INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Cut-norm and entropy minimization over weak* limits

Martin Doležal
Jan Hladký

Preprint No. 34-2017

PRAHA 2017

CUT-NORM AND ENTROPY MINIMIZATION OVER WEAK* LIMITS

MARTIN DOLEŽAL AND JAN HLADKÝ

Abstract. We prove that the accumulation points of a sequence of graphs $G_1, G_2, G_3, \ldots$ with respect to the cut-distance are exactly the weak* limit points of subsequences of the adjacency matrices (when all possible orders of the vertices are considered) that minimize the entropy over all weak* limit points of the corresponding subsequence. In fact, the entropy can be replaced by any map $W \mapsto \iint f(W(x, y))$, where f is a continuous and strictly concave function. Our proofs are elementary, and do not use the regularity lemma.

As a corollary, we obtain a self-contained proof of compactness of the cut-distance topology. In particular, it avoids the regularity lemma machinery or ultraproduct techniques.

1. Introduction

The theory of limits of dense graphs was developed in [9, 4] and has revolutionized graph theory since then. The key objects of the theory are so-called graphons. More precisely, a graphon is a symmetric Lebesgue measurable function from I^2 to $[0, 1]$ where $I = [0, 1]$ is the unit interval (equipped by the Lebesgue measure λ). In the heart of the theory is then the following statement.

Theorem 1 (Informally)**.** *Suppose that $G_1, G_2, G_3, \ldots$ is a sequence of graphs. Then there exists a subsequence $G_{k_1}, G_{k_2}, G_{k_3}, \ldots$ and a graphon $W : I^2 \to [0, 1]$ such that $G_{k_1}, G_{k_2}, G_{k_3}, \ldots$ converges to W.*

Roughly speaking, to obtain the graphon W one looks at the adjacency matrices of the graphs $(G_{k_n})_n$ from distance. One possible way an analyst might attempt to make this statement formal could be to take W as a weak* limit[1] of adjacency matrices of the graphs $(G_{k_n})_n$ represented as functions from I^2 to $\{0, 1\}$. Such a version of Theorem 1 would be just an instance of the Banach–Alaoglu Theorem. However, the weak* topology turns out to be too coarse to provide the favorable properties that are available in the contemporary theory of graph limits.[2] A good toy example is the sequence of the complete balanced bipartite graphs $(K_{n,n})_{n=1}^{\infty}$. When considering adjacency matrices of these graphs with vertices grouped into the two parts of the bipartite graphs, the corresponding weak* limit is a 2×2-chessboard function with values 0 and 1, which we denote by $W_{\text{bipartite}}$. This turns out to be a desirable limit. On the other hand, one could consider adjacency matrices ordered differently. Ordering the vertices randomly, we get the constant $W_{\text{const}} \equiv \frac{1}{2}$ as the weak* limit (almost surely). We see that it is undesirable to get W_{const} as the limit object as the only information carried by such an object is that the overall edge densities of the graphs along the sequence converge to $\frac{1}{2}$.

So, instead of the weak* topology one considers the so-called cut-norm topology, and this is also the topology to which "converges to W" in Theorem 1 refers. The cut-norm $\|\cdot\|_\square$ is a certain uniformization of the weak* topology. Indeed, recall that given symmetric measurable functions

Jan Hladký was supported by the Alexander von Humboldt Foundation. Research of Martin Doležal was supported by the GAČR project GA16-07378S, and by RVO:67985840.

[1]See the Appendix for basic information about the weak* topology.

[2]A primal example of such a favorable property is the continuity of subgraph densities.

$\Gamma : I^2 \to [0,1]$ and $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$, the two convergence notions compare as follows.

$$\Gamma_n \xrightarrow{\text{w}^*} \Gamma \qquad \iff \qquad \sup_{B \subset I} \left\{ \limsup_{n} \left| \int_{x \in B} \int_{y \in B} \Gamma_n(x,y) - \Gamma(x,y) \right| \right\} = 0 \; ,$$

$$\Gamma_n \xrightarrow{\|\cdot\|_\square} \Gamma \qquad \iff \qquad \limsup_{n} \left\{ \sup_{B \subset I} \left| \int_{x \in B} \int_{y \in B} \Gamma_n(x,y) - \Gamma(x,y) \right| \right\} = 0 \; .$$

We shall state the formal version of Theorem 1 in a somewhat bigger generality for graphons. If $\Gamma, \Gamma' : I^2 \to [0,1]$ are two graphons then we say that they are *versions* of each other if they differ only by some measure-preserving transformation of I (see Section 2 for a precise definition).

Then the formal statement of Theorem 1 reads as follows.

Theorem 2. *Suppose that $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ is a sequence of graphons. Then there exists a sequence $k_1 < k_2 < k_3 < \cdots$ of natural numbers, versions $\Gamma'_{k_1}, \Gamma'_{k_2}, \Gamma'_{k_3}, \ldots$ of $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$, and a graphon $W : I^2 \to [0,1]$ such that the sequence $\Gamma'_{k_1}, \Gamma'_{k_2}, \Gamma'_{k_3}, \ldots$ converges to W in the cut-norm.*

Prior to our work, there were two approaches to proving Theorem 2. One, taken in [9] and in [10], uses (variants of) the regularity lemma to group parts of I according to the structure of Γ_n. This way, one approximates the graphons by step-functions, and the limit graphon W is a limit of these step-functions. The other approach, taken in [6], relies on ultraproduct techniques. This later approach is extremely technical, and was developed for the (more difficult) theory of limits of hypergraphs, where for some time the regularity approach was not available.

We present a third proof of Theorem 2. We believe that our approach is simpler, both on the conceptual level and on the technical level. Another advantage is that our approach seems to be more accessible to analysts. Last, but perhaps most importantly, our proof provides for the first time a characterization of the cut-norm convergence in terms of the weak* convergence. Namely, fixing any continuous and strictly concave function $f : [0,1] \to \mathbb{R}$, we prove that there is a subsequence $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$ such that the map $W \mapsto \iint f(W(x,y))$ attains its minimum on the space of all weak* accumulation points of versions of graphons $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$, and that any such minimizer is an accumulation point of a subsequence of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-distance. This result is consistent with our toy example above. Indeed, for any strictly concave function f we have $\iint f\left(W_{\text{bipartite}}(x,y)\right) < \iint f\left(W_{\text{const}}(x,y)\right)$ by Jensen's inequality. Jensen's inequality underlies the general proof of our result.

Note that every concave function defined on an open interval is continuous. Therefore the additional assumption of the continuity of the concave function $f : [0,1] \to \mathbb{R}$ which we work with in this paper only means that f is continuous (from the appropriate sides) at 0 and at 1.

Let $f : [0,1] \to \mathbb{R}$ be an arbitrary continuous and strictly concave function. Given a graphon $\Gamma : I^2 \to [0,1]$, we write $\mathrm{INT}_f(\Gamma) := \int_{x \in I} \int_{y \in I} f(\Gamma(x,y))$. When f is the binary entropy, the integration $\mathrm{INT}_f(W)$ appears also in the work on large deviations in random graphs, [5] (which does not relate to the current work otherwise), and is called the *entropy of the graphon W*.

For a sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ of graphons, we denote by $\mathbf{ACC}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ the set of all functions $W : I^2 \to [0,1]$ for which there exist versions $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ such that W is a weak* accumulation point of the sequence $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$. We also denote by $\mathbf{LIM}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ the set of all functions $W : I^2 \to [0,1]$ for which there exist versions $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ such that W is a weak* limit of the sequence $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$. We have $\mathbf{LIM}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) \subset \mathbf{ACC}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Note that $\mathbf{LIM}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ can be empty but $\mathbf{ACC}_{\text{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ cannot be empty by the sequential Banach–Alaoglu Theorem (see the Appendix for more details). Also, note that such weak* accumulation points (and thus also limits) are necessarily symmetric, Lebesgue measurable, $[0,1]$-valued, and thus graphons.

Our main result states that, given a sequence of graphons $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$, there is a subsequence $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$ such that the minimum of $\mathrm{INT}_f(\cdot)$ over the set $\mathbf{ACC}_{\text{w}^*}(\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots)$

is attained, and the graphon attaining this minimum is an accumulation point of the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-distance.

Theorem 3. *Suppose that* $f : [0,1] \to \mathbb{R}$ *is an arbitrary continuous and strictly concave function. Suppose that* $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ *is a sequence of graphons.*

(a) Suppose that $W \in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *and that* W *is not an accumulation point of the sequence* $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ *in the cut-norm. Then there exists* $\widetilde{W} \in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *such that* $\mathrm{INT}_f(\widetilde{W}) < \mathrm{INT}_f(W)$.

(b) There exist a subsequence $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$ *and a graphon* $W_{\min} \in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots)$ *such that*

$$\mathrm{INT}_f(W_{\min}) = \inf \left\{ \mathrm{INT}_f(W) \colon W \in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots) \right\} .$$

Clearly, Theorem 3 implies Theorem 2.

The proof of Theorem 3 is given in Sections 4 and 5.

To complete the "characterization of the cut-norm convergence in terms of the weak* convergence" advertised above, we prove that weak* limit points that do not minimize $\mathrm{INT}_f(\cdot)$ cannot be limit points in the cut-norm.

Proposition 4. *Suppose that* $f : [0,1] \to \mathbb{R}$ *is an arbitrary continuous and strictly concave function. Suppose that* $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ *is a sequence of graphons. If* $W \in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *is a cut-norm limit of versions of* $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ *then* W *is a minimizer of* $\mathrm{INT}_f(\cdot)$ *over the space* $\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.

In Section 4 we show that Proposition 4 is an easy consequence of a result of Borgs, Chayes, and Lovász [3] on uniqueness of graph limits. In addition, we give a self-contained proof.

2. Notation and tools

For every function $W : I^2 \to \mathbb{R}$, we define the *cut-norm* of W by

$$\|W\|_\square = \sup_A \left| \int_A \int_A W(x,y) \right| , \tag{1}$$

where A ranges over all measurable subsets of I. Another slightly different formula is also often used in the literature where one replaces the right-hand side of (1) by $\sup_{A,B} \left| \int_A \int_B W(x,y) \right|$ where two sets A and B range over all measurable subsets of I. However, it is easy to see that for every symmetric function W, we have

$$\sup_A \left| \int_A \int_A W(x,y) \right| \le \frac{3}{2} \sup_{A,B} \left| \int_A \int_B W(x,y) \right| ,$$

and so the notion of convergence of sequences of graphons (which are symmetric) in the cut-norm is irrelevant to the choice between these two formulas.

We say that a graphon $\Gamma \colon I^2 \to [0,1]$ is a *step-graphon* with steps $I_1, I_2, \ldots, I_k \subset I$ if the sets $I_1, I_2, \ldots, I_k$ are pairwise disjoint, $I_1 \cup I_2 \cup \ldots \cup I_k = I$ and $W_{|I_i \times I_j}$ is constant (up to a null set) for every $i, j = 1, 2, \ldots, k$.

We say that a measurable function $\gamma : I \to I$ is an *almost-bijection* if there exist conull sets $J_1, J_2 \subset I$ such that $\gamma_{|J_1}$ is a bijection from J_1 onto J_2. When we talk about the inverse of such a function γ then we mean $(\gamma_{|J_1})^{-1}$ but we denote it only by γ^{-1}. Note that this inverse γ^{-1} is not unique but that does not cause any problems as any two inverses of γ differ only on a null set.

If $\Gamma, \Gamma' : I^2 \to [0,1]$ are two graphons then we say that Γ' is a *version* of Γ if there exists a measure preserving almost-bijection $\gamma : I \to I$ such that $\Gamma'(x,y) = \Gamma(\gamma^{-1}(x), \gamma^{-1}(y))$ for almost every $(x,y) \in I^2$.

Related to versions, we recall that the *cut-distance* and L^1*-distance* between two graphons W_1, W_2 are defined as $\delta_\square(W_1, W_2) = \inf \|U_1 - W_2\|_\square$ and $\delta_1(W_1, W_2) = \inf \|U_1 - W_2\|_1$ where U_1 ranges over all versions of W_1.

By an *ordered partition* of I, we mean a partition of I with a fixed order of the sets from the partition. For an ordered partition $\mathcal{J}$ of I into finitely many sets $C_1, C_2, \ldots, C_k$, we define mappings $\alpha_{\mathcal{J},1}, \alpha_{\mathcal{J},2}, \ldots, \alpha_{\mathcal{J},3} : I \to I$, and a mapping $\gamma_{\mathcal{J}} : I \to I$ by

$$
\begin{aligned}
\alpha_{\mathcal{J},1}(x) &= \int_0^x \mathbf{1}_{C_1}(y)\,\mathrm{d}(y)\ , \\
\alpha_{\mathcal{J},2}(x) &= \alpha_{\mathcal{J},1}(1) + \int_0^x \mathbf{1}_{C_2}(y)\,\mathrm{d}(y)\ , \\
&\ \vdots \\
\alpha_{\mathcal{J},k}(x) &= \alpha_{\mathcal{J},1}(1) + \alpha_{\mathcal{J},2}(1) + \ldots + \alpha_{\mathcal{J},k-1}(1) + \int_0^x \mathbf{1}_{C_k}(y)\,\mathrm{d}(y)\ , \\
\gamma_{\mathcal{J}}(x) &= \alpha_{\mathcal{J},i}(x) \quad \text{if } x \in C_i, \quad i = 1, 2, \ldots, k\ .
\end{aligned}
\tag{2}
$$

Informally, $\gamma_{\mathcal{J}}$ is defined in such a way that it maps the set C_1 to the left side of the interval I, the set C_2 next to it, and so on. Finally, the set C_k is mapped to the right side of the interval I. Clearly, $\gamma_{\mathcal{J}}$ is a measure preserving almost-bijection.

For a graphon $W : I^2 \to [0, 1]$ and an ordered partition $\mathcal{J}$ of I into finitely many sets, we denote by ${}_{\mathcal{J}}W$ the version of W defined by ${}_{\mathcal{J}}W(x, y) = W(\gamma_{\mathcal{J}}^{-1}(x), \gamma_{\mathcal{J}}^{-1}(y))$ for every $(x, y) \in I^2$.

2.1. **Lebesgue points.** Recall that whenever $W : I^2 \to \mathbb{R}$ is an integrable function then almost every point $(x, y) \in I^2$ is a *Lebesgue point* of W. This means that for every $\eta > 0$ there exists $\delta_0 > 0$ such that for every $\delta \in (0, \delta_0]$ we have

$$
\frac{1}{4\delta^2} \int_{x-\delta}^{x+\delta} \int_{y-\delta}^{y+\delta} |W(x, y) - W(w, z)\,\mathrm{d}(w)\,\mathrm{d}(z)| < \eta\ . \tag{3}
$$

Note that the integration in (3) is over the square $[x - \delta, x + \delta] \times [y - \delta, y + \delta]$ (with (x, y) in the center), and it is not possible in general to extend this formula to integration over arbitrary rectangles containing the point (x, y). However, easy (and well known) calculations show that one can extend this formula to integration over all rectangles containing the point (x, y) such that the ratio of the lengths of their sides lies in some interval with positive endpoints given in advance (e.g. when no side of the rectangle is longer than double the length of the other side). Therefore for a.e. $(x, y) \in I^2$ and for every $\eta > 0$ there exists $\delta_0 > 0$ such that whenever $[p_1, p_2] \subset I$ and $[q_1, q_2] \subset I$ are intervals such that the length of the intervals is smaller or equal to δ_0, such that the ratio of the lengths of these intervals is at least $\frac{1}{2}$ and at most 2, and such that $[p_1, p_2]$ contains x and $[q_1, q_2]$ contains y then

$$
\frac{1}{(p_2 - p_1)(q_2 - q_1)} \int_{p_1}^{p_2} \int_{q_1}^{q_2} |W(x, y) - W(w, z)\,\mathrm{d}(w)\,\mathrm{d}(z)| < \eta\ ,
$$

which clearly implies that

$$
\left| W(x, y) - \frac{1}{(p_2 - p_1)(q_2 - q_1)} \int_{p_1}^{p_2} \int_{q_1}^{q_2} W(w, z)\,\mathrm{d}(w)\,\mathrm{d}(z) \right| < \eta\ . \tag{4}
$$

2.2. **Averaged graphons.** The next definition introduces graphons derived by averaging of a given graphon W on a given partition of I. Here, we denote by $\lambda^{\oplus 2}$ the two-dimensional Lebesgue measure on I^2.

Definition 5. *Suppose that* $\Gamma : I^2 \to [0,1]$ *is a graphon. For a partition* $\mathcal{I}$ *of the unit interval into finitely many sets of positive measure,* $I = I_1 \sqcup I_2 \sqcup \ldots \sqcup I_k$*, we define a graphon* $\Gamma^{\Join\mathcal{I}}$ *which is defined on each rectangle* $I_i \times I_j$ *to be the constant* $\frac{1}{\lambda^{\oplus 2}(I_i \times I_j)} \int_{I_i} \int_{I_j} W(x,y)$.

The next lemma shows that we can replace any graphon W by its averaged graphon (on some partition of I) without changing the value of $\mathrm{INT}_f(W)$ too much.

Lemma 6. *Let* $f : [0,1] \to \mathbb{R}$ *be an arbitrary continuous and strictly concave function, and let* $\mathcal{J}$ *be an arbitrary partition of* I *into finitely many intervals of positive measure. Suppose that* $W \colon I^2 \to [0,1]$ *is a graphon, and let* $\varepsilon > 0$*. Then there exists a partition* $\mathcal{I}$ *of* I *into finitely many intervals of positive measure such that* $\mathcal{I}$ *is a refinement of* $\mathcal{J}$ *and such that* $|\mathrm{INT}_f(W) - \mathrm{INT}_f(W^{\Join\mathcal{I}})| < \varepsilon$.

Proof. As f is continuous, there is $\eta > 0$ such that $|f(x) - f(y)| < \frac{1}{2}\varepsilon$ whenever $x, y \in [0,1]$ are such that $|x - y| < \eta$. Also, as W is an integrable function, almost every point $(x,y) \in I^2$ is a Lebesgue point of W. This implies that for a.e. $(x,y) \in I^2$ there is a natural number $n(x,y)$ such that whenever $[p_1,p_2] \subset I$ and $[q_1,q_2] \subset I$ are intervals of lengths smaller or equal to $\frac{2}{n(x,y)}$ such that the ratio of the lengths is at least $\frac{1}{2}$ and at most 2, and such that $[p_1,p_2]$ contains x and $[q_1,q_2]$ contains y then the inequality (4) holds. Let us denote $C := \max_{x\in[0,1]} |f(x)|$. We find a natural number n_0 large enough such that

$$\lambda^{\oplus 2}\left(\left\{(x,y) \in I^2 \colon n(x,y) > n_0\right\}\right) < \frac{1}{4C}\varepsilon \;, \tag{5}$$

and such that $\frac{1}{n_0}$ is smaller than the length of all intervals from the partition $\mathcal{J}$. Let us denote $B := \left\{(x,y) \in I^2 \colon n(x,y) > n_0\right\}$. Now let $\mathcal{I}$ be an arbitrary refinement of the partition $\mathcal{J}$ into finitely many intervals $I_1, I_2, \ldots, I_k$, such that the length of each of these intervals is at least $\frac{1}{n_0}$ and at most $\frac{2}{n_0}$. For each $i,j = 1,2,\ldots,k$, we denote $C_{i,j} = \frac{1}{\lambda^{\oplus 2}(I_i \times I_j)} \int_{I_i} \int_{I_j} W(x,y)$. Inequality (4) then tells us that

$$|W(x,y) - C_{i,j}| < \eta \qquad \text{for every } (x,y) \in (I_i \times I_j) \setminus B, \quad i,j = 1,2,\ldots,k \;,$$

and so

$$|f(W(x,y)) - f(C_{i,j})| < \tfrac{1}{2}\varepsilon \qquad \text{for every } (x,y) \in (I_i \times I_j) \setminus B, \quad i,j = 1,2,\ldots,k \;. \tag{6}$$

So we have

$$\begin{aligned}
&\left|\mathrm{INT}_f(W) - \mathrm{INT}_f(W^{\Join\mathcal{I}})\right| \\
\leq &\iint\limits_B \left|f(W(x,y)) - f(W^{\Join\mathcal{I}}(x,y))\right| + \sum_{i,j=1}^{k} \iint\limits_{(I_i \times I_j)\setminus B} \left|f(W(x,y)) - f(W^{\Join\mathcal{I}}(x,y))\right| \\
\overset{(6)}{\leq} &2C \cdot \lambda^{\oplus 2}(B) + \frac{1}{2}\varepsilon \sum_{i,j=1}^{k} \lambda^{\oplus 2}((I_j \times I_j) \setminus B) \\
\overset{(5)}{<} &\frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \;,
\end{aligned}$$

as we wanted. □

The next lemma says that if a graphon is a weak* limit point then so is any graphon derived by averaging of the original one on a given partition of I into intervals.

Lemma 7. *Suppose that* $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ *is a sequence of graphons. Suppose that* $W \in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *and that we have a partition* $\mathcal{I}$ *of* I *into finitely many intervals of positive measure. Then* $W^{\Join\mathcal{I}} \in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.

Moreover, whenever $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ are versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ which converge to W in the weak topology then the versions $\Gamma''_1, \Gamma''_2, \Gamma''_3, \ldots$ of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ weak* converging to $W^{\bowtie \mathcal{I}}$ can be chosen in such a way that for every natural number j and for every intervals $K, L \in \mathcal{I}$ it holds*

$$\int_K \int_L \Gamma'_j(x,y) = \int_K \int_L \Gamma''_j(x,y) \ .$$

The proof of Lemma 7 follows a relatively standard probabilistic argument. Suppose for simplicity that $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ weak* converges to W. Then, for each n, we consider a version Γ'_n of Γ_n which is obtained by splitting each interval $A \in \mathcal{I}$ into n subsets of the same measure and then permuting these subsets of A at random. It can then be shown that $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ converge to $W^{\bowtie \mathcal{I}}$ almost surely. The next two definitions are needed to make precise the notion of randomly permuting parts of the graphon within a given partition.

Definition 8. *Given a set $A \subset I$ of positive measure and a number $s \in \mathbb{N}$, we can consider a partition $A = [\![A]\!]^s_1 \sqcup [\![A]\!]^s_2 \sqcup \ldots \sqcup [\![A]\!]^s_s$, where each set $[\![A]\!]^s_i$ has measure $\frac{\lambda(A)}{s}$ and for each $1 \le i < j \le s$, the set $[\![A]\!]^s_i$ is entirely to the left of $[\![A]\!]^s_j$. These conditions define the partition $A = [\![A]\!]^s_1 \sqcup [\![A]\!]^s_2 \sqcup \ldots \sqcup [\![A]\!]^s_s$ uniquely, up to null sets. For each $i, j \in [s]$ there is a natural, uniquely defined (up to null sets), measure preserving almost-bijection $\chi^{A,s}_{i,j} : [\![A]\!]^s_i \to [\![A]\!]^s_j$ which preserves the order on the real line.*

Definition 9. *Suppose that $\Gamma : I^2 \to [0,1]$ is a graphon. For a partition $\mathcal{I}$ of I into finitely many sets of positive measure, $I = I_1 \sqcup I_2 \sqcup \ldots \sqcup I_k$, and for $s \in \mathbb{N}$, we define a discrete distribution $\mathbb{W}(\Gamma, \mathcal{I}, s)$ on graphons using the following procedure. We take $\pi_1, \ldots, \pi_k : [s] \to [s]$ independent uniformly random permutations. After these are fixed, we define a sample $W \sim \mathbb{W}(\Gamma, \mathcal{I}, s)$ by*

$$W(x,y) = W\left(\chi^{I_i,s}_{p,\pi_i(p)}(x), \chi^{I_j,s}_{q,\pi_j(q)}(y)\right) \quad \text{when } x \in [\![I_i]\!]^s_p,\ y \in [\![I_j]\!]^s_q,\ i,j \in [k],\ p,q \in [s] \ .$$

This defines the sample $W : I^2 \to [0,1]$ uniquely up to null sets, and thus defines the whole distribution $\mathbb{W}(\Gamma, \mathcal{I}, s)$. Observe that $\mathbb{W}(\Gamma, \mathcal{I}, s)$ is supported on (some) versions of Γ.

We call the sets $[\![I_j]\!]^s_q$ stripes.

Proof of Lemma 7. By considering suitable versions of the graphons Γ_n, we can without loss of generality assume that the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ itself converges to W in the weak* topology. For each $n \in \mathbb{N}$, let us sample $U_n \sim \mathbb{W}(\Gamma_n, \mathcal{I}, n)$. We claim that the sequence $U_1, U_2, U_3, \ldots$ converges to $W^{\bowtie \mathcal{I}}$ in the weak* topology almost surely. As each U_n is a version of Γ_n, this will prove the lemma. So, let us now turn to proving the claim.

Let $i, j \in [k]$ be arbitrary. Further, let $0 \le p_1 < p_2 \le 1$ and $0 \le r_1 < r_2 \le 1$ be arbitrary rational numbers such that the rectangle $[p_1, p_2] \times [r_1, r_2]$ is contained (modulo a null set) in $I_i \times I_j$. Having fixed i, j, p_1, p_2, r_1, r_2, let us write c for the value of $W^{\bowtie \mathcal{I}}$ on $I_i \times I_j$. For each $n \in \mathbb{N}$, let E_n be the event that

$$\left| \iint_{[p_1,p_2]\times[r_1,r_2]} U_n \mathrm{d}(\lambda^{\oplus 2}) - c(p_2 - p_1)(r_2 - r_1) \right| > \sqrt[4]{1/n} + \tfrac{4}{n} \ .$$

Let us now bound the probability that E_n occurs. To this end, let Y_n be the value of $\iint_{[p_1,p_2]\times[r_1,r_2]} U_n \mathrm{d}(\lambda^{\oplus 2})$. We clearly have $\mathbb{E}[Y_n] = c(p_2 - p_1)(r_2 - r_1) \pm \frac{4}{n}$ (the error $\pm \frac{4}{n}$ comes from those products of pairs of stripes that intersect both $[p_1, p_2] \times [r_1, r_2]$ and its complement). Therefore, if E_n occurs then $|Y_n - \mathbb{E}[Y_n]| > \sqrt[4]{1/n}$. Suppose that we want to compute Y_n. From the k random permutations $\pi_1, \pi_2, \ldots, \pi_k : [n] \to [n]$ used in Definition 9 to define U_n, we only need to know the permutations π_i and π_j. To generate these, we toss in i.i.d. points $i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n$ into the unit interval I; the Euclidean order of the points $i_1, i_2, \ldots, i_n$ naturally defines π_i and similarly the points

$j_1, j_2, \ldots, j_n$ naturally define π_j.[3] So, we can view Y_n as a random variable on the probability space I^{2n}. Observe that if $\mathfrak{s} = (i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n)$ and $\mathfrak{s}' = (i'_1, i'_2, \ldots, i'_n, j'_1, j'_2, \ldots, j'_n)$ are two elements of I^{2n} that differ in only one coordinate, then $|Y_n(\mathfrak{s}) - Y_n(\mathfrak{s}')| \leq \frac{2}{n}$. Thus the Method of Bounded Differences (see [12]) tells us that

$$\mathbb{P}\left[E_n\right] \leq \mathbb{P}\left[|Y_n - \mathbb{E}[Y_n]| > \sqrt[4]{1/n}\right] \leq 2\exp\left(-\frac{2(\sqrt[4]{1/n})^2}{2n \cdot \left(\frac{2}{n}\right)^2}\right) = 2\exp\left(-\sqrt{n}/4\right) .$$

Because the sequence $(2\exp(-\sqrt{n}/4))_{n=1}^{\infty}$ is summable, the Borel–Cantelli lemma allows to conclude that only finitely many events E_n occur, almost surely. Thus, almost surely, for any weak* accumulation point U of the sequence $U_1, U_2, U_3, \ldots$, we have

$$\iint_{[p_1,p_2]\times[r_1,r_2]} U \mathrm{d}(\lambda^{\oplus 2}) = c(p_2 - p_1)(r_2 - r_1) . \tag{7}$$

By applying the union bound, we obtain that (7) holds for all (countably many) choices of i, j, p_1, p_2, r_1, r_2, almost surely. Since the elements of $\mathcal{I}$ are intervals, the above system of rectangles $[p_1, p_2] \times [r_1, r_2]$ generates the Borel σ-algebra on I^2. Consequently, we obtain that $U \equiv W^{\ltimes \mathcal{I}}$, almost surely.

The "moreover" part obviously follows from the proof. $\square$

2.3. **Jensen's inequality and averaged graphons.** Recall that one of the possible formulations of Jensen's inequality says that if (Ω, λ) is a measurable space with $\lambda(\Omega) > 0$, $g : \Omega \to \mathbb{R}$ is a measurable function and $f : \mathbb{R} \to \mathbb{R}$ is a concave function then

$$f\left(\frac{1}{\lambda(\Omega)} \int_\Omega g(x)\right) \geq \frac{1}{\lambda(\Omega)} \int_\Omega f(g(x)) . \tag{8}$$

We use this formulation of Jensen's inequality to prove the following simple lemma.

Lemma 10. *Let $f\colon [0,1] \to \mathbb{R}$ be a continuous and strictly concave function. Let $\Gamma\colon I^2 \to [0,1]$ be a step-graphon with steps $I_1, I_2, \ldots, I_k$, and let $W\colon I^2 \to [0,1]$ be another graphon such that $\int_{I_i \times I_j} W = \int_{I_i \times I_j} \Gamma$ for every $i, j = 1, 2, \ldots, k$. Then $\mathrm{INT}_f(W) \leq \mathrm{INT}_f(\Gamma)$.*

Proof. It clearly suffices to show that for every $i, j = 1, 2, \ldots, k$ it holds

$$\int_{I_i} \int_{I_j} f(W(x,y)) \leq \int_{I_i} \int_{I_j} f(\Gamma(x,y)) .$$

So let us fix i, j, and let $C_{i,j}$ be the constant for which $\Gamma_{|I_i \times I_j} = C_{i,j}$ almost everywhere. Then we have

$$\begin{aligned}
\int_{I_i} \int_{I_j} f(W(x,y)) &\overset{(8)}{\leq} \lambda^{\oplus 2}(I_i \times I_j) \cdot f\left(\frac{1}{\lambda^{\oplus 2}(I_i \times I_j)} \int_{I_i} \int_{I_j} W(x,y)\right) \\
&= \lambda^{\oplus 2}(I_i \times I_j) \cdot f\left(C_{i,j}\right) \\
&= \int_{I_i} \int_{I_j} f(\Gamma(x,y)) ,
\end{aligned}$$

as we wanted. $\square$

[3] The exception being when some of the points $i_1, i_2, \ldots, i_n$ or of the points $j_1, j_2, \ldots, j_n$ coincide, in which case the order of these points does not determine a permutation. This event however happens almost never.

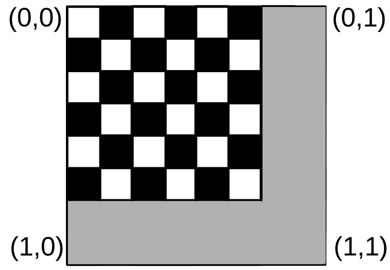

FIGURE 1. The graphon U_3 from Section 3.2. Value 0 is white, value $\frac{1}{2}$ is gray, value 1 is black.

3. Summaries of proofs

3.1. **Overview of proof of Theorem 3**(a)**.** Suppose for simplicity that the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ converges to W in the weak* topology. The key step to the proof of Theorem 3(a) is Lemma 11. There we prove that whenever we fix a sequence $(B_n)_{n=1}^{\infty}$ of measurable subsets of I and define a new version Γ'_n of Γ_n (for every n) by "shifting the set B_n to the left side of the interval I", then any weak* accumulation point $\widetilde{W}$ of the sequence $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ satisfies $\mathrm{INT}_f(\widetilde{W}) \leq \mathrm{INT}_f(W)$. As this result relies on Jensen's inequality, we actually get $\mathrm{INT}_f(\widetilde{W}) < \mathrm{INT}_f(W)$ when we choose the sets B_n carefully. "Carefully" means that each the integrals $\int_{B_n}\int_{B_n}\Gamma_n(x,y)$ differs from the integral $\int_{B_n}\int_{B_n} W(x,y)$ at least by some given $\varepsilon > 0$. But it is always possible to choose the sets B_n in this way as the graphon W is not a cut-norm accumulation point of the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$.

3.2. **Overview of proof of Theorem 3**(b)**.** Let us begin with the most straightforward attempt for a proof. For now, let us work with the simplifying assumption that all accumulation points are actually limits. As we shall see later, this simplifying assumption is a major cheat for which an extra patch will be needed. Then, let

$$m := \inf\left\{\mathrm{INT}_f(W)\colon W \in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)\right\}\ .$$

For each $k \in \mathbb{N}$, let us fix a sequence $\Gamma_1^k, \Gamma_2^k, \Gamma_3^k, \ldots$ of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ which converges in the weak* topology to a graphon $\widetilde{W_k}$ with $\mathrm{INT}_f(\widetilde{W_k}) < m + \frac{1}{k}$. Now, we might diagonalize and hope that any weak* accumulation point (whose existence is guaranteed by the Banach–Alaoglu Theorem) W^* of the sequence $\Gamma_1^1, \Gamma_2^2, \Gamma_3^3, \ldots$ satisfies $\mathrm{INT}_f(W^*) \leq m$. The reason for this hope being vain is the discontinuity of $\mathrm{INT}_f(\cdot)$ with respect to the weak* topology. As an example, let us take a situation when each $\widetilde{W_k}$ is a $2(k+2) \times 2(k+2)$-chessboard $\{0,1\}$-valued function, with the last two rows and columns having value $\frac{1}{2}$ (see Figure 1). In other words, most of each graphon $\widetilde{W_k}$ corresponds to a complete balanced bipartite graphon, to which an additional artificial subdivision to each of its parts to k subparts was introduced. These subparts were interlaced one after another, except that the vertices of the last subpart of each part were mixed together. (These graphons were clearly chosen nonoptimally in the sense that the mixing of the last two parts is undesired. We chose these graphons in this example here to have richer features to study.) All the graphons $\widetilde{W_k}$ have small values of $\mathrm{INT}_f(\cdot)$. On the other hand, the weak* limit of the sequence is the graphon $W_{\mathrm{const}} \equiv \frac{1}{2}$ whose value $\mathrm{INT}_f(\cdot)$ is bigger. There is a lesson to learn from this example. While for larger k, the versions in the sequence $\Gamma_1^k, \Gamma_2^k, \Gamma_3^k, \ldots$ will be aligned on I in a more optimal way locally, the global structure may get undesirably more convoluted as $k \to \infty$. To remedy this, we consider a sequence of version of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in which the structure of measure-preserving transformation on a rough level is inherited from measure preserving transformations leading to $\widetilde{W_1}$. Within each step corresponding to the step-graphon $\widetilde{W_1}$, the structure of the measure-preserving transformation

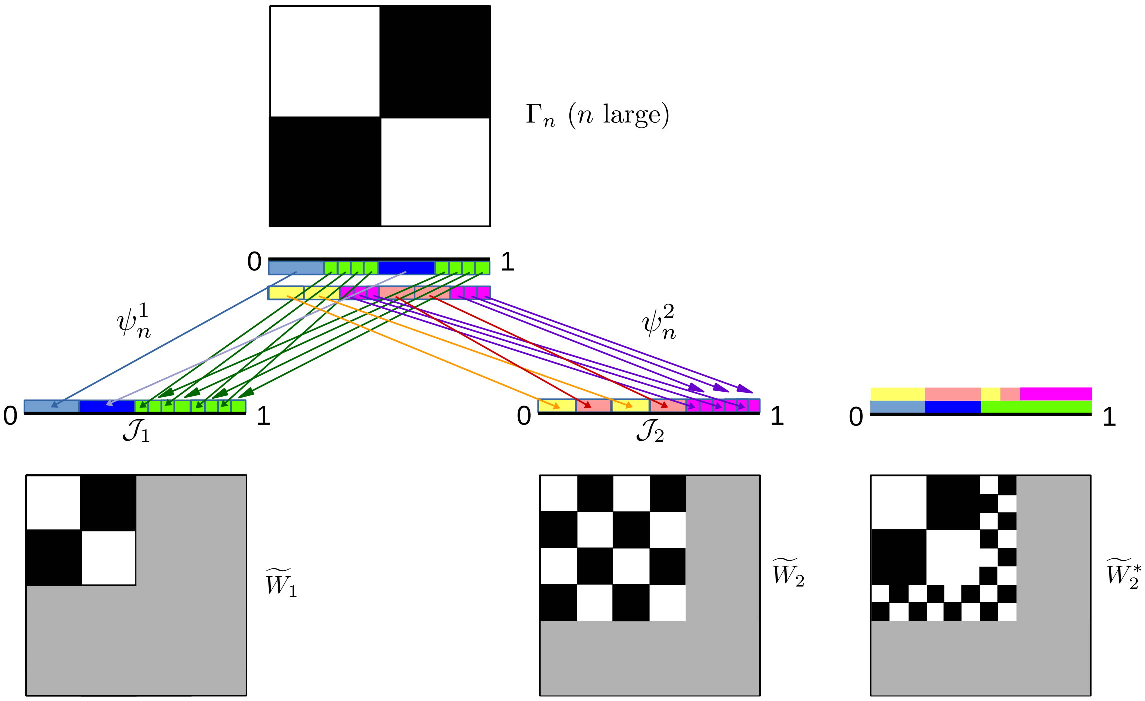

FIGURE 2. An example of reordering from Section 3.2. The top shows a graphon Γ_n, versions of which are close to $\widetilde{W}_1$ and $\widetilde{W}_2$ in the weak* topology. The two measure preserving transformations ψ_n^1 and ψ_n^2 which witness this closeness are shown with colors. The graphon $\widetilde{W}_2^*$ emerges by taking the partition whose global structure from $\widetilde{W}_1$ is refined according to the more local structure from $\widetilde{W}_2$. Iterating this process would lead to a sequence of graphons $(\widetilde{W}_k^*)_k$ which has the property that for any weak* accumulation point W^* we have $\mathrm{INT}_f(W^*) \leq \limsup_n \mathrm{INT}_f \widetilde{W}_n$.

is inherited from measure preserving transformations leading to $\widetilde{W}_2$, and so on. An example of this procedure is given in Figure 2. It can be shown that any weak* accumulation point W^* of these reordered graphons has the property that $\mathrm{INT}_f(W^*) \leq \limsup_n \mathrm{INT}_f \widetilde{W}_n$, as was needed.

Let us now explain why the assumption that all sequences converge weak* leaves a substantial gap in the proof. Recall that the information how the partition $\mathcal{J}^k$ of U_k interacts with the measure preserving almost-bijections on graphons $\mathfrak{s}_k \subset (\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ that converge to $\widetilde{W}_k$ gives us crucial directions as how to reorder and refine the subsequence of graphons $\mathfrak{s}_{k+1}$ that converges to $\widetilde{W}_{k+1}$. Let us again stress that while the existence of the subsequences $\mathfrak{s}_j$ is guaranteed by weak* compactness, we have no control on their properties. So, it can be that $\mathfrak{s}_k$ is disjoint from $\mathfrak{s}_{k+1}$. In other words, we do not get the needed information how to reorder and refine the graphons in $\mathfrak{s}_{k+1}$. To remedy this problem, we prove a lemma (Lemma 13) which says that for every sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ of graphons there exists a subsequence $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$ such that

$$\inf \left\{\mathrm{INT}_f(W)\colon W \in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots)\right\} = \inf \left\{\mathrm{INT}_f(W)\colon W \in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots)\right\} .$$

Applying this lemma first, the arguments above become sound for the subsequence $\Gamma_{k_1}, \Gamma_{k_2}, \Gamma_{k_3}, \ldots$.

4. Proof of Theorem 3(a)

The following key lemma (or its subsequent corollary) is used in both proofs of Theorem 3(a) and Theorem 3(b).

Lemma 11. *Suppose that $f : [0,1] \to \mathbb{R}$ is an arbitrary continuous and strictly concave function. Suppose that $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : I^2 \to [0,1]$ is a sequence of graphons which converges to a graphon $W : I^2 \to [0,1]$ in the weak* topology. Suppose that $B_1, B_2, B_3, \ldots$ is an arbitrary sequence of subsets of I. For each n, let $\mathcal{J}_n$ be the ordered partition of I into two sets B_n and $I \setminus B_n$ (in this order). Then every $\widetilde{W} \in \mathbf{ACC}_{\mathrm{w}^*}({}_{\mathcal{J}_1}\Gamma_1, {}_{\mathcal{J}_2}\Gamma_2, {}_{\mathcal{J}_3}\Gamma_3, \ldots)$ satisfies $\mathrm{INT}_f(\widetilde{W}) \leq \mathrm{INT}_f(W)$.*

Moreover, let $\theta \colon I \to I$ be defined by $\theta(x) = \int_0^x \psi(y)\,\mathrm{d}(y)$ where $\psi \colon I \to [0,1]$ is some accumulation point of the sequence $\mathbf{1}_{B_1}, \mathbf{1}_{B_2}, \mathbf{1}_{B_3}, \ldots$ in the weak topology. If the equality $W(x,y) = \widetilde{W}(\theta(x), \theta(y))$ does not hold on a set of full measure then $\mathrm{INT}_f(\widetilde{W}) < \mathrm{INT}_f(W)$.*

Proof. Let us fix $\widetilde{W} \in \mathbf{ACC}_{\mathrm{w}^*}({}_{\mathcal{J}_1}\Gamma_1, {}_{\mathcal{J}_2}\Gamma_2, {}_{\mathcal{J}_3}\Gamma_3, \ldots)$. By passing to a subsequence, we may assume that the sequence ${}_{\mathcal{J}_1}\Gamma_1, {}_{\mathcal{J}_2}\Gamma_2, {}_{\mathcal{J}_3}\Gamma_3, \ldots$ is convergent to $\widetilde{W}$ in the weak* topology, and that the sequence $\mathbf{1}_{B_1}, \mathbf{1}_{B_2}, \mathbf{1}_{B_3}, \ldots$ converges in the weak* topology to $\psi \colon I \to [0,1]$ (which can be chosen to be any accumulation point of the original sequence). We define $\xi \colon I \to I$ by $\xi(x) = \theta(1) + \int_0^x (1 - \psi(y))\,\mathrm{d}(y)$.

Claim 1. *For every two intervals $[p_1, p_2], [q_1, q_2] \subset I$ we have*

$$
\begin{aligned}
\int_{p_1}^{p_2} \int_{q_1}^{q_2} W(x,y) = & \int_{p_1}^{p_2} \int_{q_1}^{q_2} \widetilde{W}(\theta(x), \theta(y)) \psi(x) \psi(y) \\
& + \int_{p_1}^{p_2} \int_{q_1}^{q_2} \widetilde{W}(\theta(x), \xi(y)) \psi(x) (1 - \psi(y)) \\
& + \int_{p_1}^{p_2} \int_{q_1}^{q_2} \widetilde{W}(\xi(x), \theta(y)) (1 - \psi(x)) \psi(y) \\
& + \int_{p_1}^{p_2} \int_{q_1}^{q_2} \widetilde{W}(\xi(x), \xi(y)) (1 - \psi(x)) (1 - \psi(y)) \, .
\end{aligned} \tag{9}
$$

Proof of Claim 1. By using the fact that $\Gamma_n \xrightarrow{w^*} W$ together with the identity $ab + a(1-b) + (1-a)b + (1-a)(1-b) = 1$ we get that

$$
\begin{aligned}
\int_{p_1}^{p_2} \int_{q_1}^{q_2} W(x,y) &= \lim_{n\to\infty} \int_{p_1}^{p_2} \int_{q_1}^{q_2} \Gamma_n(x,y) \\
&= \lim_{n\to\infty} \int_{p_1}^{p_2} \int_{q_1}^{q_2} \Gamma_n(x,y) \mathbf{1}_{B_n}(x) \mathbf{1}_{B_n}(y) \\
&\quad + \lim_{n\to\infty} \int_{p_1}^{p_2} \int_{q_1}^{q_2} \Gamma_n(x,y) \mathbf{1}_{B_n}(x) (1 - \mathbf{1}_{B_n}(y)) \\
&\quad + \lim_{n\to\infty} \int_{p_1}^{p_2} \int_{q_1}^{q_2} \Gamma_n(x,y) (1 - \mathbf{1}_{B_n}(x)) \mathbf{1}_{B_n}(y) \\
&\quad + \lim_{n\to\infty} \int_{p_1}^{p_2} \int_{q_1}^{q_2} \Gamma_n(x,y) (1 - \mathbf{1}_{B_n}(x)) (1 - \mathbf{1}_{B_n}(y)) \, .
\end{aligned} \tag{10}
$$

Next we rewrite the integral following the first limit on the right-hand side of (10). To this end, we use the notation from (2) together with the obvious differentiation formula

$$
(\alpha_{\mathcal{J}_n,1})'(x) = \mathbf{1}_{B_n}(x) \quad \text{for a.e. } x \in I \tag{11}
$$

(and also, we use the fact that $\alpha_{\mathcal{J}_n,1|B_n}$ is an almost-bijection from B_n onto the interval $[0, \int_0^1 \mathbf{1}_{B_n}(y)]$, and so it makes sense to talk about its inverse). We have

$$\begin{aligned}
& \int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)\mathbf{1}_{B_n}(x)\mathbf{1}_{B_n}(y) \\
\boxed{\text{integration by substitution}} \quad &= \int_{\alpha_{\mathcal{J}_n,1}(p_1)}^{\alpha_{\mathcal{J}_n,1}(p_2)}\int_{\alpha_{\mathcal{J}_n,1}(q_1)}^{\alpha_{\mathcal{J}_n,1}(q_2)} \Gamma_n(\alpha^{-1}_{\mathcal{J}_n,1}(x), \alpha^{-1}_{\mathcal{J}_n,1}(y)) \\
\boxed{\gamma_{\mathcal{J}_n}(x)=\alpha_{\mathcal{J}_n,1}(x) \text{ for every } x\in B_n} \quad &= \int_{\alpha_{\mathcal{J}_n,1}(p_1)}^{\alpha_{\mathcal{J}_n,1}(p_2)}\int_{\alpha_{\mathcal{J}_n,1}(q_1)}^{\alpha_{\mathcal{J}_n,1}(q_2)} \Gamma_n(\gamma^{-1}_{\mathcal{J}_n}(x), \gamma^{-1}_{\mathcal{J}_n}(y)) \\
&= \int_{\alpha_{\mathcal{J}_n,1}(p_1)}^{\alpha_{\mathcal{J}_n,1}(p_2)}\int_{\alpha_{\mathcal{J}_n,1}(q_1)}^{\alpha_{\mathcal{J}_n,1}(q_2)} {}_{\mathcal{J}_n}\Gamma_n(x,y)\ .
\end{aligned} \tag{12}$$

Therefore, we have

$$\begin{aligned}
& \left| \int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)\mathbf{1}_{B_n}(x)\mathbf{1}_{B_n}(y) - \int_{\theta(p_1)}^{\theta(p_2)}\int_{\theta(q_1)}^{\theta(q_2)} {}_{\mathcal{J}_n}\Gamma_n(x,y) \right| \\
\overset{(12)}{=} & \left| \int_{\alpha_{\mathcal{J}_n,1}(p_1)}^{\alpha_{\mathcal{J}_n,1}(p_2)}\int_{\alpha_{\mathcal{J}_n,1}(q_1)}^{\alpha_{\mathcal{J}_n,1}(q_2)} {}_{\mathcal{J}_n}\Gamma_n(x,y) - \int_{\theta(p_1)}^{\theta(p_2)}\int_{\theta(q_1)}^{\theta(q_2)} {}_{\mathcal{J}_n}\Gamma_n(x,y) \right| \\
\le & |\alpha_{\mathcal{J}_n,1}(p_1)-\theta(p_1)| + |\alpha_{\mathcal{J}_n,1}(p_2)-\theta(p_2)| + |\alpha_{\mathcal{J}_n,1}(q_1)-\theta(q_1)| + |\alpha_{\mathcal{J}_n,1}(q_2)-\theta(q_2)|\ .
\end{aligned} \tag{13}$$

The fact that $\mathbf{1}_{B_n} \overset{w^*}{\to} \psi$ immediately implies that $\alpha_{\mathcal{J}_n,1}(x) \to \theta(x)$ for every $x \in I$, and so we conclude that the right-hand side, and thus also the left-hand side, of (13), tends to 0. Therefore (note that the following limits exist as $\Gamma_n \overset{w^*}{\to} W$)

$$\begin{aligned}
\lim_{n\to\infty}\int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)\mathbf{1}_{B_n}(x)\mathbf{1}_{B_n}(y) &= \lim_{n\to\infty}\int_{\theta(p_1)}^{\theta(p_2)}\int_{\theta(q_1)}^{\theta(q_2)} {}_{\mathcal{J}_n}\Gamma_n(x,y) \\
\boxed{{}_{\mathcal{J}_n}\Gamma_n \overset{w^*}{\to} \widetilde{W}} \quad &= \int_{\theta(p_1)}^{\theta(p_2)}\int_{\theta(q_1)}^{\theta(q_2)} \widetilde{W}(x,y) \\
\boxed{\text{integration by substitution}} \quad &= \int_{p_1}^{p_2}\int_{q_1}^{q_2} \widetilde{W}(\theta(x),\theta(y))\psi(x)\psi(y)\ .
\end{aligned} \tag{14}$$

In a very analogous way as we derived (14), one can verify that

$$\lim_{n\to\infty}\int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)\mathbf{1}_{B_n}(x)(1-\mathbf{1}_{B_n}(y)) = \int_{p_1}^{p_2}\int_{q_1}^{q_2} \widetilde{W}(\theta(x),\xi(y))\psi(x)(1-\psi(y))\ , \tag{15}$$

$$\lim_{n\to\infty}\int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)(1-\mathbf{1}_{B_n}(x))\mathbf{1}_{B_n}(y) = \int_{p_1}^{p_2}\int_{q_1}^{q_2} \widetilde{W}(\xi(x),\theta(y))(1-\psi(x))\psi(y)\ , \tag{16}$$

$$\lim_{n\to\infty}\int_{p_1}^{p_2}\int_{q_1}^{q_2} \Gamma_n(x,y)(1-\mathbf{1}_{B_n}(x))(1-\mathbf{1}_{B_n}(y)) = \int_{p_1}^{p_2}\int_{q_1}^{q_2} \widetilde{W}(\xi(x),\xi(y))(1-\psi(x))(1-\psi(y))\ . \tag{17}$$

By putting (10), (14), (15), (16) and (17) together, we get (9). □

Since the sets of the form $[p_1,p_2]\times[q_1,q_2]$ generate the Borel σ-algebra on I^2, we conclude from Claim 1 that for almost every $(x,y)\in I^2$ we have that

$$\begin{aligned}
W(x,y) =& \widetilde{W}(\theta(x),\theta(y))\psi(x)\psi(y) + \widetilde{W}(\theta(x),\xi(y))\psi(x)(1-\psi(y)) \\
&+ \widetilde{W}(\xi(x),\theta(y))(1-\psi(x))\psi(y) + \widetilde{W}(\xi(x),\xi(y))(1-\psi(x))(1-\psi(y))\ .
\end{aligned} \tag{18}$$

Note that the right-hand side of (18) is a convex combination of the four terms

$$\widetilde{W}(\theta(x),\theta(y))\ ,\quad \widetilde{W}(\theta(x),\xi(y))\ ,\quad \widetilde{W}(\xi(x),\theta(y))\ ,\quad \widetilde{W}(\xi(x),\xi(y))\ . \tag{19}$$

Therefore we have

$$\begin{aligned}
\mathrm{INT}_f(W) &= \int_0^1\int_0^1 f(W(x,y))\\
\boxed{f \text{ is concave}}\quad &\overset{(18)}{\geq} \int_0^1\int_0^1 f\left(\widetilde{W}(\theta(x),\theta(y))\right)\psi(x)\psi(y)\\
&\quad+\int_0^1\int_0^1 f\left(\widetilde{W}(\theta(x),\xi(y))\right)\psi(x)(1-\psi(y))\\
&\quad+\int_0^1\int_0^1 f\left(\widetilde{W}(\xi(x),\theta(y))\right)(1-\psi(x))\psi(y)\\
&\quad+\int_0^1\int_0^1 f\left(\widetilde{W}(\xi(x),\xi(y))\right)(1-\psi(x))(1-\psi(y))\\
\boxed{\text{integration by substitution}}\quad &= \int_0^{\theta(1)}\int_0^{\theta(1)} f\left(\widetilde{W}(x,y)\right)+\int_0^{\theta(1)}\int_{\theta(1)}^1 f\left(\widetilde{W}(x,y)\right)\\
&\quad+\int_{\theta(1)}^1\int_0^{\theta(1)} f\left(\widetilde{W}(x,y)\right)+\int_{\theta(1)}^1\int_{\theta(1)}^1 f\left(\widetilde{W}(x,y)\right)\\
&= \int_0^1\int_0^1 f\left(\widetilde{W}(x,y)\right)=\mathrm{INT}_f(\widetilde{W})\ .
\end{aligned} \tag{20}$$

To prove the "moreover" part, suppose that the equality $W(x,y)=\widetilde{W}(\theta(x),\theta(y))$ does not hold on a set of full measure. Then the convex combination (18) is not trivial on a set of positive measure. This is all we need as then we have a sharp inequality in (20) because f is strictly concave. $\square$

We do not use the next corollary right now but we will need it in Section 6.

Corollary 12. *Suppose that $f:[0,1]\to\mathbb{R}$ is an arbitrary continuous and strictly concave function. Suppose that $\Gamma_1,\Gamma_2,\Gamma_3,\ldots : I^2\to[0,1]$ is a sequence of graphons which converges to a graphon $W: I^2\to[0,1]$ in the weak* topology. Suppose that ℓ is a fixed natural number and that for every n, $\mathcal{J}_n$ is an ordered partition of I into ℓ sets $B_1^n,B_2^n,\ldots,B_\ell^n$. Then every $\widetilde{W}\in\mathbf{ACC}_{\mathrm{w}^*}({}_{\mathcal{J}_1}\Gamma_1,{}_{\mathcal{J}_2}\Gamma_2,{}_{\mathcal{J}_3}\Gamma_3,\ldots)$ satisfies $\mathrm{INT}_f(\widetilde{W})\leq\mathrm{INT}_f(W)$.*

Proof. For every natural number n and every $i\in\{1,\ldots,\ell\}$, we denote by $\mathcal{J}_n^i$ the ordered partition of I consisting of the sets $B_{\ell-i+1}^n,B_{\ell-i+2}^n,\ldots,B_\ell^n$ and $I\setminus\bigcup_{j=\ell-i+1}^{\ell}B_j^n$ (in this order). Consider these $\ell+1$ sequences of graphons:

$$\begin{aligned}
&\mathcal{S}_0\colon \Gamma_1,\Gamma_2,\Gamma_3,\ldots\\
&\mathcal{S}_1\colon\ {}_{\mathcal{J}_1^1}\Gamma_1,{}_{\mathcal{J}_2^1}\Gamma_2,{}_{\mathcal{J}_3^1}\Gamma_3,\ldots\\
&\mathcal{S}_2\colon\ {}_{\mathcal{J}_1^2}\Gamma_1,{}_{\mathcal{J}_2^2}\Gamma_2,{}_{\mathcal{J}_3^2}\Gamma_3,\ldots\\
&\quad\vdots\\
&\mathcal{S}_\ell\colon\ {}_{\mathcal{J}_1^\ell}\Gamma_1,{}_{\mathcal{J}_2^\ell}\Gamma_2,{}_{\mathcal{J}_3^\ell}\Gamma_3,\ldots\ ,
\end{aligned}$$

so that the sequence $\mathcal{S}_\ell$ is precisely ${}_{\mathcal{J}_1}\Gamma_1,{}_{\mathcal{J}_2}\Gamma_2,{}_{\mathcal{J}_3}\Gamma_3,\ldots$. Let us fix $\widetilde{W}\in\mathbf{ACC}_{\mathrm{w}^*}(\mathcal{S}_\ell)$. By passing to a subsequence, we may assume that the sequence $\mathcal{S}_i$ converges to some graphon W_i in the weak* topology for every $i=1,2,\ldots,\ell-1$. It remains to apply Lemma 11 ℓ-times in a row. First, we apply it on the sequence $\mathcal{S}_0$ of graphons and on the sequence $B_\ell^1, B_\ell^2, B_\ell^3, \ldots$ of subsets of I to conclude that $\mathrm{INT}_f(W_1) \leq \mathrm{INT}_f(W)$. Next, we apply it on the sequence $\mathcal{S}_1$ of graphons and on the sequence $B_{\ell-1}^1, B_{\ell-1}^2, B_{\ell-1}^3, \ldots$ of subsets of I to conclude that $\mathrm{INT}_f(W_2) \leq \mathrm{INT}_f(W_1) \leq \mathrm{INT}_f(W)$. In the last step, we apply it on the sequence $\mathcal{S}_{\ell-1}$ of graphons and on the sequence $B_1^1, B_1^2, B_1^3, \ldots$ of subsets of I to conclude that $\mathrm{INT}_f(\widetilde{W}) \leq \mathrm{INT}_f(W_{\ell-1}) \leq \ldots \leq \mathrm{INT}_f(W_1) \leq \mathrm{INT}_f(W)$. □

Now we can prove Theorem 3(a).

By passing to a subsequence, we may assume that the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ converges to W in the weak* topology. As W is not an accumulation point of the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-norm, there is $\varepsilon > 0$ and a natural number n_0 such that $\|\Gamma_n - W\|_\square \geq \varepsilon$ for every $n \geq n_0$. By passing to a subsequence, we may suppose that $\|\Gamma_n - W\|_\square \geq \varepsilon$ for every natural number n. By the definition of the cut-norm, there is a sequence $B_1, B_2, B_3, \ldots$ of subsets of I such that for every natural number n we have $\left|\int_{x\in B_n}\int_{y\in B_n}(\Gamma_n(x,y) - W(x,y))\right| \geq \varepsilon$. This means that either

$$
\begin{aligned}
(21)\qquad & \int_{x\in B_n}\int_{y\in B_n}\Gamma_n(x,y) \geq \int_{x\in B_n}\int_{y\in B_n} W(x,y) + \varepsilon \quad \text{or} \\
& \int_{x\in B_n}\int_{y\in B_n}\Gamma_n(x,y) \leq \int_{x\in B_n}\int_{y\in B_n} W(x,y) - \varepsilon\ .
\end{aligned}
$$

By passing to a subsequence, we may assume that only one of these two cases occurs. We stick to the case when (21) holds for every natural number n (the other case is analogous). By passing to a subsequence once again, we may assume that the sequence $\mathbf{1}_{B_1}, \mathbf{1}_{B_2}, \mathbf{1}_{B_3}, \ldots$ converges in the weak* topology to some $\psi\colon I \to [0,1]$. For every natural number n, let $\mathcal{J}_n$ be the ordered partition of I into two sets B_n and $I \setminus B_n$ (in this order). This allows us to define $\alpha_{\mathcal{J}_n,1}, \alpha_{\mathcal{J}_n,2}, \gamma_{\mathcal{J}_n}\colon I \to I$ as in (2), and versions ${}_{\mathcal{J}_1}\Gamma_1, {}_{\mathcal{J}_2}\Gamma_2, {}_{\mathcal{J}_3}\Gamma_3, \ldots$ of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$. We pass to a subsequence again to assure that the sequence ${}_{\mathcal{J}_1}\Gamma_1, {}_{\mathcal{J}_2}\Gamma_2, {}_{\mathcal{J}_3}\Gamma_3, \ldots$ is convergent in the weak* topology, and we denote the weak* limit by $\widetilde{W}$. Now Lemma 11 tells us that $\mathrm{INT}_f(\widetilde{W}) \leq \mathrm{INT}_f(W)$ and that to prove that this inequality is sharp, we only need to show that the equality $W(x,y) = \widetilde{W}(\theta(x),\theta(y))$ does not hold on a set of full measure. So to complete the proof, it suffices to prove the following claim.

Claim 2. *We have*

$$\int_0^1\int_0^1 \widetilde{W}(\theta(x),\theta(y))\psi(x)\psi(y) \geq \int_0^1\int_0^1 W(x,y)\psi(x)\psi(y) + \tfrac{1}{2}\varepsilon\ .$$

Proof of Claim 2. We have

$$
\begin{aligned}
\int_0^1\int_0^1 \widetilde{W}(\theta(x),\theta(y))\psi(x)\psi(y) &= \int_0^{\theta(1)}\int_0^{\theta(1)} \widetilde{W}(x,y)\\
\boxed{{}_{\mathcal{J}_n}\Gamma_n \overset{w^*}{\to} \widetilde{W}} \qquad &= \lim_{n\to\infty}\int_0^{\theta(1)}\int_0^{\theta(1)} {}_{\mathcal{J}_n}\Gamma_n(x,y)\\
\boxed{\text{for large enough } n, \text{ as } \alpha_{\mathcal{J}_n,1}(1)\to\theta(1)} \qquad &\geq \limsup_{n\to\infty}\int_0^{\alpha_{\mathcal{J}_n,1}(1)}\int_0^{\alpha_{\mathcal{J}_n,1}(1)} {}_{\mathcal{J}_n}\Gamma_n(x,y) - \tfrac{1}{2}\varepsilon\\
\boxed{\text{integration by substitution}} \qquad &= \limsup_{n\to\infty}\int_0^1\int_0^1 {}_{\mathcal{J}_n}\Gamma_n(\alpha_{\mathcal{J}_n,1}(x),\alpha_{\mathcal{J}_n,1}(y))\mathbf{1}_{B_n}(x)\mathbf{1}_{B_n}(y) - \tfrac{1}{2}\varepsilon\\
\boxed{\gamma_{\mathcal{J}_n}(x)=\alpha_{\mathcal{J}_n,1}(x) \text{ for every } x\in B_n} \qquad &= \limsup_{n\to\infty}\int_{B_n}\int_{B_n} {}_{\mathcal{J}_n}\Gamma_n(\gamma_{\mathcal{J}_n}(x),\gamma_{\mathcal{J}_n}(y)) - \tfrac{1}{2}\varepsilon\\
&= \limsup_{n\to\infty}\int_{B_n}\int_{B_n} \Gamma_n(x,y) - \tfrac{1}{2}\varepsilon\\
&\overset{(21)}{\geq} \limsup_{n\to\infty}\int_{B_n}\int_{B_n} W(x,y) + \tfrac{1}{2}\varepsilon\\
\boxed{\mathbf{1}_{B_n} \overset{w^*}{\to} \psi} \qquad &= \int_0^1\int_0^1 W(x,y)\psi(x)\psi(y) + \tfrac{1}{2}\varepsilon\ .
\end{aligned}
$$

□

Remark. *The initial step when we "shift the sets B_n to the left" crucially relies on the Euclidean order on I. This order is needless for the theory of graphons, i.e., graphons can be defined on a square of an arbitrary atomless separable probability space Ω. A linear order on Ω can be always introduced additionally, as Ω is measure-isomorphic to I. So, while our results work in full generality for an arbitrary Ω, we wonder if our argument can be modified so that the proof would naturally work without assuming a linear structure of the underlying probability space.*

5. Proof of Theorem 3(b)

The bulk of the proof is given after proving the following key lemma.

Lemma 13. *For every sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots : I^2\to[0,1]$ of graphons there exists a subsequence $\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots$ such that*

$$\inf\{\mathrm{INT}_f(W)\colon W\in \mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\} = \inf\{\mathrm{INT}_f(W)\colon W\in \mathbf{LIM}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\}\ .$$

Proof. We start by finding countably many subsequences $\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3,\ldots$ of the sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ such that for every natural number n we have:

(i) $\mathcal{S}_{n+1}$ is a subsequence of $\mathcal{S}_n$, and
(ii) there exists $W_{n+1}\in\mathbf{LIM}_{\mathrm{w}^*}(\mathcal{S}_{n+1})$ such that

$$\mathrm{INT}_f(W_{n+1}) < \inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\mathcal{S}_n)\} + \tfrac{1}{n}\ . \tag{22}$$

This is done by induction. In the first step, we just define the sequence $\mathcal{S}_1$ to be the original sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$. Next suppose that we have already defined the subsequence $\mathcal{S}_n$ for some natural number n. Then there is a graphon $W_{n+1}\in\mathbf{ACC}_{\mathrm{w}^*}(\mathcal{S}_n)$ such that

$$\mathrm{INT}_f(W_{n+1}) < \inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\mathcal{S}_n)\} + \tfrac{1}{n}\ .$$

Now we find a subsequence $\mathcal{S}_{n+1}$ of $\mathcal{S}_n$ such that some versions of the graphons from $\mathcal{S}_{n+1}$ converge to W_{n+1} in the weak* topology. This finishes the construction.

Now we use the diagonal method to define, for every natural number n, the graphon Γ_{k_n} to be the nth element of the sequence $\mathcal{S}_n$. Then we have for every n that

$$\begin{aligned}
& \inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\} \\
\boxed{\Gamma_{k_n},\Gamma_{k_{n+1}},\Gamma_{k_{n+2}},\ldots \text{ is a subsequence of } \mathcal{S}_n} \quad & \geq \inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\mathcal{S}_n)\} \\
& \overset{(22)}{>} \mathrm{INT}_f(W_{n+1})-\tfrac{1}{n} \\
\boxed{W_{n+1}\in\mathbf{LIM}_{\mathrm{w}^*}(\mathcal{S}_{n+1})\subset\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)} \quad & \geq \inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\}-\tfrac{1}{n}\ ,
\end{aligned}$$

and so

$$\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\}\geq\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots)\}\ .$$

The other inequality is trivial. □

We can now give the proof of Theorem 3(b).

By using Lemma 13 and by passing to a subsequence, we may assume that

$$\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}=\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}\ .$$

We construct the desired subsequence $\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots$ by the following construction.

In the first step, we find a graphon $W_1\in\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ such that

$$\mathrm{INT}_f(W_1)<\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}+1\ .$$

By Lemma 6, there is a partition $\mathcal{J}_1$ of I into finitely many intervals of positive measure such that $|\mathrm{INT}_f(W_1)-\mathrm{INT}_f(W_1^{\bowtie\mathcal{J}_1})|<1$. Then we clearly have

$$\mathrm{INT}_f(W_1^{\bowtie\mathcal{J}_1})<\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}+2\ .$$

By Lemma 7, the graphon $W_1^{\bowtie\mathcal{J}_1}$ is also an element of the set $\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$, and so there is a sequence $\Gamma_1^1,\Gamma_2^1,\Gamma_3^1,\ldots$ of versions of $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ that converges to $\widetilde{W}_1:=W_1^{\bowtie\mathcal{J}_1}$ in the weak* topology. We define $\Gamma_{k_1}:=\Gamma_1$, and we also define a sequence $q_1^1,q_2^1,q_3^1,\ldots$ to be the increasing sequence of all natural numbers.

Now fix a natural number n and suppose that we have already defined a finite subsequence $\Gamma_{k_1},\Gamma_{k_2},\ldots,\Gamma_{k_n}$ of $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$. Suppose also that for every $1\leq i\leq n$, we have already constructed

(i) a step-graphon $\widetilde{W}_i$ with steps given by some partition $\mathcal{J}_i$ of I into finitely many intervals of positive measure such that $\mathcal{J}_i$ is a refinement of $\mathcal{J}_{i-1}$ (if $i>1$) and such that

$$\mathrm{INT}_f(\widetilde{W}_i)<\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}+\tfrac{2}{i}\ ,$$

and

(ii) an increasing sequence $q_1^i,q_2^i,q_3^i,\ldots$ of natural numbers which is a subsequence of $q_1^{i-1},q_2^{i-1},q_3^{i-1},\ldots$ (if $i>1$), together with a sequence $\Gamma^i_{q_1^i},\Gamma^i_{q_2^i},\Gamma^i_{q_3^i},\ldots$ of versions of $\Gamma_{q_1^i},\Gamma_{q_2^i},\Gamma_{q_3^i},\ldots$ which converges to $\widetilde{W}_i$ in the weak* topology and such that (if $i>1$) for every natural number j and for every intervals $K,L\in\mathcal{J}_{i-1}$ it holds that

$$\int_K\int_L\Gamma^i_{q_j^i}(x,y)=\int_K\int_L\Gamma^{i-1}_{q_j^i}(x,y)\ .$$

Then we find a graphon $\overline{W}_{n+1}\in\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ such that

$$\mathrm{INT}_f(\overline{W}_{n+1})<\inf\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}+\tfrac{1}{n+1}\ .$$

Find a sequence $\overline{\Gamma}^{n+1}_{q_1^n},\overline{\Gamma}^{n+1}_{q_2^n},\overline{\Gamma}^{n+1}_{q_3^n},\ldots$ of versions of $\Gamma_{q_1^n},\Gamma_{q_2^n},\Gamma_{q_3^n},\ldots$ which converges to $\overline{W}_{n+1}$ in the weak* topology. For every natural number j, let $\phi_j:I\to I$ be the measure-preserving almost bijection satisfying $\overline{\Gamma}^{n+1}_{q^n_j}(x,y)=\Gamma^n_{q^n_j}(\phi_j^{-1}(x),\phi_j^{-1}(y))$ for a.e. $(x,y)\in I^2$ (such an almost-bijection exists as both $\overline{\Gamma}^{n+1}_{q^n_j}$ and $\Gamma^n_{q^n_j}$ are versions of the same graphon $\Gamma_{q^n_j}$). Let us fix some order of the sets from the partition $\mathcal{J}_n$. For every j, let $\mathcal{I}_j$ be the ordered partition of I consisting of the sets $\phi_j(K)$, $K\in\mathcal{J}_n$, with the order given by the order of the sets from $\mathcal{J}_n$. Let $r_1,r_2,r_3,\ldots$ be a subsequence of $q^n_1,q^n_2,q^n_3,\ldots$ such that for every $K\in\mathcal{J}_n$, the sequence $\mathbf{1}_{\phi_1(K)},\mathbf{1}_{\phi_2(K)},\mathbf{1}_{\phi_3(K)},\ldots$ is convergent in the weak* topology. Find an accumulation point W_{n+1} of the sequence ${}_{\mathcal{I}_1}\overline{\Gamma}^{n+1}_{r_1},{}_{\mathcal{I}_2}\overline{\Gamma}^{n+1}_{r_2},{}_{\mathcal{I}_3}\overline{\Gamma}^{n+1}_{r_3},\ldots$ (in the weak* topology). By Corollary 12, we have

$$\mathrm{INT}_f(W_{n+1})\le\mathrm{INT}_f(\overline{W}_{n+1})<\inf\left\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\right\}+\tfrac{1}{n+1}\;.$$

Let $s_1,s_2,s_3,\ldots$ be a subsequence of $r_1,r_2,r_3,\ldots$ such that the sequence ${}_{\mathcal{I}_1}\overline{\Gamma}^{n+1}_{s_1},{}_{\mathcal{I}_2}\overline{\Gamma}^{n+1}_{s_2},{}_{\mathcal{I}_3}\overline{\Gamma}^{n+1}_{s_3},\ldots$ converges to W_{n+1} in the weak* topology. Note that for every natural number j and for every intervals $K,L\in\mathcal{J}_n$, it holds that

$$\begin{aligned}\int_K\int_L {}_{\mathcal{I}_j}\overline{\Gamma}^{n+1}_{s_j}(x,y)=\int_{\phi_J(K)}\int_{\phi_j(L)}\overline{\Gamma}^{n+1}_{s_j}(x,y)&=\int_{\phi_J(K)}\int_{\phi_j(L)}\Gamma^n_{s_j}(\phi_j^{-1}(x),\phi_j^{-1}(y))\\&=\int_K\int_L\Gamma^n_{s_j}(x,y)\;.\end{aligned}\tag{23}$$

By Lemma 6, there is a partition $\mathcal{J}_{n+1}$ of I into finitely many intervals of positive measure such that $\mathcal{J}_{n+1}$ is a refinement of $\mathcal{J}_n$ and such that $|\mathrm{INT}_f(W_{n+1})-\mathrm{INT}_f(W_{n+1}^{\Join\mathcal{J}_{n+1}})|<\frac{1}{n+1}$. Then we clearly have

$$\mathrm{INT}_f(W_{n+1}^{\Join\mathcal{J}_{n+1}})<\inf\left\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\right\}+\tfrac{2}{n+1}\;.$$

By Lemma 7, the graphon $\widetilde{W}_{n+1}:=W_{n+1}^{\Join\mathcal{J}_{n+1}}$ is a limit (in the weak* topology) of the sequence of some versions $\Gamma^{n+1}_{s_1},\Gamma^{n+1}_{s_2},\Gamma^{n+1}_{s_3},\ldots$ of the graphons ${}_{\mathcal{I}_1}\overline{\Gamma}^{n+1}_{s_1},{}_{\mathcal{I}_2}\overline{\Gamma}^{n+1}_{s_2},{}_{\mathcal{I}_3}\overline{\Gamma}^{n+1}_{s_3},\ldots$. By the "moreover" part of Lemma 7, we may further assume that for every natural number j and for every intervals $P,Q\in\mathcal{J}_{n+1}$, we have

$$\int_P\int_Q {}_{\mathcal{I}_j}\overline{\Gamma}^{n+1}_{s_j}(x,y)=\int_P\int_Q\Gamma^{n+1}_{s_j}(x,y)\;,$$

which, together with (23), easily implies that for every natural number j and for every intervals $K,L\in\mathcal{J}_n$ it holds

$$\int_K\int_L\Gamma^{n+1}_{s_j}(x,y)=\int_K\int_L\Gamma^n_{s_j}(x,y)\;.\tag{24}$$

We define $\Gamma_{k_{n+1}}:=\Gamma_{s^{n+1}_{n+1}}$, and we also define the sequence $q^{n+1}_1,q^{n+1}_2,q^{n+1}_3,\ldots$ to be the sequence $s_1,s_2,s_3,\ldots$. This completes the construction of the sequence $\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots$.

Now let $W_{\min}$ be an arbitrary accumulation point (in the weak* topology) of the sequence $\Gamma_{k_1},\Gamma_{k_2},\Gamma_{k_3},\ldots$, so that in particular $W_{\min}\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. It suffices to show that it holds for every n that $\mathrm{INT}_f(W_{\min})\le\mathrm{INT}_f(\widetilde{W}_n)$ as then we clearly have by our choice of the graphons $\widetilde{W}_1,\widetilde{W}_2,\widetilde{W}_3,\ldots$ that

$$\mathrm{INT}_f(W_{\min})=\min\{\mathrm{INT}_f(W)\colon W\in\mathbf{ACC}_{\mathrm{w}^*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\}\;.$$

But for every three natural numbers $n<m$ and j and for every intervals $K,L\in\mathcal{J}_n$ it holds by (ii) that

$$\int_K\int_L\Gamma^m_{q^m_j}(x,y)=\int_K\int_L\Gamma^n_{q^m_j}(x,y)\;,$$

and so (as $\Gamma^n_{q^n_j} \xrightarrow{w^*} \widetilde{W}_n$ as $j \to \infty$ for every n)

$$\int_K \int_L \widetilde{W}_m(x,y) = \int_K \int_L \widetilde{W}_n(x,y) \ .$$

It follows that for every n it holds

$$\int_K \int_L W_{\min}(x,y) = \int_K \int_L \widetilde{W}_n(x,y) \ .$$

The rest follows by Lemma 10.

6. Proof of Proposition 4

As promised, we give two proofs of Proposition 4. The first one is somewhat quicker, but uses a theorem of Borgs, Chayes, and Lovász [2] about uniqueness of graph limits. More precisely, the theorem states that if $U' : I^2 \to [0,1]$ and $U'' : I^2 \to [0,1]$ are two cut-norm limits of versions $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ and $\Gamma''_1, \Gamma''_2, \Gamma''_3, \ldots$ of a graphon sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$, then there exists a graphon $U^* : I^2 \to [0,1]$ that is a cut-norm limit of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$, and measure preserving transformations $\psi', \psi'' : I \to I$ such that for almost every $(x,y) \in I^2$, $U'(x,y) = U^*(\psi'(x), \psi'(y))$ and $U''(x,y) = U^*(\psi''(x), \psi''(y))$. Since then, the result was proven in several different ways, see [8, p.221]. Also, let us note that while all known proofs of the Borgs–Chayes–Lovász theorem are complicated, none uses the compactness of the space of graphons or the Regularity lemma. So, using this result as a blackbox, we still obtain a self-contained characterization of cut-norm limits in terms of weak* limits.

So, suppose that $W : I^2 \to [0,1]$ is a limit of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-norm. By Theorem 3 and by passing to a subsequence, we may assume that there exists a minimizer $W' : I^2 \to [0,1]$ of $\mathrm{INT}_f(\cdot)$ over $\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ which is a limit of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-norm. Therefore, the Borgs–Chayes–Lovász theorem tells us that there exists a graphon $W^* : I^2 \to [0,1]$ and measure preserving maps $\psi, \psi' : I \to I$ such that $W(x,y) = W^*(\psi(x), \psi(y))$ and $W'(x,y) = W^*(\psi'(x), \psi'(y))$ for almost every $(x,y) \in I^2$. Since ψ and ψ' are measure preserving, we get $\mathrm{INT}_f(W) = \mathrm{INT}_f(W^*)$ and $\mathrm{INT}_f(W') = \mathrm{INT}_f(W^*)$. This finishes the proof. □

Let us now give a self-contained proof of Proposition 4. By Theorem 3 and by passing to a subsequence, we may assume that there exists a minimizer $W' : I^2 \to [0,1]$ of $\mathrm{INT}_f(\cdot)$ over $\mathbf{LIM}_{\mathrm{w}^*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ which is a limit of versions $\Gamma'_1, \Gamma'_2, \Gamma'_3, \ldots$ of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ in the cut-norm. Suppose that W is a graphon with $\mathrm{INT}_f(W) > \mathrm{INT}_f(W')$. This in particular means that there exists $\delta > 0$ so that

$$\|W' - U\|_1 > \delta \tag{25}$$

for any version U of W. We claim that there are no versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ that converge to W in the cut-norm. Indeed, suppose that such versions $\Gamma^*_1, \Gamma^*_2, \Gamma^*_3, \ldots$ exist. Observe that $\delta_1(\Gamma'_n, \Gamma^*_n) = 0$ for each n (in fact, the infimum in the definition of δ_1 is attained). Now, [11, Lemma 2.11][4] tells us that

$$0 = \liminf_n 0 = \liminf_n \delta_1(\Gamma'_n, \Gamma^*_n) \geq \delta_1(W', W) \ ,$$

which is a contradiction to (25). □

[4] Let us stress that [11, Lemma 2.11] does not rely on the Borgs–Chayes–Lovász theorem, and has a self-contained, one-page proof.

7. Concluding remarks

7.1. **Specific concave and convex functions.** Perhaps the most natural choice of continuous concave function is the binary entropy H.

An equivalent characterization to our main result is that the limit graphons are the weak* limits that *maximize* INT_g for a strictly *convex* function g. The most interesting instance of this version of the statement is that the limit graphons are weak* limits maximizing the L^2-norm.

7.2. **Regularity lemmas as a corollary.** While the cut-distance is most tightly linked to the weak regularity lemma of Frieze and Kannan [7], a short reduction given in [10] shows that Theorem 2 implies also Szemerédi's regularity lemma [13], and its "superstrong" form, [1]. So, we believe that the simplest proofs of these regularity lemmas are using the approach from this paper.[5]

The most remarkable difference of the current approach is that it does not use any index-pumping. Recall that in the conventional proofs of regularity lemmas one keeps refining a partition, and an index-pumping argument is needed to show that the number of refinements is bounded. In comparison, in our proof one refinement is sufficient for the argument. Such a shortcut is available only in the limit setting, it seems.

7.3. **Hypergraphs.** The theory of limits of dense hypergraphs of a fixed uniformity was worked out in [6] (using ultraproduct techniques) and in [15] (using hypergraph regularity lemma techniques), and is substantially more involved. It seems that the currect approach may generalize to the hypergraph setting. This is currently work in progress.

Acknowledgements

We thank Dan Král and Oleg Pikhurko for encouraging conversations on the subject.

Appendix A. The weak* topology

Suppose that X is a Banach space and denote by X^* its dual. Then the weak* topology on X^* is the coarsest topology on X^* such that all mappings of the form $X^* \ni x^* \mapsto x^*(x)$, $x \in X$, are continuous. Recall that if the space X is separable then by the sequential Banach–Alaoglu Theorem (see e.g. [14, Theorem 1.9.14]), the unit ball of X^* is sequentially compact. This means that every bounded sequence of elements of the dual space X^* contains a weak*-convergent subsequence.

In this paper, we are interested in the case when X is the Banach space $L^1(\Omega)$ of all integrable functions on some probability space Ω. (Depending on our needs, the probability space Ω will be chosen to be either the unit interval I equipped with the one-dimensional Lebesgue measure or the unit square I^2 equipped with the two-dimensional Lebesgue measure). The space $L^1(\Omega)$ is equipped with the norm $\|f\|_1 = \int_\Omega |f(x)|$, $f \in L^1(\Omega)$. In this setting, the dual $X^* = (L^1(\Omega))^*$ is isometric to the space $L^\infty(\Omega)$ of all bounded measurable functions on Ω, equipped with the norm $\|g\|_\infty = \operatorname{ess\,sup}_{x\in\Omega} |g(x)|$. The duality between $L^1(\Omega)$ and $L^\infty(\Omega)$ is given by the formula $\langle g, f\rangle = \int_\Omega f(x)g(x)$ for $g \in L^\infty(\Omega)$ and $f \in L^1(\Omega)$. This means that a sequence $g_1, g_2, g_3, \ldots$ of elements of $L^\infty(\Omega)$ converges to $g \in L^\infty(\Omega)$ if and only if $\lim_{n\to\infty} \int_\Omega f(x)g_n(x) = \int_\Omega f(x)g(x)$ for every $f \in L^1(\Omega)$.

Now consider the Banach space $X = L^1(I^2)$ of all integrable functions defined on the unit square I^2 (which is equipped with the two-dimensional Lebesgue measure). Standard arguments show that the weak* topology on its dual space $L^\infty(I^2)$ can be equivalently generated by mappings of the form $L^\infty(I^2) \ni g \mapsto \int_A \int_B g(x,y)$ where A, B are measurable subsets of I. That is, the weak* topology can be equivalently generated only by characteristic functions of measurable rectangles (instead of all integrable functions on I^2). If we restrict this topology only to the space of all

[5] With a notable drawback that we do not obtain any quantitative bounds.

graphons $W : I^2 \to [0,1]$ defined on I^2 then it is easy to see that this restricted topology is generated only by mappings of the form $W \mapsto \int_A \int_A W(x,y)$ where A is a measurable subsets of I (this is because each graphon is symmetric by the definition). This is the topology we refer to when we talk about convergence of graphons in the weak* topology. So this means that a sequence $W_1, W_2, W_3, \dots$ of graphons defined on I^2 converges to a graphon W defined on I^2 if and only if $\lim_{n\to\infty} \int_A \int_A W_n(x,y) = \int_A \int_A W(x,y)$ for every measurable subset A of I. Note that the space of all graphons defined on I^2 is a closed subset of the unit ball of $L^\infty(I^2)$, and so it is sequentially compact by the sequential Banach–Alaoglu Theorem (as the space $L^1(I^2)$ is separable).

While crucial to our arguments, it is worth noting that the Banach–Alaoglu Theorem is not a particularly deep statement and follows easily from Tychonoff's theorem for powers of compact spaces.

References

[1] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.

[2] C. Borgs, J. Chayes, and L. Lovász. Moments of Two-Variable Functions and the Uniqueness of Graph Limits. *J. Geom. and Func. Anal*, 19:1597–1619, 2009.

[3] C. Borgs, J. Chayes, and L. Lovász. Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.*, 19(6):1597–1619, 2010.

[4] 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.

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

[6] G. Elek and B. Szegedy. A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.*, 231(3-4):1731–1772, 2012.

[7] A. Frieze and R. Kannan. Quick Approximation to Matrices and Applications. *Combinatorica*, 19(2):175–220, 1999.

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

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

[10] L. Lovász and B. Szegedy. Szemerédi's Lemma for the analyst. *J. Geom. and Func. Anal*, 17:252–270, 2007.

[11] L. Lovász and B. Szegedy. Testing properties of graphs and functions. *Israel J. Math.*, 178:113–156, 2010.

[12] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

[13] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[14] T. Tao. *An epsilon of room, I: real analysis*, volume 117 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Pages from year three of a mathematical blog.

[15] Y. Zhao. Hypergraph limits: a regularity approach. *Random Structures Algorithms*, 47(2):205–226, 2015.

Institute of Mathematics, Czech Academy of Sciences. Žitná 25, 110 00, Praha, Czech Republic.

The Institute of Mathematics of the Czech Academy of Sciences is supported by RVO:67985840.

E-mail address: dolezal@math.cas.cz

Institut für Geometrie, TU Dresden, 01062 Dresden, Germany.
E-mail address: honzahladky@gmail.com