{k} := \{Y \subseteq X : |Y| = k\}$$

consist of all subsets of X of size k . The characteristic function $\mathbb{1}_X$ of X assumes value 1 on X and 0 everywhere else. We may abbreviate an unordered pair $\{x, y\}$ to xy .

We will use the following notation related to graphs. Let $G = (V, E)$ be a graph. Its complement is $\overline{G} := (V, \binom{V}{2} \setminus E)$. For $A \subseteq V(G)$, its neighbourhood

$$N(A) := \{x : \exists y \in A \text{ with } xy \in E\}$$

consists of vertices that send at least one edge to A . For graphs H_1 and H_2 , their disjoint union $H_1 \sqcup H_2$ is obtained by taking the union of vertex-disjoint copies of these graphs (with no edges across).

We will also be using the following special graphs. The k -clique K_k is the graph on $[k]$ in which every two vertices are adjacent. The k -path P_k is the path on $[k]$ that visits vertices $1, \dots, k$ in this order. The k -cycle C_k is the cycle on $[k]$ that visits vertices $1, \dots, k$ in this cyclic order. Also, $K_{k,\ell}$ denotes the complete bipartite graph with parts of sizes k and ℓ .

Let $\mathcal{G}_{k,m}$ consist of isomorphism classes of all graphs with at most k vertices and exactly m edges that do not contain any isolated vertices. For example, $\mathcal{G}_{5,3} = \{K_3, P_4, P_3 \sqcup K_2, K_{1,3}\}$.

Let us also collect some definitions related to graphons.

The unit interval $[0, 1]$ is by default equipped with the Lebesgue measure, denoted by λ . When making any statements about subsets of $[0, 1]$, we usually mean that they hold up to a set of measure 0. We will often use (Fubini–)Tonelli’s theorem (see e.g. [9, Theorem 14.2]) whose main part states, informally speaking, that non-negative measurable functions can be integrated in any order of variables. In particular, when working with $t(F, W)$ as the value of the integral in (1), we can integrate the variables x_1, \dots, x_k in any order. We will therefore often change the order of integration without mentioning this theorem explicitly.

Let W be a graphon and let $A \subseteq [0, 1]$ be a measurable subset. The *degree* (resp. *A-degree*) of $x \in [0, 1]$ is

$$\text{deg}^W(x) := \int_0^1 W(x, y) \, dy$$

(resp. $\text{deg}_A^W(x) := \int_A W(x, y) \, dy$). The degree is defined for a.e. $x \in [0, 1]$ by a part of Tonelli’s theorem. We call W *p-regular* if $\text{deg}^W(x) = p$ for a.e. $x \in [0, 1]$. The *codegree* (resp. *A-codegree*) of $(x, y) \in [0, 1]^2$ is

$$\text{codeg}^W(x, y) := \int_0^1 W(x, z)W(z, y) \, dz$$

(resp. $\text{codeg}_A^W(x, y) := \int_A W(x, z)W(z, y) \, dz$). One can view $\text{codeg}^W(x, y)$ as the density of 2-edge paths that connect x and y .

When discussing $\mathbb{G}(k, W)$, it will often be convenient to view it as a graph whose vertex set consists of the sampled points $x_1, \dots, x_k \in [0, 1]$ (which are pairwise distinct with probability 1). Thus a phrase like “ x_i is adjacent to x_j ” will be a shorthand for “ i is adjacent to j in $\mathbb{G}(k, W)$ ”, etc.

For a graph G on $[k]$, its graphon W_G is the graphon which is 1 on $[\frac{i-1}{k}, \frac{i}{k}] \times [\frac{j-1}{k}, \frac{j}{k}]$ for each edge $ij \in E(G)$, and 0 everywhere else. In other words, we partition $[0, 1]$ into k intervals of length $1/k$ and let W_G be the k -step $\{0, 1\}$ -valued graphon that naturally encodes the adjacency relation of G .

3. Some auxiliary results

Let us present here some known or easy auxiliary results that we need in this paper.

LEMMA 10. *For any graphon W and for any graphs H_1 and H_2 , we have*

$$t(H_1 \sqcup H_2, W) = t(H_1, W) t(H_2, W).$$

PROOF. Assume that $V(H_1 \sqcup H_2) = [k]$ and let $A_1 \cup A_2 = [k]$ be the partition into the vertex sets of H_1 and H_2 . The lemma follows by observing that the subgraphs induced by A_1 and A_2 in $\mathbb{G}(k, W)$ are independent of each other and, up to a relabelling of vertices, are distributed as $\mathbb{G}(|A_1|, W)$ and $\mathbb{G}(|A_2|, W)$. \square

LEMMA 11. *If W is a p -regular graphon and F' is obtained from a graph F by attaching a pendant edge then*

$$t(F', W) = p t(F, W).$$

PROOF. We can assume that $V(F) = [k]$ and that the added edge is $\{k, k+1\}$. When computing $t(F', W)$ as the integral over $(x_1, \dots, x_{k+1}) \in [0, 1]^{k+1}$ as in (1), we can first integrate over x_{k+1} . The only factor that depends on x_{k+1} is $W(x_k, x_{k+1})$. Its integral is p for a.e. x_k , so integrating out x_{k+1} amounts to multiplying by p (and replacing F' by F in (1)), proving the lemma. \square

The following result implicitly appears in Csóka [7]. For completeness, we present its proof.

LEMMA 12. *Let integers k and m satisfy $1 \leq m \leq \binom{k}{2}$. Then for every graphon W we have*

$$(3) \quad \mathbb{E}((X_k(W))_m) = \sum_{F \in \mathcal{G}_{k,m}} c_{k,F} t(F, W),$$

where $c_{k,F} > 0$ is $m!$ times the number of graphs on $[k]$ that, after discarding isolated vertices, are isomorphic to F .

PROOF. Let \mathcal{X} consist of all ordered m -tuples $(\{s_i, t_i\}_{i=1}^m)$ of pairwise distinct pairs from $\binom{[k]}{2}$. Thus, for example, its size $|\mathcal{X}|$ is the falling factorial $(\binom{k}{2})_m$. For $F \in \mathcal{G}_{k,m}$, let \mathcal{X}_F consist of those sequences in \mathcal{X} that give a graph isomorphic to F after we discard all isolated vertices. Clearly, the sets \mathcal{X}_F when F ranges over $\mathcal{G}_{k,m}$ partition \mathcal{X} .

The left-hand side of (3) is the expectation of the number of sequences in \mathcal{X} all of whose m pairs are edges when we take the k -sample $\mathbb{G}(k, W)$.

This expectation can be written as the sum over all $(\{s_i, t_i\})_{i=1}^m \in \mathcal{X}$ of the probability that each $\{s_i, t_i\}$ is an edge. The last probability is exactly $t(F, W)$ where F is the unique graph of $\mathcal{G}_{k,m}$ with $(\{s_i, t_i\})_{i=1}^m \in \mathcal{X}_F$. The lemma follows by observing that each $F \in \mathcal{G}_{k,m}$ appears exactly $|\mathcal{X}_F| = c_{k,F}$ times this way. \square

Here is a useful consequence of this lemma.

LEMMA 13. *Let U and W be graphons such that $X_k(U)$ and $X_k(W)$ have the same distributions for some $k \geq 3$. Then $t(K_2, U) = t(K_2, W)$, $t(K_2 \sqcup K_2, U) = t(K_2 \sqcup K_2, W)$ and $t(P_3, U) = t(P_3, W)$.*

PROOF. By applying Lemma 12 with $m = 1$ to U and W , we get that

$$\binom{k}{2} t(K_2, U) = \mathbb{E}(X_k(U)) = \mathbb{E}(X_k(W)) = \binom{k}{2} t(K_2, W).$$

Thus U and W have the same density of K_2 and, by Lemma 10, of $K_2 \sqcup K_2$. If $k \geq 4$, then Lemma 12 with $m = 2$ gives that

$$\begin{aligned} 2! \frac{\binom{k}{3}}{2} \cdot t(P_3, U) &= \mathbb{E}((X_k(U))_2) - 2! \frac{\binom{k}{4}}{8} \cdot t(K_2 \sqcup K_2, U) \\ &= \mathbb{E}((X_k(W))_2) - 2! \frac{\binom{k}{4}}{8} \cdot t(K_2 \sqcup K_2, W) = 2! \frac{\binom{k}{3}}{2} \cdot t(P_3, W), \end{aligned}$$

finishing the proof (since $\binom{k}{3} \neq 0$). The same calculation applies for $k = 3$ except that the $K_2 \sqcup K_2$ term is absent. \square

We will also need the following bipartite analogue of Theorem 1. While a rather elementary proof by passing to finite graphs that converge to U is possible (along the same lines as the original proof of Chung, Graham and Wilson [6], see also e.g. [18, Theorem 11.62]), we present a proof that, while requiring some analytic background, deals directly with graphons.

LEMMA 14. *Let A and B be sets of measure a and b respectively that partition $[0, 1]$. (Thus $a + b = 1$.) Let $p \in [0, 1]$. Let U be a graphon taking value 0 on $A^2 \cup B^2$ such that $t(K_2, U) = 2abp$ and $t(C_4, U) = 2a^2b^2p^4$. Then $U(x, y) = p$ for a.e. $(x, y) \in (A \times B) \cup (B \times A)$.*

PROOF. Assume that $a, b \in (0, 1)$ as otherwise there is nothing to do.

Using that U is 0 on $A^2 \cup B^2$, we have that $t(C_4, U)$ (that is, the density of the 4-cycle in U) is $2 \int_{A^2} \text{codeg}_B^U(x, y)^2 dx dy$, where the factor 2 comes from having a partition into two equiprobable events, namely that $x_1, x_3 \in A$ and that $x_2, x_4 \in A$. On the other hand, by applying the Cauchy-Schwarz inequality twice, we get that

$$2a^2b^2p^4 = 2 \int_{A^2} (\text{codeg}_B^U(x, y))^2 dx dy \geq \frac{2}{a^2} \left(\int_{A^2} \text{codeg}_B^U(x, y) dx dy \right)^2$$

$$= \frac{2}{a^2} \left(\int_B (\deg_A^U(z))^2 dz \right)^2 \geq \frac{2}{a^2 b^2} \left(\int_B \deg_A^U(z) dz \right)^4 = 2a^2 b^2 p^4.$$

Thus we have equality. This implies that $\text{codeg}_B^U(x, y) = bp^2$ for a.e. $(x, y) \in A^2$ and that $\deg_A^U(x) = ap$ for a.e. $x \in B$. The same argument applies when we count C_4 from the other side, giving that $\text{codeg}_A^U(x, y) = ap^2$ for a.e. $(x, y) \in B^2$ and $\deg_B^U(x) = bp$ for a.e. $x \in A$.

View U as the integral kernel operator defined by

$$(Uf)(x) := \int_0^1 U(x, y)f(y) dy, \quad \text{for } f \in L^2([0, 1]) \text{ and } x \in [0, 1].$$

The Cauchy–Schwarz (or Hölder’s) inequality gives that $Uf \in L^2([0, 1])$, so U is an operator on $L^2([0, 1])$. This operator is self-adjoint (since the function U is symmetric) and compact (as an integral operator with its kernel U being a bounded and thus square-integrable function on $[0, 1]^2$, see e.g. [11, Example 3 of Section 2.16]). The Spectral Decomposition Theorem (see e.g. [11, Theorem 5.1 of Section 4.5]) gives that $L^2([0, 1])$ has an orthonormal basis of eigenfunctions $(f_i)_{i \in \mathbb{N}}$ with the corresponding eigenvalues $(\lambda_i)_{i \in \mathbb{N}}$ such that $\lambda_i \rightarrow 0$ as $i \rightarrow \infty$. Then it follows that

$$(4) \quad U(x, y) = \sum_{i=1}^{\infty} \lambda_i f_i(x) f_i(y) \quad \text{for a.e. } (x, y) \in [0, 1]^2.$$

Consider the composition of U with itself. This is again an integral kernel operator and we identify it with its kernel

$$(U \circ U)(x, y) := \int U(x, z)U(z, y) dz = \text{codeg}^U(x, y), \quad x, y \in [0, 1].$$

By above, $U \circ U$ is a.e. the two-step graphon of value 0 on $(A \times B) \cup (B \times A)$, value bp^2 on A^2 and value ap^2 on B^2 . (Recall that U is 0 on $A^2 \cup B^2$.) Thus, for every $f \in L^2([0, 1])$, its image $(U \circ U)(f)$ is a function which is constant on A and on B . Thus, as an operator, $U \circ U$ has rank at most 2. On the other hand, each of the characteristic functions $\mathbb{1}_A$ and $\mathbb{1}_B$ is an eigenvector of $U \circ U$ with eigenvalue abp^2 . We conclude that the operator $U \circ U$ has exactly one non-zero eigenvalue abp^2 of multiplicity 2.

Clearly, the same functions $(f_i)_{i \in \mathbb{N}}$ and the squares $(\lambda_i^2)_{i \in \mathbb{N}}$ give a spectral decomposition of $U \circ U$. Thus the rank of U is also 2 and its non-zero eigenvalues are in $\{-p\sqrt{ab}, p\sqrt{ab}\}$. Using the established values for the (constant) degrees in U across the partition $A \cup B$, we have that $U\mathbb{1}_A = ap\mathbb{1}_B$ and $U\mathbb{1}_B = bp\mathbb{1}_A$. Consider the functions $h_1 := \sqrt{b}\mathbb{1}_A - \sqrt{a}\mathbb{1}_B$ and

$h_2 := \sqrt{b} \mathbb{1}_A + \sqrt{a} \mathbb{1}_B$ that are orthogonal to each other and have L^2 -norm $\sqrt{2ab}$. We have

$$Uh_1 = \sqrt{b}U\mathbb{1}_A - \sqrt{a}U\mathbb{1}_B = ap\sqrt{b}\mathbb{1}_B - bp\sqrt{a}\mathbb{1}_A = -p\sqrt{ab}h_1,$$

and similarly $Uh_2 = p\sqrt{ab}h_2$. Thus, up to relabelling, we have $\lambda_1 = -p\sqrt{ab}$ and $\lambda_2 = p\sqrt{ab}$ (and $\lambda_i = 0$ for all $i \geq 3$). Moreover, by the 1-dimensionality of the eigenspaces for $\lambda_1 \neq \lambda_2$, it holds that $f_i = \pm h_i/\sqrt{2ab}$ for $i = 1, 2$. By (4), we have that

$$\begin{aligned} U(x, y) &= \lambda_1 f_1(x) f_1(y) + \lambda_2 f_2(x) f_2(y) \\ &= \frac{-p\sqrt{ab}h_1(x)h_1(y) + p\sqrt{ab}h_2(x)h_2(y)}{2ab}. \end{aligned}$$

Thus U is constant on $A \times B$ and its value there can be shown to be p by using the definition of h_1 and h_2 (or by our assumption that $t(K_2, U) = 2abp$), giving the lemma. \square

4. Proof of Theorem 4

Recall that W is the 2-step graphon which is the limit of balanced bipartite p -quasirandom graphs. Let U be an arbitrary graphon such that the distribution of $X_5(U)$ is the same as the distribution of $X_5(W)$. Let us denote this common distribution by X_5 . We will be iteratively proving a sequence of claims about U , until the derived information is enough to conclude that U must be weakly isomorphic to W . Assume that $p \neq 0$, as the constant-0 graphon is clearly forced by X_5 being 0 with probability 1.

CLAIM 4.1. *The graphon U is $(p/2)$ -regular, that is, the degree function $\deg^U(x) = \int_0^1 U(x, y) dy$ is equal to $p/2$ for a.e. $x \in [0, 1]$.*

PROOF. Consider the random variable $D := \deg^U(x)$, where $x \in [0, 1]$ is uniformly random. Lemma 13 shows that $\mathbb{E}(D) = t(K_2, U)$ equals $t(K_2, W) = p/2$ and $\mathbb{E}(D^2) = t(P_3, U)$ equals $t(P_3, W) = p^2/4$. This shows that

$$0 \leq \mathbb{E}((D - p/2)^2) = \mathbb{E}(D^2) - p\mathbb{E}(D) + p^2/4 = 0,$$

and therefore $D = p/2$ with probability 1. \square

Thus, by Lemma 11, we have the following.

CLAIM 4.2. *If $t(H, U) = t(H, W)$ for some graph H , then $t(H', U) = t(H', W)$ for any graph H' that is obtained from H by adding a pendant edge. In particular, $t(F, U) = t(F, W)$ for any forest F . \square*

We therefore obtain the following.

CLAIM 4.3. *It holds that $t(H, U) = t(H, W)$ for each H in $\mathcal{G}_{5,3} = \{K_{1,3}, P_4, P_2 \sqcup K_2, K_3\}$. In particular, $t(K_3, U) = t(K_3, W) = 0$ and thus $t(H, U) = 0$ for every graph H that contains a triangle.*

PROOF. For the first three graphs of $\mathcal{G}_{5,3}$, this follows immediately from Claim 4.2.

Lemma 12, when applied with $m = 3$ to each of U and W , gives the same linear relation (with all coefficients non-zero) relating the densities of the four graphs in $\mathcal{G}_{5,3}$. Since we have already established that $t(H, U) = t(H, W)$ for every $H \in \mathcal{G}_{5,3} \setminus \{K_3\}$, we must also have that $t(K_3, U) = t(K_3, W)$. \square

Thus by Claims 4.2 and 4.3, for each $m \in [10]$, any graph in $\mathcal{G}_{5,m}$ that has different densities in U and W must belong to \mathcal{H}_m , which we define to consist of $H \in \mathcal{G}_{5,m}$ such that H is triangle-free and H is not a forest. Clearly, every such H must have an induced cycle of length 4 or 5, which makes it easy to enumerate all graphs in \mathcal{H}_m .

We have that $\mathcal{H}_4 = \{C_4\}$ consists only of the 4-cycle and thus Lemma 12 gives that

$$(5) \quad t(C_4, U) = t(C_4, W).$$

Next, \mathcal{H}_5 consists only of the 5-cycle C_5 and C'_4 , a 4-cycle with a leaf attached to one of its vertices. By (5) and Claim 4.2, we have that $t(C'_4, U) = t(C'_4, W)$. Thus, by Lemma 12, the other graph in \mathcal{H}_5 , namely the 5-cycle, must also have the same density in U as in W .

Since \mathcal{H}_6 contains only the complete bipartite graph $K_{2,3}$, Lemma 12 gives that $t(K_{2,3}, U) = t(K_{2,3}, W)$.

We now consider the random variable

$$Z := \text{codeg}^U(x, y) = \int_0^1 U(x, z)U(z, y) dz,$$

the density of copies of P_2 which have x, y as endpoints, where x and y are chosen uniformly and independently from $[0, 1]$. The following identities show that the first three moments of Z remain the same if we replace U by W :

$$\begin{aligned} \mathbb{E}(Z) &= t(P_3, U) = t(P_3, W) = p^2/4, \\ \mathbb{E}(Z^2) &= t(C_4, U) = t(C_4, W) = p^4/8, \\ \mathbb{E}(Z^3) &= t(K_{2,3}, U) = t(K_{2,3}, W) = p^6/16. \end{aligned}$$

We now observe that

$$\mathbb{E}(Z(Z - p^2/2)^2) = \mathbb{E}(Z^3) - p^2 \mathbb{E}(Z^2) + \frac{p^4}{4} \mathbb{E}(Z) = \frac{p^6}{16} - p^2 \frac{p^4}{8} + \frac{p^2}{4} \cdot \frac{p^2}{4} = 0.$$

Since $Z(Z - p^2/2)^2 \geq 0$ deterministically, we have that $Z \in \{0, p^2/2\}$ with probability 1. By $\mathbb{E}(Z) = p^2/4$, we conclude that

$$\mathbb{P}(Z = 0) = \mathbb{P}(Z = p^2/2) = \frac{1}{2}.$$

Thus almost all pairs come in two types. Let C consist of those $(x, y) \in [0, 1]^2$ for which $\text{codeg}^U(x, y) = p^2/2$. Its complement consists a.e. of pairs with zero codegree. Since the measure of $C \subseteq [0, 1]^2$ is $1/2$, we have $\int_0^1 \text{deg}_C(x) dx = 1/2$, where $\text{deg}_C(x)$ denotes the measure of

$$N_C(x) := \{y : (x, y) \in C\}, \quad \text{for } x \in [0, 1].$$

We know from Claim 4.2 that $t(P_5, U) = t(P_5, W)$ where P_5 is the path with 5 vertices. One can compute $t(P_5, U)$ as follows. Recall that P_5 visits its vertices $1, \dots, 5$ in this order. First sample $x_1, x_3, x_5 \in [0, 1]$ and then pick common neighbours x_2 and x_4 of x_1x_3 and x_3x_5 respectively. The measure of choices of $(x_2, x_4) \in [0, 1]^2$ is $(p^2/4)^2$ if $x_1x_3, x_3x_5 \in C$ and 0 otherwise (apart from a null set of (x_1, x_3, x_5)). The same argument applies to $t(P_5, W)$. Since $p \neq 0$, the measure of $(x_1, x_3, x_5) \in [0, 1]^3$ with x_1x_3 and x_3x_5 in C must be the same as the analogous quantity for W , that is, $1/4$. Thus, by the Cauchy–Schwarz inequality, we have

$$\frac{1}{4} = \int_0^1 (\text{deg}_C(x_3))^2 dx_3 \geq \left(\int_0^1 \text{deg}_C(x_3) dx_3 \right)^2 = \frac{1}{4}.$$

We conclude that deg_C is the constant- $1/2$ function a.e.

For a.e. $x \in [0, 1]$, the set $N_C(x)$ is independent in U . Indeed, the density of C_5 in U can be written as the integral over $x_3 \in [0, 1]$ and then over $x_1, x_5 \in N_C(x_3)$ of $U(x_1, x_5)(p^2/2)^2$. On the other hand, we know that $t(C_5, U) = t(C_5, W) = 0$, giving the claim.

Pick a typical $x \in [0, 1]$, that is, with $A' := N_C(x)$ being measurable, independent and of measure $1/2$. By above and the $(p/2)$ -regularity of U , each vertex of A' has degree $p/2$ in U and almost all these edges connect A' to its complement $B' := [0, 1] \setminus A'$. Thus

$$\frac{p}{4} = \int_{A'} \text{deg}_{B'}^U(x) dx = \int_{A' \times B'} U(x, y) dx dy \leq \frac{1}{2} t(K_2, U),$$

where the factor $1/2$ arises because integrating over $B' \times A'$ would give exactly the same result, since U is symmetric. But $t(K_2, U) = t(K_2, W) = p/2$, so these are all the edges, that is, B' is an independent set.

By applying a measure-preserving transformation to U , we can assume that $A' = A$ and $B' = B$. These sets are independent in both U and W . Since W takes constant value p on $A \times B$, Lemma 14 gives that U is also p a.e. on $A \times B$, finishing the proof of Theorem 4.

5. Graphons with large density of independent k -sets

We will need the following auxiliary result which, informally speaking, states that if a graph has many independent sets of large but fixed size k then the graph has a large almost independent set. Let $\mathcal{I}(G)$ denote the family of all independent sets in a graph G and let

$$\mathcal{I}_k(G) := \{I \in \mathcal{I}(G) : |I| = k\}$$

consist of all independent sets of size k .

THEOREM 15. *For every $\delta > 0$ there exists $\varepsilon > 0$ such that for any $k \geq 1/\varepsilon$ there exists n_0 such that for every graph G on $n \geq n_0$ vertices and every real α , if $|\mathcal{I}_k(G)| \geq (\alpha - \varepsilon)^k \binom{n}{k}$, then there exists $A \subseteq V(G)$ with $|A| \geq (\alpha - \delta)n$ and $e(G[A]) \leq \delta n^2$.*

PROOF. Given $\delta > 0$, choose sufficiently small $\varepsilon > 0$; in particular, assume that $\varepsilon < \delta/2$. Given any $k \geq 1/\varepsilon$, let n be sufficiently large. Let a graph $G = (V, E)$ and a real α satisfy the assumptions of the lemma. Assume that $\alpha > \delta$ as otherwise we can trivially let A be the empty set.

We use a basic version of the container method that was introduced in high generality independently by Balogh, Morris and Samotij [2] and by Saxton and Thomason [26], and whose roots go back to Kleitman and Winston [16,17] and Sapozhenko [24,25]. Roughly speaking, we will encode each independent set I of G by a very small set $T \subseteq I$ together with a decoding procedure that produces a container $C = C(T)$ that necessarily contains I and spans few edges. Then we take for A a container $C(T(I))$ of the maximal size over all choices of $I \in \mathcal{I}_k(G)$.

Formally, we proceed as follows. Assume that $V = [n]$ with the natural order.

Take any $I \in \mathcal{I}(G)$. Enumerate $I = \{i_1 < \dots < i_m\}$ using the natural order on $V = [n]$. The *encoding* procedure produces $T = T(I)$ as follows. Initially, let $T := \emptyset$ and $j := 1$. Iterate the following step. Given the current values of $j \leq m$ and $T \subseteq \{i_1, \dots, i_{j-1}\}$, add i_j into T if and only if

$$(6) \quad |N(T \cup \{i_j\})| \geq |N(T)| + \delta n/2,$$

that is, i_j has at least $\delta n/2$ neighbours outside of $N(T)$. Then increase j by 1 and, if the new j is still at most m , repeat the iteration step.

Let $T = T(I)$ be the final set T . Since only the vertices of I were considered for inclusion into T , we have that $T \subseteq I$. Also, every time we add a vertex to T , the size of the neighbourhood $N(T)$ increases by at least $\delta n/2$. Thus

$$(7) \quad |T| \leq 2/\delta.$$

Now, let us describe the *decoding* procedure which constructs the *container* $C(T)$ for any independent set $T \subseteq V$ of the graph G . Let $t := |T|$. Enumerate $V \setminus T = \{v_1 < \dots < v_{n-t}\}$, again using the natural order on $[n]$. Initially, let $C := T$ and $j := 1$. Repeat the following step given the current values of $j \leq n - t$ and $C \subseteq T \cup \{v_1, \dots, v_{j-1}\}$: include v_j into C if and only if $v_j \notin N(T)$ and

$$(8) \quad |N(T_j \cup \{v_j\})| < |N(T_j)| + \delta n/2, \quad \text{where } T_j := \{v \in T : v < v_j\}.$$

Then increase j by 1 and, if $j \leq n - t$, repeat the iteration step.

By construction, the final set $C = C(T)$ contains T and is disjoint from $N(T)$. Also, in the notation of (8), each vertex v_j of $C \setminus T$ has fewer than $\delta n/2$ neighbours in $V \setminus N(T_j)$. Note that the last set contains $V \setminus N(T) \supseteq C$ (since $T_j \subseteq T$ and $C \cap N(T) = \emptyset$). Thus C spans at most $|T|n + |C|\delta n/2$ edges. This is at most $2n/\delta + \delta n^2/2 < \delta n^2$ if T satisfies (7).

Let us also show that

$$(9) \quad I \subseteq C(T(I)), \quad \text{for every independent set } I \text{ of } G.$$

Let $T := T(I)$. Since $T \subseteq C(T)$, it remains to show that every $i_s \in I \setminus T$ belongs to $C(T)$, where $i_1 < i_2 < \dots$ enumerate all elements of I . The reason why i_s was not included into T when it was considered at the appropriate encoding step must be that (6) fails for $j = s$, that is, i_s adds fewer than $\delta n/2$ new neighbours when added to $\{v \in T : v < i_s\}$. This is exactly the statement in (8). Also, since $I \supseteq T$ is an independent set, we have that $i_s \notin N(T)$. Thus $i_s \in C(T)$ by the definition of the decoding procedure, proving (9) as desired.

By (7) and (9), we get the following upper bound on the number of independent sets of size exactly k :

$$|\mathcal{I}_k(G)| \leq \sum_{t=0}^{\lfloor 2/\delta \rfloor} \sum_{T \in \mathcal{I}_t(G)} \binom{|C(T) \setminus T|}{k-t}.$$

Fix an index t , between 0 and $\lfloor 2/\delta \rfloor$, such that the t -th summand is at least the average value, which is at least $1/(2/\delta + 1)$ times $|\mathcal{I}_k(G)|$. Given this t , let A be a maximum-size container $C(T)$ over all independent t -sets $T \subseteq V$. Then

$$(10) \quad \frac{1}{2/\delta + 1} (\alpha - \varepsilon)^k \binom{n}{k} \leq \binom{n}{t} \binom{|A| - t}{k - t}.$$

The set A , as some container $C(T)$ for a set T with $|T| \leq 2/\delta$, spans at most δn^2 edges in G . Thus, in order to finish the proof of the theorem,

we have to show that $|A| \geq (\alpha - \delta)n$. When n tends to infinity (with $\delta \gg \varepsilon \geq 1/k$ fixed and $t \leq 2/\delta$ bounded), the inequality in (10) gives that $|A| \rightarrow \infty$ and, in fact, $|A| \geq ((\delta/3)^{1/k} + o(1))c_t n$, where $c_t := ((\alpha - \varepsilon)^k / \binom{k}{t})^{\frac{1}{k-t}}$. It is enough to show that e.g. $c_t \geq (\alpha - \delta/2)$. The last inequality is equivalent to

$$\left(\frac{\alpha - \varepsilon}{\alpha - \delta/2}\right)^k \geq \binom{k}{t} (\alpha - \delta/2)^{-t},$$

which holds for all large k because the left-hand side grows exponentially in k , while the right-hand side grows at most polynomially. (Recall that $t \leq 2/\delta$ is bounded.)

Thus the set A has all required properties. \square

For $k \in \mathbb{N}$, let

$$\alpha_k(W) := \mathbb{P}(X_k(W) = 0)$$

be the probability that the k -sample $\mathbb{G}(k, W)$ has no edges. We have

$$\alpha_{k+m}(W) \leq \alpha_k(W) \alpha_m(W), \quad \text{for all } k, m \in \mathbb{N},$$

because the subgraphs of $\mathbb{G}(k+m, W)$ spanned by the first k and the last m vertices are independent and distributed as $\mathbb{G}(k, W)$ and $\mathbb{G}(m, W)$ respectively. Thus, by the Fekete Lemma, the limit

$$(11) \quad \alpha_\infty(W) := \lim_{k \rightarrow \infty} (\alpha_k(W))^{1/k}$$

exists. Clearly, $\alpha_\infty(W)$ remains the same if we replace W by any weakly isomorphic graphon.

In order to prove Theorem 6, which states in the above notation that $\alpha(W) = \alpha_\infty(W)$, we need to present some definitions and results related to measure theoretic aspects of graphons from [18, Chapters 8 and 13]. Let U and W be graphons. We define the *cut-norm*

$$\|U - W\|_\square := \sup_{A, B \subseteq [0,1]} \left| \int_{A \times B} (U(x, y) - W(x, y)) \, dx \, dy \right|,$$

where the supremum is taken over all pairs of measurable subsets of $[0, 1]$. For a measure-preserving function $\phi: [0, 1] \rightarrow [0, 1]$, the *pull-back* U^ϕ of U along ϕ is defined by

$$U^\phi(x, y) := U(\phi(x), \phi(y)), \quad x, y \in [0, 1].$$

It is routine to see that U^ϕ is a graphon which is weakly isomorphic to U . The *cut-distance* is defined as

$$(12) \quad \delta_\square(U, W) := \inf_\phi \|U^\phi - W\|_\square,$$

where the infimum is taken over all invertible measure-preserving maps $\phi: [0, 1] \rightarrow [0, 1]$. See [18, Section 8.2] for more details and, in particular, [18, Theorem 8.13] for some alternative definitions that give the same distance. It can be easily verified that δ_{\square} is a pseudo-metric on the space of graphons. Moreover, two graphons are weakly isomorphic if and only if they are at cut-distance 0, see e.g. [18, Theorem 13.10].

PROOF OF THEOREM 6. The inequality $\alpha_{\infty}(W) \geq \alpha(W)$ is easy. Indeed, pick an independent set $A \subseteq [0, 1]$ in W of measure $\lambda(A) = \alpha(W)$ (which exists by [12, Lemma 2.4]) and observe that the probability of seeing no edges in the k -sample $\mathbb{G}(k, W)$ is at least $\lambda(A)^k$, the probability that all vertices land in A .

Let us show the converse inequality $\alpha_{\infty}(W) \leq \alpha(W)$. Let $\alpha := \alpha_{\infty}(W)$ and assume that $\alpha > 0$ as otherwise there is nothing to prove. Do the following for every $m \in \mathbb{N}$. Let $\varepsilon > 0$ be sufficiently small, in particular to satisfy Theorem 15 for $\delta := 1/m$. By (11), pick $k \geq 1/\varepsilon$ such that $\alpha_k(W) \geq 4(\alpha - \varepsilon)^k$. Let n be sufficiently large and take the n -sample $G \sim \mathbb{G}(n, W)$. Let W_G be its graphon, that is, W_G is the n -step $\{0, 1\}$ -valued graphon that encodes the adjacency relation of G .

It is easy to see that the mean of $\alpha_k(W_G)$ over $G \sim \mathbb{G}(n, W)$ is exactly $\alpha_k(W)$. We claim that in fact $\alpha_k(W_G)$ is concentrated around this mean, for which we apply Azuma's inequality (see e.g. [15, Theorem 2.25]). Observe that $\alpha_k(W_G)$ can only change by at most k/n if a vertex of G is altered, i.e. the vertex-exposure martingale revealing G and tracking $\alpha_k(W_G)$ is (k/n) -Lipschitz. Setting $t := 2(\alpha - \varepsilon)^k$, Azuma's inequality states that

$$\begin{aligned} \mathbb{P}(\alpha_k(W_G) \leq 2(\alpha - \varepsilon)^k) &\leq \mathbb{P}(\alpha_k(W_G) \leq \mathbb{E}(\alpha_k(W_G)) - t) \\ &\leq \exp\left(\frac{-t^2}{2n(k/n)^2}\right) = o(1), \end{aligned}$$

where asymptotics are as $n \rightarrow \infty$.

It is also known that, as $n \rightarrow \infty$, the probability that the cut-distance between W_G and W is more than $o(1)$ is at most $o(1)$, specifically (see e.g. [18, Lemma 10.16])

$$\mathbb{P}(\delta_{\square}(W_G, W) > 22/\sqrt{\log n}) \leq \exp(-n/(2 \log n)).$$

Thus, for large enough n , there is a graph G on $[n]$ with $\alpha_k(W_G) \geq 2(\alpha - \varepsilon)^k$ and $\delta_{\square}(W, W_G) \leq \delta/2$ because $\mathbb{G}(n, W)$ satisfies each of these properties with probability $1 - o(1)$ as $n \rightarrow \infty$. Since $|\mathcal{I}_k(G)|/\binom{n}{k} = \alpha_k(W_G) + o(1)$, Theorem 15 applies to the graph G for large enough n and returns a set A' of vertices of size at least $(\alpha - \delta)n$ spanning at most δn^2 edges. Let $A := \bigcup_{i \in A'} \left[\frac{i-1}{n}, \frac{i}{n}\right)$ be the subset of $[0, 1]$ corresponding to A' when

we pass from G to its graphon W_G . Take an invertible measure-preserving map $\phi: [0, 1] \rightarrow [0, 1]$ such that $\|W^\phi - W_G\|_\square < \delta$. Defining $S_m := \phi(A)$ to be the image of the set A , we obtain that

$$\begin{aligned} \int_{S_m^2} W(x, y) \, dx \, dy &= \int_{A^2} W^\phi(x, y) \, dx \, dy \\ &\leq \int_{A^2} W_G(x, y) \, dx \, dy + \|W^\phi - W_G\|_\square \leq 2\delta = 2/m. \end{aligned}$$

We now proceed as in [12]. Recall that a sequence of functions f_1, f_2, \dots in $L^\infty([0, 1], \lambda)$, the dual space of $L^1([0, 1], \lambda)$, *weak-** converges to $f \in L^\infty([0, 1], \lambda)$ if

$$(13) \quad \lim_{n \rightarrow \infty} \int_0^1 f_n(x)g(x) \, dx = \int_0^1 f(x)g(x) \, dx, \quad \text{for every } g \in L^1([0, 1], \lambda).$$

By the Sequential Banach–Alaoglu Theorem (see e.g. [30, Theorem 1.9.14]), the sequence of the characteristic functions of the sets S_m viewed as elements of $L^\infty([0, 1], \lambda)$ has a subsequence that weak- $*$ converges to some function $f \in L^\infty([0, 1], \lambda)$. Note that $f(x) \geq 0$ for λ -a.e. $x \in [0, 1]$. Indeed, letting $g := \mathbb{1}_X$ be the characteristic function of the measurable set $X := \{x \in [0, 1] : f(x) < 0\}$, we get from (13) that

$$0 \geq \int_X f \, d\lambda = \int_0^1 fg \, d\lambda = \lim_{m \rightarrow \infty} \int_0^1 \mathbb{1}_{S_m} g \, d\lambda \geq 0,$$

from which it follows that $\lambda(X) = 0$. Likewise, we obtain that $f \leq 1$ a.e. on $[0, 1]$.

Let the *support* of a function $g: [0, 1] \rightarrow \mathbb{R}$ be the set

$$\text{supp}(g) := \{x \in [0, 1] : g(x) \neq 0\}.$$

In fact, [12, Lemma 2.4] states that, for any graphon W , the support of the weak- $*$ limit of the characteristic functions of W -independent sets is W -independent. Thus $S := \text{supp}(f)$, the support of f , is an independent set in W . By the definition of weak- $*$ convergence and the fact that $\|f\|_\infty \leq 1$, we have that

$$\begin{aligned} \alpha(W) \geq \lambda(S) &\geq \int_0^1 f(x) \, dx = \lim_{m \rightarrow \infty} \int_0^1 \mathbb{1}_{S_m}(x) \, dx \\ &= \lim_{m \rightarrow \infty} \lambda(S_m) = \alpha_\infty(W). \end{aligned}$$

This shows that $\alpha(W) = \alpha_\infty(W)$, proving Theorem 6. \square

6. Proof of Theorem 3

Recall that W is the 2-step graphon with steps A and B which is 0 on A^2 , p on $A \times B$ and q on B^2 , and where $A = [0, a)$. Let U be an arbitrary graphon such that for every $k \in \mathbb{N}$ the distributions of $X_k(U)$ and $X_k(W)$ are the same; let us denote this random variable by X_k . We have to show that U is weakly isomorphic to W . Assume that $a \in (0, 1)$ as otherwise W is weakly isomorphic to a 1-step graphon and the conclusion follows from the results in [7].

By Theorem 6 we conclude that $\alpha(U) = \lim_{k \rightarrow \infty} (\mathbb{P}(X_k = 0))^{1/k} = \alpha(W)$. Thus, by [12, Lemma 2.4], there is a set A' of measure a with U being 0 on $A' \times A'$ a.e. By taking a measure-preserving Borel isomorphism $\phi: [0, 1] \rightarrow [0, 1]$ with $\phi(A') = A$ and replacing U with U^ϕ (which is weakly isomorphic to U), we can assume that $A' = A$. Recall that $\deg_A^U(x) := \int_A U(x, y) dy$ for $x \in [0, 1]$.

CLAIM 6.1. *For almost every $x \in B$, we have $\deg_A^U(x) \geq ap$.*

PROOF OF CLAIM. If the claim is false, then by the continuity of measure, there is $\varepsilon > 0$ such that the measure of $B' := \{x \in B : \deg_A^U(x) \leq ap - \varepsilon\}$ is at least ε . Take k sufficiently large. Let us lower bound $\alpha_k(U)$, the probability that the k -sample $\mathbb{G}(k, U)$ spans no edges. Recall that we sample uniform $x_1, \dots, x_k \in [0, 1]$ and then make each pair ij an edge with probability $U(x_i, x_j)$, with all choices being mutually independent. With probability a^k , all elements x_i belong to A (when almost surely $\mathbb{G}(k, U)$ has no edges). A disjoint event is that x_1 belongs to $B' \subseteq B$, which has probability is at least $\varepsilon > 0$. Conditioned on this event, the probability of having no edges is at least

$$\left(\int_A (1 - U(x_1, y)) dy \right)^{k-1} = (a - \deg_A^U(x_1))^{k-1} \geq (a(1 - p) + \varepsilon)^{k-1},$$

since, when ignoring null sets, it is enough that all other $k - 1$ vertices belong to A and are all non-adjacent to x_1 .

Thus $\alpha_k(U) \geq a^k + \varepsilon(a(1 - p) + \varepsilon)^{k-1}$. For W , it is easy to write an explicit formula, where i denotes the number of sampled vertices that belong to B :

$$\begin{aligned} \alpha_k(W) &= \sum_{i=0}^k \binom{k}{i} (1 - a)^i a^{k-i} (1 - p)^{(k-i)i} (1 - q)^{\binom{i}{2}} \\ &= a^k + k(1 - a)a^{k-1}(1 - p)^{k-1} + \dots \end{aligned}$$

Note that the first term a^k matches that for U . Of course, we have $\alpha_k(U) = \mathbb{P}(X_k = 0) = \alpha_k(W)$. Thus a desired contradiction, namely that $\alpha_k(U) > \alpha_k(W)$, will follow if we show that for every $i \in [k]$,

$$(14) \quad \varepsilon(a(1-p) + \varepsilon)^{k-1} > k \cdot \binom{k}{i} (1-a)^i a^{k-i} (1-p)^{(k-i)i} (1-q)^{\binom{i}{2}}.$$

Informally speaking, if i is small, then the main terms are $(a(1-p) + \varepsilon)^k$ versus $(a(1-p)^i)^k$; otherwise either $(1-p)^{(k-i)i}$ or $(1-q)^{\binom{i}{2}}$ (and thus the right-hand side of (14)) is very small. Formally, given $\varepsilon > 0$ as above, fix a large constant $M \gg 1/\varepsilon$ and then let $k \rightarrow \infty$. If $1 \leq i \leq M$ then the ratio of the right-hand side to the left-hand side of (14) is at most

$$O\left(\left(\frac{a(1-p)}{a(1-p) + \varepsilon}\right)^k \cdot k^{M+1}\right) = o(1).$$

If $M < i \leq k$, then since (slightly crudely) $\max\left((k-i)i, \binom{i}{2}\right) \geq ki/4$ the ratio is at most

$$\max(1-p, 1-q)^{ik/4} \cdot \frac{k^{i+1}}{\varepsilon(a(1-p) + \varepsilon)^k} = o(1),$$

where the last estimate holds since $p, q \neq 0$ by assumption. This proves (14) for all large k and finishes the proof of the claim. \square

Let U' be the graphon obtained from U by averaging it over $(A \times B) \cup (B \times A)$ and over B^2 . That is, U' is the 2-step graphon with parts A and B which assumes value 0 on A^2 , value $q' := \frac{1}{(1-a)^2} \int_{B^2} U(x, y) dx dy$ on B^2 , and value

$$(15) \quad p' := \frac{1}{a(1-a)} \int_{A \times B} U(x, y) dx dy = \frac{1}{a(1-a)} \int_B \deg_A^U(y) dy \geq p$$

on $(A \times B) \cup (B \times A)$, where we applied Claim 6.1 in (15). Consider the density of P_3 , the path visiting vertices 1, 2, 3 in this order. Its density, say in U , can be written as $t(P_3, U) = \int (\deg^U(x_2))^2 dx_2$. Clearly, when we pass from U to U' then the degrees inside A (resp. degrees inside B) are all replaced by their average value. Thus the average of $(\deg(x))^2$ over x in A (resp. B) does not increase by the Cauchy–Schwarz inequality. By adding up these two averages weighted by a and $1-a$ respectively, we get the average of $(\deg(x))^2$ over $x \in [0, 1]$, which is the density of P_3 . Thus $t(P_3, U) \geq t(P_3, U')$. The W -degrees of x in A and B are the constants $(1-a)p$ and $ap + (1-a)q$ respectively. These constants for U' are $(1-a)p'$ and $ap' + (1-a)q'$ respectively. Since $p' \geq p$ by (15), we have that $(1-a)p' \geq (1-a)p$. Furthermore, since U' and W have the same edge density (which

can be computed as the convex combination of the average degrees in A and B weighted by a and $1 - a$), we have that $ap' + (1 - a)q' \leq ap + (1 - a)q$. Thus,

$$(1 - a)p' \geq (1 - a)p \geq ap + (1 - a)q \geq ap' + (1 - a)q',$$

where the middle inequality is an assumption of the theorem. We see that when we pass from U' to W , the degrees get more even (with the average staying the same) and so the density of P_3 , the average of $(\deg(x))^2$, does not increase. However, the density of P_3 is determined by X_3 by Lemma 13. Thus the degree functions of U , U' and W coincide a.e. Thus $p' = p$, $q' = q$,

$$(16) \quad \deg_A^U(x) = ap, \quad \deg_B^U(x) = (1 - a)q \text{ for a.e. } x \in B, \\ \deg_B^U(y) = (1 - a)p \text{ for a.e. } y \in A.$$

CLAIM 6.2. *For almost every $(x, y) \in B^2$, we have that $\text{codeg}_A^U(x, y) = ap^2$. (Recall that we denote $\text{codeg}_A^U(x, y) := \int_A U(x, z)U(y, z) dz$.)*

PROOF OF CLAIM. Suppose first that $\text{codeg}_A^U(x, y) > ap^2$ for some set of $(x, y) \in B^2$ of positive measure. Then by the continuity of measure there exists $\varepsilon > 0$ such that the measure of

$$B' := \{(x, y) \in B^2 : \text{codeg}_A^U(x, y) \geq ap^2 + \varepsilon\}$$

is at least ε .

When we compute $\mathbb{P}(X_k \leq 1)$ using U or W , the k -tuples of vertices that have at most one point in B contribute the same amount. For example, let us condition on x_i being the unique sampled vertex that belongs to B . Then each other $x_\ell \in A$ is adjacent to x_i with probability $\frac{1}{a} \deg_A^U(x_i)$ which is equal to $\frac{1}{a} \deg_A^W(x_i)$ by (16). Moreover, these choices for different choices of ℓ are independent of each other, both in U and in W . Thus any particular adjacency pattern of $x_i \in B$ to the other $k - 1$ vertices from A has the same conditional probabilities in U and in W .

Consider the remaining contribution to $\mathbb{P}(X_k \leq 1)$, i.e. when at least two sampled vertices belong to B . First, take U . With probability at least ε , the pair (x_1, x_2) belongs to the set B' . Conditioned on this, each other vertex x_ℓ is adjacent to neither x_1 nor x_2 with probability

$$\int_A (1 - U(x_1, y))(1 - U(x_2, y)) dy = a - 2ap + \text{codeg}_A^U(x_1, x_2),$$

with these choices being mutually independent for different values of ℓ . This contributes at least $\varepsilon(a(1 - p)^2 + \varepsilon)^{k-2}$ to $\mathbb{P}(X_k \leq 1)$. On the other hand,

an explicit summation formula can be written for W . First assuming that $p, q \neq 1$, we have by above that

$$\begin{aligned} & \varepsilon(a(1-p)^2 + \varepsilon)^{k-2} \\ & \leq \sum_{i=2}^k \binom{k}{i} a^{k-i} (1-a)^i (1-p)^{(k-i)i} (1-q)^{\binom{i}{2}} \left(1 + \binom{i}{2} \frac{q}{1-q} + (k-i)i \frac{p}{1-p} \right). \end{aligned}$$

As before by taking $k \rightarrow \infty$ and looking at the cases $i = O(1)$ and $i \gg 1$ separately, one can argue that $\varepsilon(a(1-p)^2 + \varepsilon)^{k-2}$ is strictly larger than k times the maximum term in the sum, giving a contradiction. If either p or q is 1, the inequality must be rewritten to avoid dividing by zero, and in fact the only non-zero terms come from $i = 2$ (if $q = 1$) or $i = k$ (if $p = 1$). It is then easy to obtain the necessary contradiction directly.

Thus $\text{codeg}_A^U(x, y) \leq ap^2$ for a.e. $(x, y) \in B^2$. The integral of $\text{codeg}_A(x, y)$ over all $(x, y) \in B^2$ can be written as the integral over $z \in A$ of $(\text{deg}_B(z))^2$. By (16), the latter integral is the same for U as for W . Thus

$$(1-a)^2 ap^2 \geq \int_{B^2} \text{codeg}_A^U(x, y) \, dx \, dy = \int_{B^2} \text{codeg}_A^W(x, y) \, dx \, dy = (1-a)^2 ap^2,$$

and the first integrand must be ap^2 a.e. on B^2 , proving the claim. \square

It follows that the density of triangles with one vertex in A and two vertices in B is the same in U as in W . Indeed, first sample two vertices x, y in B , which are adjacent in U and in W with the same probability q by (16), and observe that the probability of a vertex from A being adjacent to both is exactly $\frac{1}{a} \text{codeg}_A(x, y)$. Also, U and W have zero density of triangles with at least two vertices in A . Since U and W have the same triangle density, namely $\mathbb{P}(X_3 = 3)$, they must have the same density of triangles that lie inside B .

Thus the density of triangles with any given partition of their vertices between A and B is the same for U as for W . Since the degrees in U are the same as the degrees in W by (16), this allows us to conclude by a version of Lemma 11 that the density of K'_3 , the triangle with a pendant edge, is the same in U as in W . (Indeed, this is true even if we specify, relative to A and B , where the vertices of the triangle in K'_3 lie.) Lemma 12 applied to $\mathcal{G}_{4,4} = \{C_4, K'_3\}$ gives that U and W have the same density of 4-cycles.

Claim 6.2 implies that the density of the 4-cycle (which we assume to visit the vertices 1, 2, 3, 4 in this order) conditioned on $x_1, x_3 \in A$ and $x_2, x_4 \in B$ is p^4 , the fourth power of edge density p between A and B . (Indeed, to sample such a 4-cycle, we can first sample uniform $(x_2, x_4) \in B^2$; then we independently sample two vertices of A , each being adjacent to both x_2 and x_4 with probability $(\frac{1}{a} \text{codeg}_A(x_2, x_4))^2$.) Thus if we let U' be obtained

from U by letting it be 0 on B^2 , then all assumptions of Lemma 14 are satisfied. The lemma gives that U' (and thus U) assumes the constant value p between A and B .

Thus we know all about U except its values on B^2 . Our knowledge about U is enough to compute the density of all types of 4-cycles except those entirely inside B . For example, if the sampled 4-cycle is to visit parts B, B, B, A in this order, then we can first sample a 3-vertex path x_1, x_2, x_3 in B (knowing its density since $\deg_B^U(x) = \deg_B^W(x) = (1 - a)q$ for almost every $x \in B$) and then use Claim 6.2 to see that, conditioned on $x_4 \in A$, the probability of x_4 being adjacent to both x_1 and x_3 is exactly p^2 (same as in W). Since the graphons U and W have the same overall density of 4-cycles, they have the same density of 4-cycles inside B . This relative density of 4-cycles is the fourth power of the relative edge density since W is constant on B . By the Chung–Graham–Wilson Theorem (Theorem 1), the graphon U assumes the constant value q on B^2 . We conclude that $U = W$ a.e., proving Theorem 3.

7. Proof of Theorem 5

Recall that W is the 2-step graphon with the first step $A = [0, a]$ being an independent set, while W assumes value 1 on B^2 for $B := [a, 1]$ and value p on $A \times B$. We have to show that if a graphon U satisfies $\alpha(U) = a$, $\omega(U) = 1 - a$ and $X_4(U) = X_4(W) =: X_4$, then U is weakly isomorphic to W .

By Theorem 6 (and [12, Lemma 2.4]), there are subsets $C, D \subseteq [0, 1]$ of measures a and $1 - a$ respectively such that U is 0 on C^2 a.e. and U is 1 on D^2 a.e. The intersection $C \cap D$ must have measure 0, so we can assume that C and D partition $[0, 1]$. By applying a measure-preserving transformation to U , assume that $C = A$ and $D = B$.

CLAIM 7.1. *Almost every $x \in A$ satisfies $\deg_B^U(x) = (1 - a)p$.*

PROOF OF CLAIM. The graphon U has the same K_4 -density as W by $t(K_4, U) = \mathbb{P}(X_4 = 6)$. Since A is an independent set, there are only two types of K_4 in U : those inside B (and their contribution to the overall density is $(1 - a)^4$, the same as the analogous quantity for W) and those that have three vertices in B and one vertex in A . Since $U = 1$ on B^2 , the latter type of 4-cliques determines $\frac{1}{a} \int_A (\deg_B^U(x))^3 dx$, the third moment of the random variable $Y := \deg_B^U(x)$, where x is a uniform element of A . This third moment is the same as for W , which is $((1 - a)p)^3$ as $\deg_B^W(x)$ is the constant function $(1 - a)p$. Also, we have by Lemma 13 that

$$(17) \quad \mathbb{E}(Y) = \frac{1}{a} \int_A \deg_B(x) dx = \frac{\frac{1}{2} \mathbb{P}(X_2 = 1) - (1 - a)^2}{a} = (1 - a)p.$$

Thus $(\mathbb{E}(Y))^3 = \mathbb{E}(Y^3)$ which for a non-negative variable Y is possible only if Y is constant a.e. Of course, the constant value of Y must be $\mathbb{E}(Y) = (1 - a)p$, proving the claim. \square

Consider the random variable $Z := \deg_A^U(x)$ for uniform $x \in B$. A calculation analogous to that in (17) shows that $\mathbb{E}(Z) = ap$. Our knowledge about U directly gives the density of all possible copies of P_3 , except ABA -paths (that is, copies of P_3 that have the middle vertex in B and the other two in A). For example, the density of BAB -paths can be computed by sampling $x_2 \in A$ first and then using Claim 7.1. Since the total P_3 -density in U is the same as that for W by Lemma 13, we conclude that the density of ABA -paths in U is also the same as in W , which is $(1 - b)(ap)^2$. Thus $\mathbb{E}(Z^2) = (ap)^2 = (\mathbb{E}(Z))^2$. This implies that for a.e. $x \in B$ we have $\deg_A^U(x) = ap$. (Alternatively, this conclusion can be reached by applying the argument of Claim 7.1 to the complementary graphon $1 - U$.)

Let K_4^- be the 4-clique minus an edge, the unique graph on 4 vertices with 5 edges. Clearly, we have $t(K_4^-, U) = \frac{1}{6} \mathbb{P}(X_4 = 5) + \mathbb{P}(X_4 = 6) = t(K_4^-, W)$. The graphon U has 2 types of K_4^- of positive density. The first type consists of those copies of K_4^- that have exactly 3 vertices in B (and since U is 1 a.e. on B^2 , the corresponding density is determined by the degree distribution on A , which we know by Claim 7.1). Thus we know the density in U of the other copies of K_4^- which have 2 points in A and 2 points in B (and this matches that for W). Thus U has the same density of $ABAB$ -cycles as W . Also, their densities of AB -edges coincide, e.g. by Claim 7.1. Since W is constant- p on $A \times B$, the same must hold for U by Lemma 14. Thus U and W are weakly isomorphic.

8. Proofs of Propositions 8 and 9

PROOF OF PROPOSITION 8. We have to show that the family of graphs with at most one cycle is not forcing. Here we use the observation that if W is an n -step graphon with parts of measure $1/n$ and its values are encoded by a symmetric $n \times n$ matrix $A \in [0, 1]^{n \times n}$, then

$$(18) \quad t(C_k, W) = \frac{1}{n^k} \sum_{i=1}^n \lambda_i^k, \quad \text{for every } k \geq 3,$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A , repeated with their multiplicities. Indeed, $t(C_k, W)$ is the sum over all ordered k -tuples $(v_0, \dots, v_{k-1}) \in [n]^k$ of $\frac{1}{n^k} \prod_{i=0}^{k-1} A_{v_i, v_{i+1}}$, (where $v_k := v_0$). On the other hand, the j -th diagonal entry of A^k is the sum of $\prod_{i=0}^{k-1} A_{v_i, v_{i+1}}$ taken over all $(v_0, \dots, v_{k-1}) \in [n]^k$

with $v_0 = j$. Summing this over all $j \in [n]$, we get that $t(C_k, W)$ is the trace of $\frac{1}{n^k} A^k$, giving the identity in (18).

Take, for example, the following unit vectors:

$$\mathbf{x}_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{x}_3 := \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix},$$

and let e.g. $\varepsilon := 1/4$. Note that \mathbf{x}_2 and \mathbf{x}_3 are orthogonal to \mathbf{x}_1 . It routinely follows that the symmetric 3×3 matrices

$$(19) \quad A := \mathbf{x}_1 \mathbf{x}_1^T + \varepsilon \mathbf{x}_2 \mathbf{x}_2^T \quad \text{and} \quad A' := \mathbf{x}_1 \mathbf{x}_1^T + \varepsilon \mathbf{x}_3 \mathbf{x}_3^T$$

have the same eigenvalues (namely 1, ε and 0), all entries in $[0, 1]$ and all row sums equal (namely, to 1, which is the row sum of $\mathbf{x}_1 \mathbf{x}_1^T$).

Let W and W' be the 3-step graphons, with steps of measure $1/3$, whose values are given by the symmetric matrices $A, A' \in [0, 1]^{3 \times 3}$. The maximum entry of A' is $1/3 + 2\varepsilon/3$, which is strictly larger than the maximum entry $1/3 + \varepsilon/2$ of A , so the graphons W and W' are not weakly isomorphic by e.g. considering the density of K_k as $k \rightarrow \infty$ (and noting that the maximum entry of A' is on the diagonal).

On the other hand, W and W' have the same cycle densities by (18). Since they are both $(1/3)$ -regular, they have the same homomorphism density for every graph with at most one cycle by Lemma 11. This proves Proposition 8. \square

PROOF OF PROPOSITION 9. We have to show that the family of graphs of diameter at most d is not forcing. Let W (resp. W') be the graphon of the disjoint union $G := P_{d+2} \sqcup P_{d+2}$ (resp. $G' := P_{d+3} \sqcup P_{d+1}$). (Recall that P_n denotes the path with n vertices.) In other words, each of W and W' is the $\{0, 1\}$ -valued step graphon with $2d + 4$ steps of equal measure that encodes the adjacency relation of the corresponding graph. They are not weakly isomorphic because the induced density of P_{d+3} is zero in W but not in W' .

On the other hand, $t(F, W) = t(F, W')$ for every graph F of diameter at most d . Indeed, if $V(F) = [k]$, then for each choice of $\mathbf{x} = (x_1, \dots, x_k) \in [0, 1]^k$ for which the integrand in (1) is positive, the union of the corresponding vertices of G or G' must induce a subgraph of diameter at most d , that is, a sub-path with $i \leq d + 1$ vertices. The number of sub-paths of any given order $i \leq d + 1$ is the same for G and G' (namely, $2d + 4 - 2i$). Thus $t(F, W) = t(F, W')$.

We conclude that the family of graphs of diameter at most d is not forcing. \square

REMARK 16. One can view the construction in the proof of Proposition 9 as first taking the graphon W_{C_m} of the m -cycle C_m for $m := 2d + 4$ and then

decreasing density to zero on some two edges. Instead, we could have first multiplied the whole m -cycle graphon W_{C_m} by some real $p \in (0, 1)$ and then defined W and W' by modifying pW_{C_m} on the the same pairs of edges of C_m as before but in a way such that the new graphons are still $(2p/m)$ -regular but not weakly isomorphic. This modified construction shows by Lemma 11 that we can increase the family in Proposition 9 by taking all graphs of diameter at most d and then attaching any number of pendant trees.

9. Concluding remarks

The only information that our results on Question 2 used was the conclusions about the densities for graphs on at most 5 vertices that follow from the distribution of X_2, \dots, X_5 , and the probabilities that X_k is 0, 1, $\binom{k}{2} - 1$, or $\binom{k}{2}$ as $k \rightarrow \infty$. It would be interesting to find new types of arguments that use much more substantial information about the random variables X_k .

An intriguing open question (which is a considerable weakening of Question 2) is whether the sequence $(X_k(W))_{k \in \mathbb{N}}$ determines the essential supremum $\|W\|_\infty$ of an arbitrary graphon W . If true, this would greatly enlarge the set of 2-step graphons W for which we can prove that $W \in \mathcal{W}_5$.

References

- [1] D. J. Aldous, Representations for partially exchangeable arrays of random variables, *J. Multivariate Anal.*, **11** (1981), 581–598.
- [2] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, **28** (2015), 669–709.
- [3] C. Borgs, J. Chayes and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits, *Geom. Funct. Anal.*, **19** (2010), 1597–1619.
- [4] C. Borgs, J. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi, Counting graph homomorphisms, in: *Topics in Discrete Mathematics*, Algorithms Combin., vol 26, Springer (2006), pp. 315–371.
- [5] S. Chatterjee and S. R. S. Varadhan, The large deviation principle for the Erdős–Rényi random graph, *European J. Combin.*, **32** (2011), 1000–1017.
- [6] F. R. K. Chung, R. L. Graham and R. M. Wilson, Quasi-random graphs, *Combinatorica*, **9** (1989), 345–362.
- [7] E. Csóka, On the graph limit question of Vera T. Sós, *J. Combin. Theory (B)*, **116** (2016), 306–311.
- [8] P. Diaconis and S. Janson, Graph limits and exchangeable random graphs, *Rend. Mat.*, **28** (2008), 33–61.
- [9] E. DiBenedetto, *Real Analysis*, Birkhäuser Advanced Texts, Basler Lehrbücher, Birkhäuser/Springer (New York, 2016).
- [10] J. Fox, Personal communication.
- [11] I. Gohberg, S. Goldberg and M. A. Kaashoek, *Basic Classes of Linear Operators*, Birkhäuser Verlag (Basel, 2003).
- [12] J. Hladký, P. Hu and D. Piguet, Komlós’s tiling theorem via graphon covers, *J. Graph Theory*, **90** (2019), 24–45.

- [13] J. Hladký and I. Rocha, Independent sets, cliques and colorings in graphons, *European J. Combin.*, **88** (2020), 103108, 18pp.
- [14] D. Hoover, Relations on probability spaces and arrays of random variables, preprint, Institute for Advanced Study (Princeton, NJ, 1979).
- [15] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience New York, 2000).
- [16] D. J. Kleitman and K. J. Winston, The asymptotic number of lattices, *Ann. Discrete Math.*, **6** (1980), 243–249.
- [17] D. J. Kleitman and K. J. Winston, On the number of graphs without 4-cycles, *Discrete Math.*, **41** (1982), 167–172.
- [18] L. Lovász, *Large Networks and Graph Limits*, Colloquium Publications. Amer. Math. Soc. (2012).
- [19] L. Lovász and V. T. Sós, Generalized quasirandom graphs, *J. Combin. Theory (B)*, **98** (2008), 146–163.
- [20] L. Lovász and B. Szegedy, Limits of dense graph sequences, *J. Combin. Theory (B)*, **96** (2006), 933–957.
- [21] L. Lovász and B. Szegedy, Szemerédi’s lemma for the analyst, *Geom. Funct. Anal.*, **17** (2007), 252–270.
- [22] L. Lovász and B. Szegedy, Testing properties of graphs and functions, *Israel J. Math.*, **178** (2010), 113–156.
- [23] L. Lovász and B. Szegedy, Finitely forcible graphons, *J. Combin. Theory (B)*, **101** (2011), 269–301.
- [24] A. A. Sapozhenko, On the number of independent sets in extenders, *Diskret. Mat.*, **13** (2001), 56–62 (in Russian); translation in *Discrete Math. Appl.*, **11** (2001), 155–161.
- [25] A. A. Sapozhenko, Asymptotics of the number of sum-free sets in abelian groups of even order, *Dokl. Akad. Nauk*, **383** (2002), 454–457 (in Russian); translation in *Dokl. Math.*, **79** (2009), 79–80.
- [26] D. Saxton and A. Thomason, Hypergraph containers, *Invent. Math.*, **201** (2015), 925–992.
- [27] A. Shapira and M. Tyomkyn, Quasirandom graphs and the pantograph equation, arXiv: 2101.08173 (2021).
- [28] J. Sliacăn, Flag algebras and flagmatic, MPhil Thesis, University of Warwick (2015).
- [29] V. T. Sós, Some open problems on dense graphs, in: *Conference on Graphs and Analysis*, Institute for Advanced Study (Princeton, NJ, 2012).
- [30] T. Tao, *An Epsilon of Room. I: Real Analysis*, Pages from year three of a mathematical blog, Graduate Studies in Mathematics, vol. 117, Amer. Math. Soc. (Providence, RI, 2010).

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.