arXiv:1806.07368v3 [math.CO] 15 Apr 2020

RELATING THE CUT DISTANCE AND THE WEAK* TOPOLOGY FOR GRAPHONS

MARTIN DOLEŽAL, JAN GREBÍK, JAN HLADKÝ, ISRAEL ROCHA, VÁCLAV ROZHOŇ

ABSTRACT. The theory of graphons is ultimately connected with the so-called cut norm. In this paper, we approach the cut norm topology via the weak* topology (when considering a predual of L^1-functions). We prove that a sequence $W_1, W_2, W_3, \ldots$ of graphons converges in the cut distance if and only if we have equality of the sets of weak* accumulation points and of weak* limit points of all sequences of graphons $W_1', W_2', W_3', \ldots$ that are weakly isomorphic to $W_1, W_2, W_3, \ldots$. We further give a short descriptive set theoretic argument that each sequence of graphons contains a subsequence with the property above. This in particular provides an alternative proof of the theorem of Lovász and Szegedy about compactness of the space of graphons. We connect these results to «multiway cut» characterization of cut distance convergence from [Ann. of Math. (2) 176 (2012), no. 1, 151-219].

These results are more naturally phrased in the Vietoris hyperspace $K(\mathcal{W}_0)$ over graphons with the weak* topology. We show that graphons with the cut distance topology are homeomorphic to a closed subset of $K(\mathcal{W}_0)$, and deduce several consequences of this fact.

From these concepts a new order on the space of graphons emerges. This order allows to compare how structured two graphons are. We establish basic properties of this «structuredness order».

1. INTRODUCTION

Graphons emerged from the work of Borgs, Chayes, Lovász, Sós, Szegedy, and Vesztergombi [18, 4] on limits of sequences of finite graphs. We write $\mathcal{W}_0$ for the space of all *graphons*, i.e., all symmetric measurable functions from Ω^2 to $[0,1]$, after identifying graphons that are equal almost everywhere. Here as well as in the rest of the paper, Ω is an arbitrary separable atomless probability space with probability measure ν. While it is meaningful to investigate the space $\mathcal{W}_0$ with respect to several metrics and topologies, the two that relate the most to graph theory are the metrics $d_\square$ and $\delta_\square$ based on the so-called *cut norm* defined on the space $L^1(\Omega^2)$ by

$$\|Y\|_\square = \sup_{S,T\subseteq\Omega}\left|\int_{S\times T} Y\right| \quad \textit{for each } Y \in L^1(\Omega^2)\,.$$

Given $U, W \in \mathcal{W}_0$ we set

$$d_\square(U,W) := \|U-W\|_\square = \sup_{S,T\subseteq\Omega}\left|\int_{S\times T} U - \int_{S\times T} W\right|\,, \text{ and}$$

$$\delta_\square(U,W) := \inf_{\varphi} d_\square(U, W^\varphi)\,, \tag{1.1}$$

Key words and phrases. graphon; graph limit; cut norm; weak* convergence.

Research of *Doležal* was supported by the GAČR project No. 17-27844S and RVO: 67985840. *Grebík* was supported by the GAČR project 17-33849L and RVO: 67985840. *Hladký* was supported by the Alexander von Humboldt Foundation and by the GAČR project No. 18-01472Y and RVO: 67985840. *Rocha* and *Rozhoň* were supported by the Czech Science Foundation, grant number GJ16-07822Y.

where φ ranges over all measure preserving bijections of Ω and the graphon W^φ is defined by

$$W^\varphi(x,y) = W(\varphi(x),\varphi(y)) \; . \tag{1.2}$$

We call $d_\square$ the *cut norm distance* and $\delta_\square$ the *cut distance*. We call graphons of the form W^φ *versions* of W. Passing to a version is an infinitesimal counterpart to considering another adjacency matrix of a graph, in which the vertices are reordered.

The key property of the space $\mathcal{W}_0$ is its compactness with respect to the cut distance $\delta_\square$. The result was first proven by Lovász and Szegedy [18] using the regularity lemma,[a] and then by Elek and Szegedy [10] using ultrafilter techniques, by Austin [1] and Diaconis and Janson [6] using the theory of exchangeable random graphs, and finally by Doležal and Hladký [8] by optimizing a suitable parameter over the set of weak* limits.

Theorem 1.1. *For every sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ of graphons there is a subsequence $\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots$ and a graphon Γ such that $\delta_\square(\Gamma_{n_i},\Gamma)\to 0$.*

1.1. **Overview of the results.** We view graphons as functions in the Banach space $L^\infty(\Omega^2)$, with a predual Banach space $L^1(\Omega^2)$ to which we associate the concept of weak* convergence. That means that a sequence of graphons $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ *converges weak** to a graphon W if we have

$$\sup_{Q\subseteq\Omega^2}\lim_{n\to\infty}\int_Q \Gamma_n - \int_Q W = 0 \; . \tag{1.3}$$

Since the sigma-algebra of all measurable sets $Q\subseteq\Omega^2$ is generated by sets of the form $S\times T$, where S and T are measurable subsets of Ω, we can equivalently rewrite (1.3) in a way which is more convenient for us,

$$\sup_{S,T\subseteq\Omega}\lim_{n\to\infty}\int_{S\times T} \Gamma_n - \int_{S\times T} W = 0 \; . \tag{1.4}$$

The weak* topology is weaker than the topology generated by $d_\square$, of which the former can be viewed as a certain uniformization. Indeed, recall that a sequence of graphons $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ converges to W in the cut norm if

$$\lim_{n\to\infty}\sup_{S,T\subseteq\Omega}\left\{\int_{S\times T} \Gamma_n - \int_{S\times T} W\right\} = 0 \; , \tag{1.5}$$

that is, (1.4) and (1.5) differ only in the order of the limit and the supremum.

In Section 3 and Section 4, which we consider the main contribution of the paper, we show that the interplay between the cut distance and weak* topology creates a rich theory.

In particular, we can make use of the weak* convergence to prove Theorem 1.1. To this end, we look at the set $\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ of all weak* accumulation points of sequences

$$\left\{\Gamma'_1,\Gamma'_2,\Gamma'_3,\ldots : \Gamma'_n \text{ is a version of } \Gamma_n\right\} \; .$$

Similarly, denote by $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ the set of all graphons W 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 $\Gamma'_1,\Gamma'_2,\Gamma'_3,\ldots$. Note that equivalently, we could have required $\Gamma'_1,\Gamma'_2,\Gamma'_3,\ldots$ to be weakly isomorphic to $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$, rather than being versions of $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$. The set $\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ is non-empty by the Banach–Alaoglu Theorem. In the set $\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ we cleverly select one graphon Γ. The selection is done so that in addition to being a weak* accumulation point, Γ is also a cut

[a] see also [19] and [20] for variants of this approach

distance accumulation point (the latter being clearly a stronger property). In [8], Doležal and Hladký carried out a similar program[b] where they showed that for the «clever selection» we can take Γ as the maximizer of an arbitrary graphon parameter of the form

$$\mathrm{INT}_f(W) := \int_x \int_y f\left(W(x,y)\right) \tag{1.6}$$

where $f : [0,1] \to \mathbb{R}$ is a fixed but arbitrary continuous strictly convex function.[c]

Our main result, Theorem 3.5, says that a sequence of graphons $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ is cut distance convergent if and only if $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. This is complemented by Theorem 3.3 which says that from any sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ of graphons, we can choose a subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ such that $\mathbf{LIM}_{w*}(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots) = \mathbf{ACC}_{w*}(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots)$. In particular, this yields Theorem 1.1. In Section 7 we show that Theorem 3.5 is actually equivalent to one of the main results of [5] on so-called «multiway cuts».

It turns out that these results can be naturally phrased in terms of the so-called Vietoris hyperspace $K(\mathcal{W}_0)$ over graphons with the weak* topology (see Section 2.5 for definitions). To each graphon $W : \Omega^2 \to [0,1]$, we associate its *envelope* $\langle W \rangle = \mathbf{ACC}_{w*}(W, W, W, \ldots)$ which is a subset of $L^\infty\left(\Omega^2\right)$. We show that cut distance convergence of graphons is equivalent to convergence of the corresponding envelopes in $K(\mathcal{W}_0)$, and that envelopes form a closed set in $K(\mathcal{W}_0)$. As $K(\mathcal{W}_0)$ is known to be compact, this connection in particular provides an alternative proof of Theorem 1.1. However, the transference between the space of graphons and $K(\mathcal{W}_0)$ has other applications.

From these proofs a new partial order on the space of graphons naturally emerges. We say that U is *more structured* than W if $\langle U \rangle \supsetneq \langle W \rangle$, and write $U \succ W$. One illustrative example is to take U to be a complete balanced bipartite graphon and $W \equiv \frac{1}{2}$. Obviously, $\langle W \rangle = \{W\}$. By considering versions of U which create finer and finer chessboards, we have $\langle U \rangle \ni W$. Thus, $U \succeq W$. We establish basic properties of this «structuredness order». We investigate these properties in Section 8. In particular, in Proposition 8.5 we characterize minimal and maximal elements. In Section 4.4, we introduce the *range frequencies* of a graphon W as a pushforward probability measure on $[0,1]$ defined by

$$\mathbf{\Phi}_W(A) := \nu^{\otimes 2}\left(W^{-1}(A)\right) ,$$

for a measurable set $A \subseteq [0,1]$, and introduce a certain *flatness order* on these range frequencies. Roughly speaking, one probability measure on $[0,1]$ is flatter than another, if the former can be obtained by a certain sort of averaging of the latter. As we show in Proposition 4.17, this order is compatible with the structuredness order on the corresponding graphons. In [9] we use range frequencies to reprove in a very quick way the result of Doležal and Hladký (even for discontinuous functions, as mentioned in Footnote [c]).

The main motivation for introducing and studying the structuredness order is its connection to Theorem 1.1. Indeed, the «clever selection» in our proof of Theorem 1.1 is to take a maximal element (inside $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$) with respect to the structuredness order. However, the abstract theory of the weak* approach to cut norm convergence which we introduce in this paper has also applications in classical problems in extremal graph theory. More specifically,

[b]which in actuality is more complicated for reasons sketched in Section 3.2

[c]Most of [8] deals with *minimizing* $\mathrm{INT}_f(W)$ for a fixed continuous strictly *concave* function. This is obviously equivalent. Also, note that in [9] it is shown that the assumption of continuity is not needed.

one of the main results which we prove in [9] says that if a graph is «step Sidorenko» then it is «weakly norming», thus answering an open question from [14]. Our proof in [9] relies heavily on the theory of the structuredness order introduced in the current paper, and we do not see an alternative proof which would avoid it.[d]

2. Preliminaries

2.1. **General notation.** We write $\stackrel{\varepsilon}{\approx}$ for equality up to ε. For example, $1 \stackrel{0.2}{\approx} 1.1 \stackrel{0.2}{\approx} 1.3$. We write P_k for a path on k vertices and C_k for a cycle on k vertices.

If A and B are measure spaces then we say that a map $f : A \to B$ is an *almost-bijection* if there exist measure zero sets $A_0 \subseteq A$ and $B_0 \subseteq B$ so that $f_{\restriction A \setminus A_0}$ is a bijection between $A \setminus A_0$ and $B \setminus B_0$. Note that in (1.1), we could have worked with measure preserving almost-bijections φ instead.

2.2. **Graphon basics.** Our notation is mostly standard, following [17]. We write $\mathcal{W}_0$ for the space of all graphons, that is, symmetric measurable functions from Ω^2 to $[0,1]$, modulo differences on null-sets.

Graphons U and W are called *weakly isomorphic* if $\delta_\square(U,W) = 0$. Note that in this case there is not necessarily a measure preserving bijection φ for which $d_\square(U, W^\varphi) = 0$; see [17, Figure 7.1]. In other words, being versions and being weakly isomorphic are two slightly different notions. Let us denote the compact space of graphons after the weak isomorphism factorization as $\widetilde{\mathcal{W}_0}$. For every $W \in \mathcal{W}_0$ we denote its equivalence class $[\![W]\!] \in \widetilde{\mathcal{W}_0}$.

The probability measure underlying Ω is ν. We write $\nu^{\otimes k}$ for the product measure on Ω^k.

Remark 2.1. *Every separable atomless probability space is isomorphic to the unit interval with the Lebesgue measure. While most of our arguments are abstract and work with an arbitrary separable atomless probability space Ω, there are some other, where we will assume that graphons are defined on the square of the unit interval, and then will make use of the usual order on* $[0,1]$.

If $W\colon \Omega^2 \to [0,1]$ is a graphon and φ, ψ are two measure preserving bijections of Ω then we use the short notation $W^{\psi\varphi}$ for the graphon $W^{\psi\circ\varphi}$, i.e. $W^{\psi\varphi}(x,y) = W(\psi(\varphi(x)), \psi(\varphi(y))) = W^\psi(\varphi(x), \varphi(y)) = \left(W^\psi\right)^\varphi(x,y)$ for $(x,y) \in \Omega^2$. Thus we have a right action of the group of all measure preserving isomorphisms of $[0,1]$ on $\mathcal{W}_0$.

The graphons that take values 0 or 1 almost everywhere are called *0-1 valued graphons*.

We call the quantity $\int_x \int_y W(x,y)$ the *edge density of* W. Recall also that for $x \in \Omega$, we have the *degree of* x in W defined as $\deg_W(x) = \int_y W(x,y)$. Recall that measurability of W gives that $\deg_W(x)$ exists for almost each $x \in \Omega$. We say that W is *p-regular* if for almost every $x \in \Omega$, $\deg_W(x) = p$.

2.2.1. *The stepping operator.* Suppose that $W : \Omega^2 \to [0,1]$ is a graphon. We say that W is a *step graphon* if W is constant on each $\Omega_i \times \Omega_j$, for a suitable a finite partition $\mathcal{P}$ of Ω, $\mathcal{P} = \{\Omega_1, \Omega_2, \ldots, \Omega_k\}$. We recall the definition of the stepping operator.

Definition 2.2. Suppose that $\Gamma : \Omega^2 \to [0,1]$ is a graphon. For a finite partition $\mathcal{P}$ of Ω, $\mathcal{P} = \{\Omega_1, \Omega_2, \ldots, \Omega_k\}$, we define a graphon $\Gamma^{\Join\mathcal{P}}$ by setting it on the rectangle $\Omega_i \times \Omega_j$ to be the constant $\frac{1}{\nu^{\otimes 2}(\Omega_i \times \Omega_j)} \int_{\Omega_i} \int_{\Omega_j} \Gamma(x,y)$. We allow graphons to have not well-defined values on null sets which handles the cases $\nu(\Omega_i) = 0$ or $\nu(\Omega_j) = 0$.

[d]See Remark 3.29 of [9].

In [17], a stepping is denoted by $\Gamma_{\mathcal{P}}$ rather than $\Gamma^{\Join\mathcal{P}}$.

Finally, we say that a graphon U *refines* a graphon W, if W is a step graphon for a suitable partition $\mathcal{P}$ of Ω, $\mathcal{P} = \{\Omega_1, \Omega_2, \ldots, \Omega_k\}$, and $U^{\Join\mathcal{P}} = W$.

2.3. **Topologies on $\mathcal{W}_0$.** There are several natural topologies on $\mathcal{W}_0$. The $\|\cdot\|_1$ topology inherited from the normed space $L^1(\Omega^2)$, the topology given by the $\|\cdot\|_\square$ norm, and the weak* topology (when $\mathcal{W}_0$ is viewed as a subset of the dual space of $L^1(\Omega^2)$ i.e. a subset of $L^\infty(\Omega^2)$). Note that $\mathcal{W}_0$ is closed in L^1. We write $d_1(\cdot,\cdot)$ for the distance derived from the $\|\cdot\|_1$ norm. Recall also that by the Banach–Alaoglu Theorem, $\mathcal{W}_0$ equipped with the weak* topology is compact and that the weak* topology on $\mathcal{W}_0$ is metrizable. We shall denote by $d_{\mathrm{w}^*}(\cdot,\cdot)$ any metric compatible with this topology. For example, we can take some countable dense measure subalgebra $\{A_n\}_{n\in\mathbb{N}}$ of all measurable subsets of Ω, and define

$$d_{\mathrm{w}^*}(U,W) := \sum_{n,k\in\mathbb{N}} 2^{-(n+k)} \left| \int_{A_n\times A_k} (U-W)\,\mathrm{d}\nu \right| . \tag{2.1}$$

The following fact summarizes the relation of the above topologies.

Fact 2.3. *The following identity maps are continuous: $(\mathcal{W}_0, d_1) \to (\mathcal{W}_0, d_\square) \to (\mathcal{W}_0, d_{\mathrm{w}^*})$.*

Proof. The continuity of the first map is an easy consequence of the definitions and the second is explained in Subsection 1.1. □

2.4. **Auxiliary results about L^1-spaces.** We prove two auxiliary lemmas about L^1-spaces. Lemma 2.4 is an easy result about functions that do not converge in L^1.

Lemma 2.4. *Suppose that Λ is a probability measure space with measure λ. If we have functions $g, g_1, g_2, g_3, \ldots : \Lambda \to [0,1]$ for which $g_n \overset{\|\cdot\|_1}{\not\longrightarrow} g$, then there exists an interval $J \subseteq [0,1]$ and a number $c > 0$ such that for the interval $J^+ := \{x + d : x \in J, d \in [-c, c]\}$ we have $\lambda\left(g^{-1}(J) \setminus g_n^{-1}(J^+)\right) \not\longrightarrow 0$.*

Proof. By passing to a subsequence, we may assume that there exists a constant $\varepsilon > 0$ so that for each n, $\|g_n - g\|_1 > \varepsilon$. Take $k := \lceil 4/\varepsilon \rceil$ and a partition of $[0,1]$ into k intervals $J_1, J_2, \ldots, J_k$ of lengths at most $\frac{\varepsilon}{4}$ (ordered from left to right; it is not important if they are open, closed, or semiopen). For $j \in [k]$, $L_j := g^{-1}(J_j)$. For each n, there exists a number $j(n) \in [k]$ so that we have $\int_{L_{j(n)}} |g_n - g| > \varepsilon \cdot \lambda(L_{j(n)})$. Observe that the strict inequality forces that $\lambda(L_{j(n)}) > 0$. In particular, we then have that

$$|g_n(x) - g(x)| > \frac{\varepsilon}{2} \tag{2.2}$$

for a set of points $x \in L_{j(n)}$ of measure at least $\frac{\varepsilon}{2} \cdot \lambda\left(L_{j(n)}\right)$. Let j be a number that repeats infinitely often in the sequence $j(1), j(2), j(3), \ldots$. By passing to a subsequence once again, we can assume that $j = j(1) = j(2) = \ldots$. We set $J := J_j$, $c := \frac{\varepsilon}{4}$, and J^+ as in the statement of the lemma. Observe that whenever $x \in L_j$ satisfies (2.2), then $g_n(x) \notin J^+$. Therefore, we conclude that for each n, $\lambda\left(g^{-1}(J) \setminus g_n^{-1}(J^+)\right) \geq \frac{\varepsilon}{2} \cdot \lambda\left(L_j\right)$. This concludes the proof. □

We can now state the second lemma of this section.

Lemma 2.5. *For every graphon $\Gamma : \Omega^2 \to [0,1]$ and every $\varepsilon > 0$ there exists a finite partition $\mathcal{P}$ of Ω such that $\left\|\Gamma - \Gamma^{\Join\mathcal{P}}\right\|_1 < \varepsilon$.*

For the proof of Lemma 2.5, the following fact will be useful.

Fact 2.6. *Suppose that $f \in L^1(\Lambda)$ is an arbitrary function on a finite measure space Λ with measure λ. Set $a := \frac{1}{\lambda(\Lambda)} \cdot \int_\Lambda f$. Then for each $b \in \mathbb{R}$ we have that $\|f-a\|_1 \le 2\|f-b\|_1$.*

Proof. We have

$$\begin{aligned}\|f-a\|_1 &= \int_\Lambda |f(x)-a| \le \int_\Lambda |f(x)-b| + \int_\Lambda |a-b| = \|f-b\|_1 + \lambda(\Lambda)\cdot|a-b| \\ &= \|f-b\|_1 + \left|\int_\Lambda (f(x)-b)\right| \le 2\|f-b\|_1 \ .\end{aligned}$$

□

Proof of Lemma 2.5. Since sets of the form $A \times B$, where $A, B \subseteq \Omega$ are measurable sets, generate the product sigma-algebra on Ω^2, there exists a finite partition $\mathcal{P}$ of Ω and a function $S : \Omega^2 \to \mathbb{R}$ such that S is constant on each rectangle of $\mathcal{P} \times \mathcal{P}$, and such that $\|\Gamma - S\|_1 < \frac{\varepsilon}{2}$. Now, for each rectangle $(A, B) \in \mathcal{P} \times \mathcal{P}$, we apply Fact 2.6 on the restricted function $\Gamma_{\restriction A\times B}$ and the constant $S_{\restriction A\times B}$. Summing up the contributions coming from these applications of Fact 2.6, we get that $\left\|\Gamma - \Gamma^{\Join\mathcal{P}}\right\|_1 \le 2\|\Gamma - S\|_1 < \varepsilon$. □

We call $\Gamma^{\Join\mathcal{P}}$ with properties as in Lemma 2.5 *averaged L^1-approximation of Γ by a step-graphon for precision ε.*

2.5. **Hyperspace $K(\mathcal{W}_0)$.** Let X be a metrizable compact space. We denote as $K(X)$ the space of all compact subsets of X with the topology generated by sets of the form $\{L \in K(X) : L \subseteq U\}$ and $\{L \in K(X) : L \cap U \neq \emptyset\}$ where $U \subseteq X$ ranges over all open sets of X. Then $K(X)$ is called the *hyperspace of X with the Vietoris topology.*

Fact 2.7 ((4.22) and (4.26) in [13])**.** *Let X be a metrizable compact space with compatible metric ρ. Then $K(X)$ is metrizable compact (and hence separable). Furthermore, the Hausdorff metric on $K(X)$,*

$$d^{\mathrm{Hf}}_\rho(L, M) = \max\left\{\max_{x\in L}\{\rho(x, M)\}, \max_{y\in M}\{\rho(y, L)\}\right\} \tag{2.3}$$

is compatible with the Vietoris topology on $K(X)$.

Remark 2.8. *We will be interested in the situation where $X = \mathcal{W}_0$ is endowed with the weak* topology. By the discussion in Section 2.3, X is indeed metrizable compact.*

3. Weak* convergence and the cut distance

3.1. **Becoming familiar with $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.** Let us observe some basic properties of the sets $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. We have $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) \subseteq \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. The set $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ can be empty (for example when $\Gamma_1 \equiv 0, \Gamma_2 \equiv 1, \Gamma_3 \equiv 0, \Gamma_4 \equiv 1, \ldots$) but $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is non-empty by the Banach–Alaoglu Theorem. Actually, we can describe some elements of $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ fairly easily. Let $T \subseteq [0, 1]$ be the set of the accumulation points of the edge densities of the graphons $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$, i.e., T is the set of the accumulation points of the sequence $\left(\int_x \int_y \Gamma_n(x, y)\right)_n$. Now, a constant $c \in [0, 1]$ (viewed as a constant graphon) lies in $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ if and only if $c \in T$. The direction that $c \in \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ implies $c \in T$ is obvious. Now, suppose that $c \in T$. That is, for some subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ the densities converge to c.

Partition each Γ_{n_i} into i sets of measure $\frac{1}{i}$ and consider a version $\widehat{\Gamma_{n_i}}$ of Γ_{n_i} obtained by a measure preserving bijection permuting these sets randomly. Then almost surely, $\widehat{\Gamma_{n_1}}, \widehat{\Gamma_{n_2}}, \widehat{\Gamma_{n_3}}, \ldots$ weak* converge to c. This is included here just to get familiar with $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and a proof is not needed at this point. However, the statement follows from Lemma 4.2(b).

The first non-trivial fact we will prove about the set $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is that it is closed.

Lemma 3.1. *Let $\Gamma_1, \Gamma_2, \Gamma_3, \ldots \in \mathcal{W}_0$ be a sequence of graphons. Then the following holds for the set $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.*

- *(a)* $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *is weak* closed in* $L^\infty\left(\Omega^2\right)$.
- *(b)* $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *is weak* compact in* $L^\infty\left(\Omega^2\right)$.
- *(c)* $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *is closed in* $L^1\left(\Omega^2\right)$.

Proof of Part (a). Suppose that $L_1, L_2, L_3, \ldots$ are elements of $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ such that $L_k \xrightarrow{\mathrm{w}^*} L$ for $k \to \infty$. For every k let $\Gamma_1^k, \Gamma_2^k, \Gamma_3^k, \ldots$ be a sequence of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ converging to L_k. We find an increasing sequence $i_1, i_2, i_3, \ldots$ of integers such that for every k and for every $n \geq i_k$ we have $d_{\mathrm{w}^*}(\Gamma_n^{k+1}, L_{k+1}) < \frac{1}{k}$. Then the following sequence of versions of $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ weak* converges to L:

$$\Gamma_1^1, \Gamma_2^1, \ldots, \Gamma_{i_1-1}^1, \Gamma_{i_1}^2, \Gamma_{i_1+1}^2, \ldots, \Gamma_{i_2-1}^2, \Gamma_{i_2}^3, \Gamma_{i_2+1}^3, \ldots, \Gamma_{i_3-1}^3, \ldots.$$

Proof of Part (b): Recall that the closed unit ball is compact in the weak* topology. Since $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ lies in this ball, it is weak* compact.

Proof of Part (c): The unit ball B of $L^\infty\left(\Omega^2\right)$ is closed in $L^1\left(\Omega^2\right)$. $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is a weak* closed subset of B, and so it is also closed in B in the topology inherited from $L^1\left(\Omega^2\right)$ (by Fact 2.3). So, $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is closed in $L^1\left(\Omega^2\right)$. □

Remark 3.2. *In* [8]*, an analogous closeness property of* $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ *than Lemma 3.1*(c) *was established and used, namely that the set* $\left\{\mathrm{INT}_f(W) : W \in \mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)\right\}$ *attains it supremum (here, f is a fixed continuous strictly convex function). Section 7.4 of* [8] *contains an example, due to Jon Noel, which shows that the set* $\left\{\mathrm{INT}_f(W) : W \in \mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)\right\}$ *need not even achieve its supremum.*

3.2. **Differentiating between $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ in [8] and in the present paper.** In the proof of Theorem 1.1 given in [8], which is in some sense a precursor of the current work, quite some work is put into zigzagging between $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Let us explain this in more detail. Let us fix a continuous strictly convex function f. The idea for finding the graphon Γ in Theorem 1.1 in [8] is as follows. Denoting by X either *(i)* $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ or *(ii)* $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, we take $\Gamma \in X$ that maximizes $\mathrm{INT}_f(\Gamma)$. Using the definition of X, there exist versions $\Gamma'_{n_1}, \Gamma'_{n_2}, \Gamma'_{n_3}, \ldots$ of $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ that converge to Γ weak*.[e] The aim is to prove that $\Gamma'_{n_1}, \Gamma'_{n_2}, \Gamma'_{n_3}, \ldots$ actually converge to Γ also in the cut norm — that would obviously prove Theorem 1.1. Now, the key step in [8] is to prove that if $\Gamma'_{n_1}, \Gamma'_{n_2}, \Gamma'_{n_3}, \ldots$ do not converge to Γ in the cut norm, then there exist versions $\Gamma''_{n_{k_1}}, \Gamma''_{n_{k_2}}, \Gamma''_{n_{k_3}}, \ldots$ of a suitable subsequence of $\Gamma'_{n_1}, \Gamma'_{n_2}, \Gamma'_{n_3}, \ldots$ that weak* converge to a graphon Γ' with $\mathrm{INT}_f(\Gamma') > \mathrm{INT}_f(\Gamma)$. Since $\Gamma''_{n_{k_1}}, \Gamma''_{n_{k_2}}, \Gamma''_{n_{k_3}}, \ldots$ witness that $\Gamma' \in X$, this is a contradiction. Now, let us explain why we need favorable properties of

[e]Note that in variant (i), we actually have $n_1 = 1, n_2 = 2, n_3 = 3, \ldots$.

both (i) and (ii) for the proof. Firstly, note that in the sentence «Since $\Gamma''_{n_{k_1}}, \Gamma''_{n_{k_2}}, \Gamma''_{n_{k_3}}, \ldots$ witness that $\Gamma' \in X$» we are referring to a subsequence, so this is a correct justification only in case $X = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. On the other hand, in the sentence «we take $\Gamma \in X$ that maximizes $\mathrm{INT}_f(\Gamma)$» we need the maximum to be achieved. Such a closeness property is enjoyed by $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ as we saw in Lemma 3.1(c), but not by $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ as we saw in Remark 3.2.

So, while differences between $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ were viewed in [8] as a nuisance that required a subtle and technical treatment, in this section we shall show that these differences capture the essence of the cut norm convergence. Namely, we shall prove in Theorem 3.3 that each sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ of graphons contains a subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ such that

$$\mathbf{LIM}_{w*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right) = \mathbf{ACC}_{w*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right) \;, \tag{3.1}$$

and in Theorem 3.5 we shall prove that (3.1) is equivalent to cut distance convergence of $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$. Of course, a proof of Theorem 1.1 then follows immediately.

3.3. **Main results: subsequences with $\mathbf{LIM}_{w*} = \mathbf{ACC}_{w*}$.** As we observed earlier we have $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) \subseteq \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and equality usually does not hold. The next theorem however says that we can always achieve equality after passing to a subsequence.

Theorem 3.3. *Let $\mathcal{S} = (\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ be a sequence of graphons. Then there exists a subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ such that $\mathbf{LIM}_{w*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right) = \mathbf{ACC}_{w*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right)$.*

The proof of Theorem 3.3 proceeds by transfinite induction. Crucially, we rely on a well-known fact from descriptive set theory, below referred to [13], which says that a strictly increasing transfinite sequence of closed sets in a second countable topological space must be of at most countable length. We shall apply this to the space $\left(L^\infty(\Omega^2), w^*\right)$ which is second countable because it is metrizable and separable.

Proof. For two sequences[f] of graphons $\mathcal{U}$ and $\mathcal{T}$ we write $\mathcal{U} \leq^\star \mathcal{T}$ if deleting finitely many terms from $\mathcal{U}$ gives us a subsequence of $\mathcal{T}$. Note that the relation $\leq^\star$ is transitive. Note that if $\mathcal{U} \leq^\star \mathcal{T}$ then

$$\mathbf{LIM}_{w*}\left(\mathcal{T}\right) \subseteq \mathbf{LIM}_{w*}\left(\mathcal{U}\right) \;. \tag{3.2}$$

In the following, we construct a countable ordinal α_0 and a transfinite sequence $\left(\mathcal{S}_\alpha\right)_{\alpha \leq \alpha_0}$ of subsequences of $\mathcal{S}$ such that for every pair of ordinals $\gamma < \delta$ it holds that

$$\mathcal{S}_\delta \leq^\star \mathcal{S}_\gamma \;, \tag{3.3}$$

and also that $\mathbf{LIM}_{w*}\left(\mathcal{S}_\gamma\right)$ is a proper subset of $\mathbf{LIM}_{w*}\left(\mathcal{S}_\delta\right)$.

In the first step, we put $\mathcal{S}_0 = \mathcal{S}$. Now suppose that for some countable ordinal α, we have already constructed $\mathcal{S}_\beta$ for every $\beta < \alpha$. Either $\alpha = \beta + 1$ for some ordinal β or α is a limit ordinal. Suppose first that $\alpha = \beta + 1$ for some ordinal β. We distinguish two cases. If $\mathbf{LIM}_{w*}(\mathcal{S}_\beta) = \mathbf{ACC}_{w*}(\mathcal{S}_\beta)$ then we define $\alpha_0 = \beta$ and the construction is finished. Otherwise there is some graphon $W \in \mathbf{ACC}_{w*}(\mathcal{S}_\beta) \setminus \mathbf{LIM}_{w*}(\mathcal{S}_\beta)$. Then we proceed the construction by finding a subsequence $\mathcal{S}_{\beta+1}$ of $\mathcal{S}_\beta$ such that some versions of the graphons from $S_{\beta+1}$ converge to W. This way we have $\mathcal{S}_{\beta+1} \leq^\star \mathcal{S}_\beta$ and $W \in \mathbf{LIM}_{w*}(\mathcal{S}_{\beta+1}) \setminus \mathbf{LIM}_{w*}(\mathcal{S}_\beta)$. Now suppose

[f]By a *sequence*, we mean a system indexed by a countable initial segment of ordinals.

that α is a countable limit ordinal. We find an increasing sequence $\beta_1, \beta_2, \beta_3, \ldots$ of ordinals such that $\beta_i \to \alpha$ for $i \to \infty$ (this is possible as α has countable cofinality). Now we use the diagonal method to define a sequence S_α such that $S_\alpha \leq^\star S_{\beta_i}$ for every i. Combined with (3.3) and with $\beta_i \to \alpha$, we get that $\mathcal{S}_\alpha \leq^\star \mathcal{S}_\beta$ for every $\beta < \alpha$. Plugging in (3.2), we conclude $\bigcup_{\beta<\alpha} \mathbf{LIM}_{\mathrm{w}*}(\mathcal{S}_\beta) \subseteq \mathbf{LIM}_{\mathrm{w}*}(\mathcal{S}_\alpha)$.

The obtained transfinite sequence $(\mathbf{LIM}_{\mathrm{w}*}(\mathcal{S}_\alpha))_{\alpha \leq \alpha_0}$ is a strictly increasing sequence of subsets of unit ball in $L^\infty(\Omega^2)$. By Lemma 3.1(a), all of these subsets are weak* closed. By [13, Theorem 6.9], the sequence is at most countable, i.e. the previous construction stopped at some countable ordinal α_0. This means that $\mathbf{LIM}_{\mathrm{w}*}(\mathcal{S}_{\alpha_0}) = \mathbf{ACC}_{\mathrm{w}*}(\mathcal{S}_{\alpha_0})$. □

Remark 3.4. *Theorem 3.3 substantially extends the key Lemma 13 from* [8] *which states that any sequence of graphons $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ contains a subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ such that*

$$\sup\left\{\mathrm{INT}_f(\Gamma) \ : \ \Gamma \in \mathbf{LIM}_{\mathrm{w}*}(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots)\right\} = \sup\left\{\mathrm{INT}_f(\Gamma) \ : \ \Gamma \in \mathbf{ACC}_{\mathrm{w}*}(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots)\right\},$$

for a continuous strictly convex function $f : [0,1] \to \mathbb{R}$. Lemma 13 in [8] *is proved by induction (over natural numbers) without any appeal to descriptive set theory.*

As promised, we shall now state that the property asserted in Theorem 3.3 is necessary and sufficient for cut distance convergence.

Theorem 3.5. *Let $\Gamma_1, \Gamma_2, \Gamma_3, \ldots \in \mathcal{W}_0$. The following are equivalent:*

(a) The sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ is Cauchy with respect to the cut distance $\delta_\square$,
(b) $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) = \mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.

Furthermore, in case (a) and (b) hold, we can take a maximal element W in $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ with respect to the structuredness order (defined in Section 4 below) and then $\Gamma_1, \Gamma_2, \Gamma_3, \ldots \xrightarrow{\delta_\square} W$.

We provide a proof of Theorem 3.5 in Section 5, after building key tools in Section 4. In Section 6 we state and prove Theorem 6.1 which extends Theorem 3.5 and relates cut distance convergence to convergence in the hyperspace $K(\mathcal{W}_0)$.

4. Envelopes and the structuredness order

Suppose that $W \in \mathcal{W}_0$ is a graphon. We call the set $\langle W \rangle := \mathbf{LIM}_{\mathrm{w}*}(W, W, W, \ldots)$ the *envelope* of W. Envelopes allow us to introduce structuredness order on graphons. Intuitively, less-structured graphons have smaller envelopes. Extreme examples of this are constant graphons $W \equiv c$ (for some $c \in [0,1]$), which are obviously the only graphons for which $\langle W \rangle = \{W\}$. This leads us to say that a graphon U is *at most as structured as a graphon W* if $\langle U \rangle \subseteq \langle W \rangle$. We write $U \preceq W$ in this case. We write $U \prec W$ if $U \preceq W$ but it does not hold that $W \preceq U$. Observe that $\preceq$ is a quasiorder on the space of graphons and if $U \preceq W$ then also $U^\varphi \preceq W$ for every measure preserving bijection φ. As we shall see in Lemma 8.1, it is actually an order on the space of graphons modulo weak isomorphism. To prove these results we shall need several auxiliary results.

Lemma 4.1 (Lemma 7 in [8])**.** *Suppose that $\Gamma_1, \Gamma_2, \Gamma_3, \ldots : \Omega^2 \to [0,1]$ 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 $\mathcal{P}$ of Ω into finitely many sets. Then $W^{\bowtie\mathcal{P}} \in \mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.*

Lemma 4.2. *Suppose that $W \in \mathcal{W}_0$. Then*

(a) If $Q \subseteq \langle W \rangle$ then the weak* closure of Q is also contained in $\langle W \rangle$,
(b) $W^{\ltimes \mathcal{P}} \in \langle W \rangle$ for every finite partition $\mathcal{P}$ of Ω,
(c) $U \in \langle W \rangle$ if and only if $U \preceq W$,
(d) if $\delta_\square(W, U) = 0$ then $\langle W \rangle = \langle U \rangle$.

Proof. Item (a) follows from Lemma 3.1(a). Item (b) is a special case of Lemma 4.1.

Let us now turn to Item (c). If $U \preceq W$ then $U \in \langle W \rangle$ follows from the definition of $\preceq$ and the fact that $U \in \langle U \rangle$. To prove the opposite implication observe that if $(\varphi_n : \Omega \to \Omega)_n$ is a sequence of measure preserving bijections, $W^{\varphi_n} \xrightarrow{w^*} U$ and $\psi : \Omega \to \Omega$ is a measure preserving bijection then $W^{\varphi_n \psi} \xrightarrow{w^*} U^{\psi}$. Then we have that every version of U is in $\langle W \rangle$. Because $\langle U \rangle$ is exactly the weak* closure of the set of all versions of U we obtain that $U \preceq W$.

Let us now prove Item (d). If $\delta_\square(W, U) = 0$ then we have a sequence $W^{\varphi_n} \xrightarrow{\|\cdot\|_\square} U$ which by Fact 2.3 implies that $W^{\varphi_n} \xrightarrow{w^*} U$. Therefore $U \preceq W$. A symmetric argument gives $W \preceq U$ and we may conclude that $\langle W \rangle = \langle U \rangle$. $\square$

4.1. **Beyond Lemma 4.2(b).** In this short section we borrow results stated and proven further below to show, in Theorem 4.3, that Lemma 4.2(b) can be strengthened. We do not have applications of Theorem 4.3. On the other hand, together with the complement of Theorem 4.3, which we include as Problem 4.4, this would lead to a correspondence between the structuredness order and order on sub-sigma-algebras of the Borel sigma-algebra. Let us now give details.

Let $\mathcal{B}$ be the sigma-algebra of measurable sets on Ω. Suppose that $W : \Omega^2 \to [0,1]$ and $\mathcal{P}$ is a finite partition of Ω into measurable sets. Let $\mathcal{P}^*$ be the algebra generated by $\mathcal{P}$; clearly $\mathcal{P}^*$ is finite. We can express $W^{\ltimes \mathcal{P}}$ as a conditional expectation of W with respect to a sub-sigma-algebra, $W^{\ltimes P} = \mathbf{E}[W | \mathcal{P}^* \times \mathcal{P}^*]$. Then Lemma 4.2(b) tells us that this particular conditional expectation lies in $\langle W \rangle$. Here, the fact that we are taking a conditional expectation with respect to a finite sub-sigma-algebra is not needed.

Theorem 4.3. *Suppose $\mathcal{B}$ is the sigma-algebra of measurable sets on Ω, $\mathcal{A}$ is a sub-sigma-algebra of $\mathcal{B}$, and that $W : \Omega^2 \to [0,1]$ is a graphon. Then $\mathbf{E}[W | \mathcal{A} \times \mathcal{A}] \in \langle W \rangle$.*

Proof. Recall that we assume that Ω is a separable atomless measure space. This implies that there is an increasing sequence $\{\mathcal{A}_n\}_{n \in \mathbb{N}}$ of finite subalgebras of $\mathcal{A}$ such that $\bigcup_{n \in \mathbb{N}} \mathcal{A}_n$ generates $\mathcal{A}$, i.e., $\mathcal{A}$ is the smallest sub-sigma-algebra of $\mathcal{B}$ that contains $\mathcal{A}_n$ for every $n \in \mathbb{N}$. Clearly, there are finite partitions $\mathcal{P}_n$ of Ω for every $n \in \mathbb{N}$ such that $\mathcal{P}_n^* = \mathcal{A}_n$, where we use the notation from the beginning of this subsection. It follows that $W^{\ltimes P_n} \in \left\langle W^{\ltimes P_{n+1}} \right\rangle$ and by Corollary 8.3 together with Lemma 4.5 we have that $W^{\ltimes P_n} \xrightarrow{\delta_\square} U$ where $U \in \langle W \rangle$. It follows from the Martingale convergence theorem [2, Theorem 35.5] that $W^{\ltimes P_n} \xrightarrow{\|\cdot\|_1} \mathbf{E}[W | \mathcal{A} \times \mathcal{A}]$ and as a consequence we get that $\delta_\square(\mathbf{E}[W | \mathcal{A} \times \mathcal{A}], U) = 0$. $\square$

As advertised, we pose the complement of Theorem 4.3, which would give one-to-one correspondence between the structuredness order and containment of sub-sigma-algebras of $\mathcal{B}$, as an open problem.

Problem 4.4. Suppose $\mathcal{B}$ is the sigma-algebra of measurable sets on Ω, and that $U \preceq W$ are two graphons. Do there exist graphons U', W', and a sub-sigma-algebra $\mathcal{A}$ of $\mathcal{B}$ such that $\delta_\square(U, U') = \delta_\square(W, W') = 0$ and $U' = \mathbf{E}[W' | \mathcal{A} \times \mathcal{A}]$?

4.2. **$\preceq$-maximal elements in $\mathbf{LIM}_{w*}$.** The main result of this section, Lemma 4.9, says that if

$$\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$$

then there exists a $\preceq$-maximal element in $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Most of the work for the proof of Lemma 4.9 is done in Lemma 4.8 which is stated for step graphons only. To infer that certain favorable properties of a sequence of step graphons (on which Lemma 4.8 can be applied) can be transferred even to a graphon they approximate, Lemma 4.5 is introduced.

Lemma 4.5. *Suppose $U_1, U_2, U_3, \ldots$ is a sequence of graphons that converges weak* to U, and suppose that W is a graphon. Suppose that for each $n \in \mathbb{N}$ we have that $U_n \preceq W$. Then $U \preceq W$.*

Proof. Follows immediately from Lemma 4.2(a). □

For the key Lemma 4.8, we shall refine the structure of a graphon by «moving some parts to the left». To this end, it is convenient to work on $[0,1]$ (see Remark 2.1). We introduce the following definitions.

Definition 4.6. By an *ordered partition* $\mathcal{P}$ of a set S, we mean a finite partition of S, $S = P_1 \sqcup P_2 \sqcup P_3 \sqcup \ldots \sqcup P_k$, $\mathcal{P} = (P_1, P_2, P_3, \ldots, P_k)$ in which the sets $P_1, P_2, P_3, \ldots, P_k$ are linearly ordered (in the way they are enumerated in $\mathcal{P}$).

Definition 4.7. For an ordered partition $\mathcal{J}$ of $I = [0,1]$ into finitely many sets $C_1, C_2, \ldots, C_k$, we define mappings $\alpha_{\mathcal{J},1}, \alpha_{\mathcal{J},2}, \ldots, \alpha_{\mathcal{J},k} : I \to I$, and a mapping $\gamma_{\mathcal{J}} : I \to I$ by

$$\begin{aligned} \alpha_{\mathcal{J},i}(x) &= \lambda\left(\bigcup_{i=1}^{j-1} C_j\right) + \lambda\left(C_i \cap [0,x]\right), \quad i = 1,2,\ldots,k\,. \\ \gamma_{\mathcal{J}}(x) &= \alpha_{\mathcal{J},i}(x) \quad \text{if } x \in C_i, \quad i = 1,2,\ldots,k\,. \end{aligned} \tag{4.1}$$

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.

Last, given a graphon $W : I^2 \to [0,1]$, we define a graphon ${}_{\mathcal{J}}W : I^2 \to [0,1]$ by

$$ {}_{\mathcal{J}}W(x,y) := W\left(\gamma_{\mathcal{J}}^{-1}(x), \gamma_{\mathcal{J}}^{-1}(y)\right)\,. \tag{4.2}$$

When $\mathcal{J}$ has only two parts, $\mathcal{J} = (A, I \setminus A)$, in order to simplify notation, we write

$$ {}_{A}W := {}_{\mathcal{J}}W\,. \tag{4.3}$$

Lemma 4.8. *Let $\Gamma_1, \Gamma_2, \Gamma_3, \ldots \in \mathcal{W}_0$ be a sequence of graphons on $[0,1]$. Suppose that $U, V \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ are two step graphons. Then there exists a step graphon $W \in \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \ldots)$ that refines U and such that $U, V \preceq W$.*

Proof. We at first assume that the ordered partition $\mathcal{P} = (P_1, \ldots, P_m)$ of U and the ordered partition $\mathcal{Q} = (Q_1, \ldots, Q_n)$ of V are composed of intervals, since each partition $\mathcal{J}$ can be reordered to intervals by the measure preserving almost-bijection $\gamma_{\mathcal{J}}$. We further assume without loss of generality that the sequence of measure preserving bijections certifying that $U \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \ldots)$ contains only identities, i.e., that $\Gamma_1, \Gamma_2, \ldots \xrightarrow{w^*} U$. Let $\varphi_1, \varphi_2, \ldots$ be measure preserving bijections such that $\Gamma_1^{\varphi_1^{-1}}, \Gamma_2^{\varphi_2^{-1}}, \ldots \xrightarrow{w^*} V$.

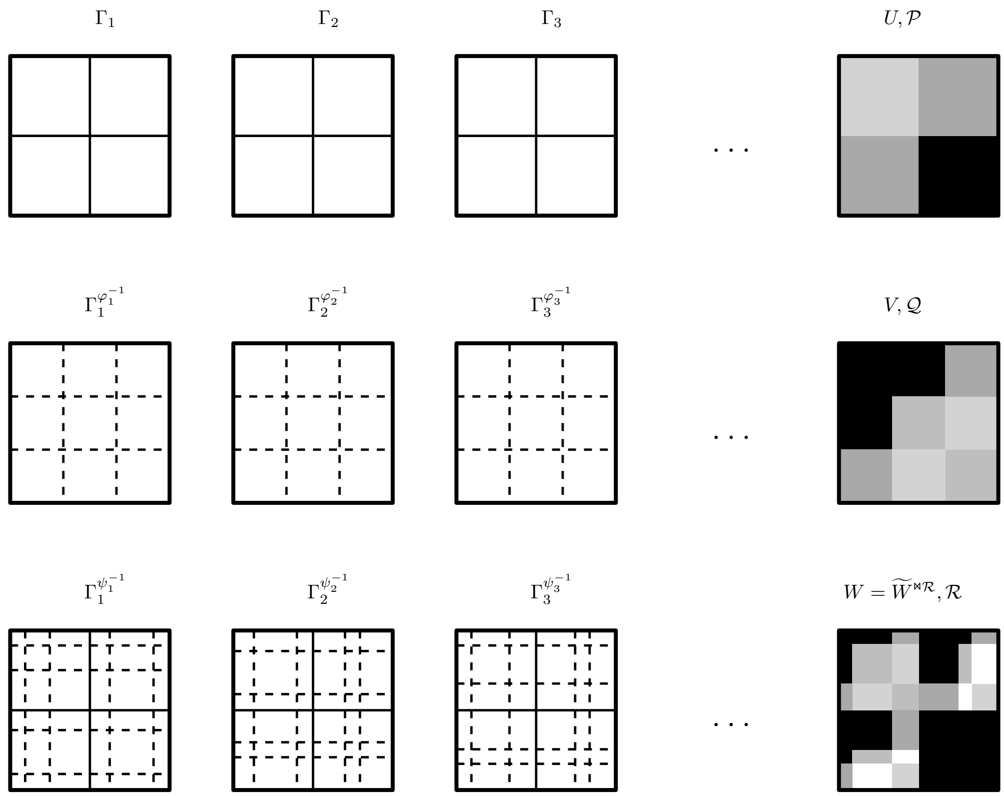

FIGURE 4.1. Two graphons U, V are step graphons with partitions $\mathcal{P}$ and $\mathcal{Q}$. A subsequence of $\Gamma_1^{\psi_1^{-1}}, \Gamma_2^{\psi_2^{-1}}, \ldots$ converges to $\widetilde{W}$ and the corresponding partitions converge to the partition $\mathcal{R}$ that refines both $\mathcal{P}$ and $\mathcal{Q}$. The graphon $W = \widetilde{W}^{\ltimes\mathcal{R}}$ is the desired step graphon that is structured more than both U and V.

We now describe a sequence of measure preserving bijections $\psi_1, \psi_2, \ldots$ such that the weak star limit of $\Gamma_1^{\psi_1^{-1}}, \Gamma_2^{\psi_2^{-1}} \ldots$ gives the desired graphon W. For the ℓ-th graphon Γ_ℓ we define its partition $\mathcal{H}^{(\ell)} = \left(H_{1,1}^{(\ell)}, H_{1,2}^{(\ell)}, \ldots, H_{1,n}^{(\ell)}, H_{2,1}^{(\ell)} \ldots, H_{m,n}^{(\ell)}\right)$ where $H_{i,j}^{(\ell)} = P_i \cap \varphi_\ell^{-1}(Q_j)$ and set $\psi_\ell = \gamma_{\mathcal{H}^{(\ell)}}$ using Definition 4.7. The intuition behind ψ_ℓ is that it refines each block P_i of the partition $\mathcal{P}$ with the partition $\mathcal{Q}$; it can be, indeed, seen that for each i we have $\psi_\ell(P_i) = \psi_\ell\left(\bigcup_j H_{i,j}^{(\ell)}\right) = P_i$, where the equalities hold up to a null set.

We pass to a subsequence mn times to get that both endpoints of each of the intervals $\psi_\ell(H_{i,j}^{(\ell)})$ converge to some fixed numbers from $[0,1]$, thus giving us a limit partition $\mathcal{R} = (R_{1,1}, \ldots, R_{m,n})$ into intervals (some of them may be degenerate intervals of length 0). Note that as we know that for each i we have $\psi_\ell\left(\bigcup_j H_{i,j}^{(\ell)}\right) = P_i$, it is also true that $\bigcup_j R_{i,j} = P_i$. We also have for all j that $\nu\left(\bigcup_i H_{i,j}^{(\ell)}\right) = \nu(Q_j)$, hence $\nu\left(\bigcup_i R_{i,j}\right) = \nu(Q_j)$. Now we use the fact that the set of accumulation points of our sequence is non-empty due to Banach–Alaoglu theorem, thus after passing to a subsequence yet again we get a subsequence $\Gamma_{k_1}^{\psi_{k_1}^{-1}}, \Gamma_{k_2}^{\psi_{k_2}^{-1}}, \ldots \xrightarrow{w^*} \widetilde{W}$.

Define W as $\widetilde{W}^{\ltimes\mathcal{R}}$. We apply Lemma 3.1(a) to $\mathbf{LIM}_{w*}\left(\Gamma_{k_1},\Gamma_{k_2},\dots\right)$ and Lemma 4.2(b) to $\widetilde{W}$ and $\widetilde{W}^{\ltimes\mathcal{R}}$ to get that $\widetilde{W}^{\ltimes\mathcal{R}}\in\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\dots)$.

At first we prove that $\widetilde{W}^{\ltimes\mathcal{R}}$ refines U. Since U is constant on each step $P_i\times P_j$, it suffices to prove that for any $\varepsilon>0$ and any step $P_i\times P_j$ we have

$$\int_{P_i\times P_j} U=\int_{P_i\times P_j}\widetilde{W}^{\ltimes\mathcal{R}}. \tag{4.4}$$

Take ℓ sufficiently large, so that

$$\left|\int_{P_i\times P_j} U-\int_{P_i\times P_j}\Gamma_{k_\ell}\right|<\varepsilon \tag{4.5}$$

and

$$\left|\int_{P_i\times P_j}\widetilde{W}-\int_{P_i\times P_j}\Gamma_{k_\ell}^{\psi_{k_\ell}^{-1}}\right|<\varepsilon. \tag{4.6}$$

Putting this together with the facts that $P_i=\psi_{k_\ell}(P_i)$ up to a null set and $R_{i,j}\subseteq P_i$ for all i,j, we get that

$$\begin{aligned}
&\int_{P_i\times P_j} U\\
\boxed{\text{Eq. (4.5)}}\quad &\overset{\varepsilon}{\approx}\int_{P_i\times P_j}\Gamma_{k_\ell}\\
\boxed{P_i=\psi_{k_\ell}(P_i)}\quad &=\int_{\psi_{k_\ell}(P_i)\times\psi_{k_\ell}(P_j)}\Gamma_{k_\ell}\\
&=\int_{P_i\times P_j}\Gamma_{k_\ell}^{\psi_{k_\ell}^{-1}}\\
\boxed{\text{Eq. (4.6)}}\quad &\overset{\varepsilon}{\approx}\int_{P_i\times P_j}\widetilde{W}\\
&=\int_{P_i\times P_j}\widetilde{W}^{\ltimes\mathcal{R}}.
\end{aligned}$$

Since this holds for every $\varepsilon>0$, we get the desired Equation (4.4).

Now we prove that $\widetilde{W}^{\ltimes\mathcal{R}}\succeq V$. Since V is constant on each step $Q_i\times Q_j$, it suffices to prove that for any $\varepsilon>0$ and any step $Q_i\times Q_j$ of V we have

$$\int_{Q_i\times Q_j} V=\sum_{\substack{1\le g\le m\\1\le h\le n}}\int_{R_{g,i}\times R_{h,j}}\widetilde{W}^{\ltimes\mathcal{R}}. \tag{4.7}$$

Take ℓ sufficiently large so that

$$\left|\int_{Q_i\times Q_j} V-\int_{Q_i\times Q_j}\Gamma_{k_\ell}^{\varphi_{k_\ell}^{-1}}\right|<\varepsilon, \tag{4.8}$$

$$\left|\sum_{g,h}\int_{R_{g,i}\times R_{h,j}}\widetilde{W}^{\ltimes\mathcal{R}}-\sum_{g,h}\int_{R_{g,i}\times R_{h,j}}\Gamma_{k_\ell}^{\psi_{k_\ell}^{-1}}\right|<\varepsilon \tag{4.9}$$

and, moreover, the length of each interval $\psi_{k_\ell}\left(H_{i,j}^{(k_\ell)}\right)$ differs from the length of interval $R_{i,j}$ by at most $\frac{\varepsilon}{4m^2n^2}$. Now we can bound the measure of overlap of each pair of rectangles $R_{g,i}\times R_{h,j}$ and $\psi_{k_\ell}\left(H_{g,i}^{(k_\ell)}\right)\times\psi_{k_\ell}\left(H_{h,j}^{(k_\ell)}\right)$. More precisely, we have

$$\nu^{\otimes 2}\left(\left(R_{g,i}\times R_{h,j}\right)\ \triangle\ \left(\psi_{k_\ell}\left(H_{g,i}^{(k_\ell)}\right)\times\psi_{k_\ell}\left(H_{h,j}^{(k_\ell)}\right)\right)\right) < 4mn\cdot\frac{\varepsilon}{4m^2n^2}=\frac{\varepsilon}{mn}, \tag{4.10}$$

where $4mn$ comes from the facts that we bound the displacement of all four sides of the rectangles and that their displacement depends on the displacement of all preceding intervals. Putting all of this together, we get that

$$\begin{aligned}
&\sum_{g,h}\int_{R_{g,i}\times R_{h,j}}\widetilde{W}^{\ltimes\mathcal{R}}\\
\boxed{\text{Eq. (4.9)}}\quad &\overset{\varepsilon}{\approx}\sum_{g,h}\int_{R_{g,i}\times R_{h,j}}\Gamma_{k_\ell}^{\psi_{k_\ell}^{-1}}\\
\boxed{\text{Eq. (4.10)}}\quad &\overset{\varepsilon}{\approx}\sum_{g,h}\int_{\psi_{k_\ell}\left(H_{g,i}^{(k_\ell)}\right)\times\psi_{k_\ell}\left(H_{h,j}^{(k_\ell)}\right)}\Gamma_{k_\ell}^{\psi_{k_\ell}^{-1}}\\
&=\sum_{g,h}\int_{\left(P_g\cap\varphi_{k_\ell}^{-1}(Q_i)\right)\times\left(P_h\cap\varphi_{k_\ell}^{-1}(Q_j)\right)}\Gamma_{k_\ell}\\
&=\sum_{g,h}\int_{\left(\varphi_{k_\ell}(P_g)\cap Q_i\right)\times\left(\varphi_{k_\ell}(P_h)\cap Q_j\right)}\Gamma_{k_\ell}^{\varphi_{k_\ell}^{-1}}\\
&=\int_{Q_i\times Q_j}\Gamma_{k_\ell}^{\varphi_{k_\ell}^{-1}}\\
\boxed{\text{Eq. (4.8)}}\quad &\overset{\varepsilon}{\approx}\int_{Q_i\times Q_j}V\ .
\end{aligned}$$

This yields the desired equation (4.7). □

Lemma 4.9. *Let $\Gamma_1,\Gamma_2,\Gamma_3,\ldots\in\mathcal{W}_0$ be a sequence of graphons on $[0,1]$ for which $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)=\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. Then $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ contains a maximum element with respect to the structuredness order.*

Proof. The space $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ is separable metrizable since the space $\mathcal{W}_0$ with the weak* topology is separable metrizable and therefore we may find a countable set $P\subseteq\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ such that its weak* closure is $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. For each $W\in P$ and $k\in\mathbb{N}$, consider a suitable graphon, denoted by $W(k)$, that is an averaged L^1-approximation of W by a step graphon for precision $\frac{1}{k}$. Such a graphon $W(k)$ exists by Lemma 2.5. Note also that if $W(k)$ is chosen as $W^{\ltimes\mathcal{P}}$ for some finite partition $\mathcal{P}$ (as in Lemma 2.5) then $W(k)\in\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ by Lemma 4.1.

Let us now consider the set $Q:=\{W(k): W\in P, k\in\mathbb{N}\}$. The set Q is countable, contained in $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ and its weak* closure is $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. Let $U_1,U_2,U_3,\ldots$ be an enumeration of the elements of Q. Let $M_1:=U_1$ and define inductively a sequence $\{M_n\}_{n\in\mathbb{N}}\subseteq\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ as follows. Suppose that we have $M_n\in\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ and use Lemma 4.8 with M_n in place of U and U_{n+1} in place of V to find $M_{n+1}\in\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. It follows from the assumption that $\mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)=\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ that, in fact, $M_{n+1} \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Note that we have $U_n \preceq M_n$ for every $n \in \mathbb{N}$. Moreover, it follows from the construction that there is a sequence of finite partitions $\{\mathcal{P}_n\}_{n\in\mathbb{N}}$ such that $M_n = M_{n+1}^{\Join \mathcal{P}_n}$ and $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$ for every $n \in \mathbb{N}$.

By the Martingale convergence theorem [2, Theorem 35.5] we find M such that $M_n \xrightarrow{\|\cdot\|_1} M$ and $M^{\Join \mathcal{P}_n} = M_n$ for every $n \in \mathbb{N}$. Since $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is weak* closed by Lemma 3.1(a) we have $M \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and $M \succeq U_n$ for every $n \in \mathbb{N}$.

Finally, we claim that M is a maximal element of $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Indeed, let $\Gamma \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ be arbitrary. Since Q is weak* dense in $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, we can find a sequence $U_{n_1}, U_{n_2}, U_{n_3}, \ldots$ weak* converging to Γ. Note that we have $U_{n_i} \preceq M$ for each $i \in \mathbb{N}$. Lemma 4.5 gives $\Gamma \preceq M$, as was needed. □

4.3. **Cut distance identifying graphon parameters.** In this section, we mention the notion of cut distance identifying graphon parameters, which we then extensively study in a follow-up paper [9]. A *graphon parameter* is any function $\theta : \mathcal{W}_0 \to \mathbb{R}$ such that $\theta(W_1) = \theta(W_2)$ for any two graphons W_1 and W_2 with $\delta_\square(W_1, W_2) = 0$. We say that a graphon parameter $\theta(\cdot)$ is a *cut distance identifying graphon parameter* (*CDIP*) if we have that $W_1 \prec W_2$ implies $\theta(W_1) < \theta(W_2)$. In particular, the key role of $\prec$-maximal element, such as in Theorem (b), can be in this setting expressed as θ-maximal elements. Note that $\mathrm{INT}_f(\cdot)$, where f is a strictly convex convex, is CDIP by [9, Section 3.3]. Hence the proof of Theorem 1.1 in [8] is a particular instance of a CDIP-version of Theorem (b), the general version then being given in Theorem 3.3 and Theorem 3.4 in [9]. As we show in [9], CDIP have more applications: they can be used for «index-pumping» in the proof of the Frieze-Kannan regularity lemma, and they tell us quite a bit about graph norms.

4.4. **Values and degrees with respect to the structuredness order.** Given a graphon $W : \Omega^2 \to [0,1]$, we can define a pushforward probability measure on $[0,1]$ by

$$\mathbf{\Phi}_W(A) := \nu^{\otimes 2}\left(W^{-1}(A)\right) , \tag{4.11}$$

for a set $A \subseteq [0,1]$. The measure $\mathbf{\Phi}_W$ gives us the distribution of the values of W, and we call it *the range frequencies of* W. Similarly, we can take the pushforward measure of the degrees,

$$\mathbf{Y}_W(A) := \nu\left(\deg_W^{-1}(A)\right) , \tag{4.12}$$

for a set $A \subseteq [0,1]$. We call $\mathbf{Y}_W$ the *degree frequencies of* W. The measures $\mathbf{\Phi}_W$ and $\mathbf{Y}_W$ do not characterize W in general but certainly give us substantial information about W. Therefore, given two graphons $U \preceq W$ it is natural to ask how $\mathbf{\Phi}_U$ compares to $\mathbf{\Phi}_W$ and how $\mathbf{Y}_U$ compares to $\mathbf{Y}_W$. To this end, we introduce the following concept.

Definition 4.10. Suppose that Λ_1 and Λ_2 are two finite measures on $[0,1]$. We say that Λ_1 is at *least as flat* as Λ_2 if there exists a finite measure Ψ on $[0,1]^2$ such that Λ_1 is the marginal of Ψ on the first coordinate, Λ_2 is the marginal of Ψ on the second coordinate, and for each measurable $D \subseteq [0,1]$ we have

$$\int_{D\times[0,1]} x \,\mathrm{d}\Psi(x,y) = \int_{D\times[0,1]} y \,\mathrm{d}\Psi(x,y) . \tag{4.13}$$

In addition, we say that Λ_1 is *strictly flatter* than Λ_2 if $\Lambda_1 \neq \Lambda_2$.

The condition that the marginals of Ψ are Λ_1 and Λ_2 is well-known in probability theory and referred to as a *coupling of* Λ_1 *and* Λ_2. However, we are not aware of this concept being combined with the condition (4.13).

Example 4.11. We should understand Λ_1 as a certain averaging of Λ_2. For example, suppose that Λ_1 has an atom $a \in [0,1]$, say $\Lambda_1(\{a\}) = m > 0$. Then taking $D = \{a\}$, (4.13) tells us that by averaging y according to the normalized (by $\frac{1}{m}$) restriction of the measure Ψ to $\{a\} \times [0,1]$, we get a. As we show in Lemma 4.12, such a property extends also to non-atoms.

Let us prove two basic lemmas. In Lemma 4.14 we give a useful characterization of strictly flatter pairs of measures. In Lemma 4.15 we prove that the flatness relation is actually an order. Even though we do not need these lemmas, we believe that the theory we develop here would not be complete without them.

We say that a finite measure Ψ on $[0,1]^2$ is *diagonal* if we have $\Psi\left(\{(x,y) \in [0,1]^2 : x \neq y\}\right) = 0$.

Let us now recall the notion of disintegration of a measure. Suppose that (X, μ) is a probability Borel measure space on a Polish space X. Let $f : X \to Y$ be a Borel map onto another Polish space Y and denote as $f^*\mu$ the push-forward measure via f. Then the Disintegration Theorem tells us that there is a system $\{F_y\}_{y \in Y}$ of probability Borel measures on X such that

(D1) $F_y\left(f^{-1}(y)\right) = 1$ for every $y \in Y$, and
(D2) $\int_X h(x)\ \mathrm{d}\mu(x) = \int_Y \left(\int_X h(x)\ \mathrm{d}F_y(x)\right)\ \mathrm{d}f^*\mu(y)$ for every Borel map $h : X \to [0,1]$.

We will use the disintegration exclusively in the situation where $X = [0,1]^2$, f is the projection on the i-th coordinate, $\mu = \Phi$ and $f^*\mu = \Lambda_i$ where Φ is a witness for the fact that Λ_1 is at least as flat as Λ_2. When we use the variable x for the first coordinate and y for the second coordinate then we obtain a disintegration $\left\{\Phi^1_x\right\}_{x \in [0,1]}$ and $\left\{\Phi^2_y\right\}_{y \in [0,1]}$. Moreover, to simplify notation we will always assume that each Φ^i_z lives on the interval $[0,1]$ instead of the corresponding (horizontal or vertical) strip. For example in case of disintegration on the second coordinate, the two conditions above then look like this:

(D1) $\Phi^2_y([0,1]) = 1$ for every $y \in [0,1]$, and
(D2) $\int_{[0,1]^2} h(x,y)\ \mathrm{d}\Phi(x,y) = \int_{[0,1]} \left(\int_{[0,1]} h(x,y)\ \mathrm{d}\Phi^2_y(x)\right)\ d\Lambda_2(y)$ for every Borel map $h : [0,1]^2 \to [0,1]$.

Lemma 4.12. *Suppose that Λ_1 is at least as flat as Λ_2, witnessed by a measure Φ. Then for Λ_1-almost every $x \in [0,1]$ we have*

$$x = \int_{[0,1]} y\,\mathrm{d}\Phi^1_x(y)\,. \tag{4.14}$$

Proof. Assume not, then we may assume without loss of generality that there is a set $D \subseteq [0,1]$ of positive Λ_1-measure such that $x > \int_{[0,1]} y\,\mathrm{d}\Phi^1_x(y)$ for each $x \in D$. Then we have

$$\int_{D\times[0,1]} x\,\mathrm{d}\Phi(x,y) = \int_D x\,\mathrm{d}\Lambda_1(x) > \int_D \left(\int_{[0,1]} y\,\mathrm{d}\Phi^1_x(y)\right)\,\mathrm{d}\Lambda_1(x) \overset{(D2)}{=} \int_{D\times[0,1]} y\,\mathrm{d}\Phi(x,y)$$

which contradicts (4.13). $\square$

Lemma 4.13. *Suppose that Λ_1 is at least as flat as Λ_2. Then we have $\Lambda_1([0,1]) = \Lambda_2([0,1])$.*

Proof. Let Ψ be a witness that Λ_1 is at least as flat as Λ_2. Then the marginal condition of Definition 4.10 tells us that

$$\Lambda_1([0,1]) = \Psi([0,1] \times [0,1]) = \Lambda_2([0,1]) .$$

$\square$

Lemma 4.14. *Suppose that Λ_1 and Λ_2 are finite measures on $[0,1]$, and that Λ_1 is at least as flat as Λ_2. Then $\Lambda_1 = \Lambda_2$ if and only if the only measure which witnesses that Λ_1 is at least as flat as Λ_2 is diagonal.*

Proof. Suppose that we have a diagonal measure Φ whose marginals are Λ_1 and Λ_2. Then for every measurable set D we have

$$\Lambda_1(D) = \Phi(D \times [0,1]) = \Phi(D \times D) = \Phi([0,1] \times D) = \Lambda_2(D) ,$$

and so $\Lambda_1 = \Lambda_2$.

On the other hand, suppose that we have a non-diagonal measure Φ whose marginals are Λ_1 and Λ_2. By Lemma 4.13 we may assume that both Λ_1 and Λ_2 are probability measures. Let us fix a strictly convex function $f : [0,1] \to [0,1]$. Let us consider the disintegration $\{\Phi^1_x\}_{x\in[0,1]}$ of Φ. Recall that by (D1), for each $x \in [0,1]$, Φ^1_x is a probability measure. In particular, Jensen's inequality gives us

$$f\left(\int y \,\mathrm{d}\Phi^1_x(y)\right) \le \int f(y)\, \mathrm{d}\Phi^1_x(y) .$$

Observe that for a positive Λ_1-measure of x's, we have that Φ^1_x is not a Dirac measure. For each such x, the inequality above is strict. Then we have

$$\begin{aligned}\int f(x)\,\mathrm{d}\Lambda_1(x) &= \int f\left(\int y\,\mathrm{d}\Phi^1_x(y)\right)\,\mathrm{d}\Lambda_1(x)\\ \boxed{\text{Jensen's inequality as above}}\quad &< \int\left(\int f(y)\,\mathrm{d}\Phi^1_x(y)\right)\,\mathrm{d}\Lambda_1(x) = \int f(y)\,\mathrm{d}\Phi(x,y) = \int f(y)\,\mathrm{d}\Lambda_2(y).\end{aligned}$$

In particular, we can conclude that $\Lambda_1 \neq \Lambda_2$. $\square$

Lemma 4.15. *Suppose that $\Lambda_{\mathrm{A}}, \Lambda_{\mathrm{B}}, \Lambda_{\mathrm{C}}$ are three finite measures on $[0,1]$. Suppose thatΛ_{A} is at least as flat as Λ_{B} and that Λ_{B} is at least as flat as Λ_{C}. Then Λ_{A} is at least as flat as Λ_{C}. If, in addition at least one of these flatness relations is strict, then Λ_{A} is strictly flatter than Λ_{C}.*

Proof. By Lemma 4.13, the measures Λ_{A}, Λ_{B} and Λ_{C} have the same total measure, say $\Lambda_{\mathrm{A}}([0,1]) = \Lambda_{\mathrm{B}}([0,1]) = \Lambda_{\mathrm{C}}([0,1]) = m$. By multiplying these measures, and all the corresponding measures witnessing the flatness relation by $\frac{1}{m}$, it is enough to restrict ourselves to the case of probability measures from now on.

Let $\widetilde{\Phi}$ be a witness that Λ_{A} is at least as flat as Λ_{B}, and let $\widehat{\Phi}$ be a witness that Λ_{B} is at least as flat as Λ_{C}. Let us disintegrate $\widetilde{\Phi}$ on the second coordinate, and $\widehat{\Phi}$ on the first coordinate. This gives us families $\left\{\widetilde{\Phi}^2_y\right\}_{y\in[0,1]}$ and $\left\{\widehat{\Phi}^1_y\right\}_{y\in[0,1]}$ of probability Borel measures on $[0,1]$. This way, for each measurable set $S \subseteq [0,1]^2$ we have

$$\widetilde{\Phi}(S) = \int_y \widetilde{\Phi}^2_y(\{x : (x,y) \in S\})\, \mathrm{d}\Lambda_{\mathrm{B}}(y) \text{ and } \widehat{\Phi}(S) = \int_y \widehat{\Phi}^1_y(\{z : (y,z) \in S\})\, \mathrm{d}\Lambda_{\mathrm{B}}(y) .$$

Now, for every set $S \subseteq [0,1]^2$ of the form $S = \bigsqcup_{i=1}^n A_i \times B_i$ where A_i and B_i are measurable subsets of $[0,1]$, we define

$$\Xi(S) := \sum_{i=1}^n \int_{y\in[0,1]} \widetilde{\Phi}^2_y\left(A_i\right) \cdot \widehat{\Phi}^1_y\left(B_i\right)\ \mathrm{d}\Lambda_{\mathrm{B}}(y)\ . \tag{4.15}$$

It is tedious but straightforward to verify that the value of $\Xi(S)$ does not depend on the choice of the decomposition $S = \bigsqcup_{i=1}^n A_i \times B_i$. Then Carathéodory's extension theorem allows us to extend Ξ to a Borel measure on $[0,1]^2$, which we still denote by Ξ. We claim that this measure witnesses that Λ_{A} is at least as flat as Λ_{C}. Firstly, let us check that the marginals of Ξ are Λ_{A} and Λ_{C}, respectively. For $D \subseteq [0,1]$, we have that

$$\begin{aligned}\Xi\left(D\times[0,1]\right) &\overset{4.15}{=} \int_{y\in[0,1]} \widetilde{\Phi}^2_y(D)\cdot\widehat{\Phi}^1_y\left([0,1]\right)\ \mathrm{d}\Lambda_{\mathrm{B}}(y)\\ \boxed{\text{by (D2), }\widehat{\Phi}^1_y([0,1])=1}\quad &= \int_{y\in[0,1]} \widetilde{\Phi}^2_y(D)\ \mathrm{d}\Lambda_{\mathrm{B}}(y) = \widetilde{\Phi}\left(D\times[0,1]\right) = \Lambda_{\mathrm{A}}\left(D\right)\ .\end{aligned}$$

Similarly, one can verify that $\Lambda_C(D) = \Xi\left([0,1]\times D\right)$.

Let $D \subseteq [0,1]$. We have the following

$$\begin{aligned}\int_{D\times[0,1]} x\,\mathrm{d}\Xi(x,z) &= \int_D x\,\mathrm{d}\Lambda_{\mathrm{A}}(x)\\ &= \int_{D\times[0,1]} x\,\mathrm{d}\tilde{\Phi}(x,y)\\ &= \int_{D\times[0,1]} y\,\mathrm{d}\tilde{\Phi}(x,y) = \int_{[0,1]}\left(\int_{[0,1]} y\cdot\mathbf{1}_{D\times[0,1]}(x,y)\,\mathrm{d}\widetilde{\Phi}^2_y(x)\right)\ \mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \int_{[0,1]} y\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_B(y)\\ &= \int_{[0,1]}\left(\int_{[0,1]} z\,\mathrm{d}\widehat{\Phi}^1_y\left(z\right)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \int_{D\times[0,1]} z\,\mathrm{d}\Xi(x,z)\end{aligned}$$

where the last equality follows from the following claim.

Claim 4.16. Let $g : [0,1] \to [0,1]$ be a measurable function. Then

$$\int_{[0,1]}\left(\int_{[0,1]} g\left(z\right)\ \mathrm{d}\widehat{\Phi}^1_y\left(z\right)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y) = \int_{D\times[0,1]} g\left(z\right)\ \mathrm{d}\Xi(x,z).$$

Proof. Let $A \subseteq [0,1]$ be a measurable set and consider its characteristic function $\mathbf{1}_A$. We have

$$\begin{aligned}\int_{[0,1]}\left(\int_{[0,1]} \mathbf{1}_A\left(z\right)\ \mathrm{d}\widehat{\Phi}^1_y\left(z\right)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y) &= \int_{[0,1]}\widehat{\Phi}^1_y\left(A\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \Xi\left(D\times A\right) = \int_{D\times[0,1]} \mathbf{1}_A\left(z\right)\ \mathrm{d}\Xi(x,z).\end{aligned}$$

This implies that the claim holds for every step function. Assume now that $g_n \to g$ uniformly and $\{g_n\}_{n\in\mathbb{N}}$ are step functions. We have

$$\begin{aligned}\int_{[0,1]}\left(\int_{[0,1]} g\,(z)\ \mathrm{d}\widehat{\Phi}^1_y\,(z)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y) &= \int_{[0,1]}\left(\int_{[0,1]} \lim_n g_n\,(z)\ \mathrm{d}\widehat{\Phi}^1_y\,(z)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \int_{[0,1]}\lim_n\left(\int_{[0,1]} g_n\,(z)\ \mathrm{d}\widehat{\Phi}^1_y\,(z)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \lim_n\int_{[0,1]}\left(\int_{[0,1]} g_n\,(z)\ \mathrm{d}\widehat{\Phi}^1_y\,(z)\right)\widetilde{\Phi}^2_y(D)\,\mathrm{d}\Lambda_{\mathrm{B}}(y)\\ &= \lim_n\int_{D\times[0,1]} g_n\,(z)\ \mathrm{d}\Xi(x,z) = \int_{D\times[0,1]} g\,(z)\ \mathrm{d}\Xi(x,z)\end{aligned}$$

(the second, third and fifth equalities follow by the uniform convergence) and that finishes the proof. □

It remains to prove the additional part. So suppose that either Λ_{A} is strictly flatter than Λ_{B} or Λ_{B} is strictly flatter than Λ_{C}. Fix a strictly convex function $f\colon [0,1]\to[0,1]$. In the very same way as in the proof of Lemma 4.14 it follows that

$$\int_{[0,1]} f(x)\,\mathrm{d}\Lambda_{\mathrm{A}}(x) \le \int_{[0,1]} f(y)\,\mathrm{d}\Lambda_{\mathrm{B}}(y) \le \int_{[0,1]} f(z)\,\mathrm{d}\Lambda_{\mathrm{C}}(z)\ ,$$

and at least one of the inequalities above is strict. Therefore $\Lambda_{\mathrm{A}} \neq \Lambda_{\mathrm{C}}$. □

The next two propositions answer the question of relating $\mathbf{\Phi}_U$ to $\mathbf{\Phi}_W$ and $\mathbf{Y}_U$ to $\mathbf{Y}_W$ using the above concept of flatter measures. While we consider these result interesting per se, let us note that in [9] we give several quick and fairly powerful applications of these results. For example, we show that the result of Doležal and Hladký can be extended to discontinuous functions, as advertised in Footnote [c].

Proposition 4.17. *Suppose that we have two graphons $U \preceq W$. Then the measure $\mathbf{\Phi}_U$ is at least as flat as the measure $\mathbf{\Phi}_W$. Similarly, the measure $\mathbf{Y}_U$ is at least as flat as the measure $\mathbf{Y}_W$. Lastly, if $U \prec W$ then $\mathbf{\Phi}_U$ is strictly flatter than $\mathbf{\Phi}_W$.*

Example 4.18. We cannot conclude that $\mathbf{Y}_U$ is strictly flatter than $\mathbf{Y}_W$ if $U \prec W$. To this end, it is enough to take U the constant-p graphon (for some $p \in (0,1)$) and W some p-regular but non-constant graphon. Then $\mathbf{Y}_U$ and $\mathbf{Y}_W$ are both equal to the Dirac measure on p.

Example 4.19. A probabilist might say that the information inherited from $\mathbf{\Phi}_W$ to $\mathbf{\Phi}_U$ is only «annealed», and not «quenched». Let us explain this on an example. Suppose that U is the constant-$\frac{1}{2}$ graphon and W attains each of the values $0, \frac{1}{2}, 1$ on sets of measure $\frac{1}{3}$ each. Obviously, we have that $U \prec W$ and thus there is a sequence $W^{\pi_1}, W^{\pi_2}, W^{\pi_3}, \ldots \xrightarrow{\mathrm{w}^*} U \equiv \frac{1}{2}$. Now, observe that there are many different scenarios where the values $\frac{1}{2}$ can arise in the limit, of which we give two extreme ones. The first possibility is that around each $(x,y) \in \Omega^2$, we have alternations of values 0's, $\frac{1}{2}$'s, and 1's in the graphons W^{π_n}, each with frequency $\frac{1}{3}$. The second possibility is that for the measure $\frac{1}{3}$ of (x,y)'s, the graphons W^{π_n} attain values only $\frac{1}{2}$ around (x,y), and for the remaining (x,y)'s of measure $\frac{2}{3}$, we have alternations of values 0's, and 1's, each with density $\frac{1}{2}$.

In the proof of Proposition 4.17, we will need some basic facts about the weak* convergence of measures on $[0,1]^2$ (this convergence is also often called weak convergence or narrow convergence in the literature) which we recall here. We say that a bounded sequence of finite positive measures $\Psi_1, \Psi_2, \Psi_3, \ldots$ on $[0,1]^2$ converges in the weak* topology to a finite positive measure Ψ if for every continuous real function f defined on $[0,1]^2$ we have

$$\lim_{n\to\infty} \int_{[0,1]^2} f(x,y)\,\mathrm{d}\Psi_n(x,y) = \int_{[0,1]^2} f(x,y)\,\mathrm{d}\Psi(x,y).$$

This definition has many equivalent reformulations but we will need only the following one: A sequence $\Psi_1, \Psi_2, \Psi_3, \ldots$ converges to Ψ in the weak* topology if and only if $\lim_{n\to\infty} \Psi_n(A) = \Psi(A)$ for every Borel subset A of $[0,1]^2$ which satisfies that the Ψ-measure of its boundary in $[0,1]^2$ is 0.[g] Recall also that every sequence of probability measures on a compact separable space has a weak* convergent subsequence.

The following lemma is the key ingredient of the proof of Lemma 4.21. However, it may be of independent interest as it connects our research with Choquet theory. (We will not use Choquet theory, and the rest of this paragraph is meant only to hint the connection.) Recall that from the point of view of Choquet theory, if Λ_1, Λ_2 are finite positive measures on some compact convex subset C of a normed space then we say that Λ_1 *is smaller than* Λ_2 if $\int_C f\,\mathrm{d}\Lambda_1 \leq \int_C f\,\mathrm{d}\Lambda_2$ for every continuous convex function $f\colon C \to \mathbb{R}$. The next lemma states that the relation «being flatter» is naturally embedded into this Choquet ordering of measures (when the compact convex set C is the unit interval $[0,1]$).

Lemma 4.20. *Let Λ_1, Λ_2 be finite measures on $[0,1]$ such that Λ_1 is at least as flat as Λ_2. Then for every continuous convex function $f\colon [0,1] \to \mathbb{R}$ it holds $\int_{[0,1]} f\,\mathrm{d}\Lambda_1 \leq \int_{[0,1]} f\,\mathrm{d}\Lambda_2$. Moreover, if Λ_1 is strictly flatter than Λ_2 then there is a continuous convex function $g\colon [0,1] \to \mathbb{R}$ such that $\int_{[0,1]} g\,\mathrm{d}\Lambda_1 < \int_{[0,1]} g\,\mathrm{d}\Lambda_2$.*

Proof. Let Ψ be a witness for Λ_1 being at least as flat as Λ_2, as in Definition 4.10. Fix a continuous convex function $f\colon [0,1] \to \mathbb{R}$ and $\varepsilon > 0$. Find a natural number n such that $f(x) - f(y) < \varepsilon$ whenever $x, y \in [0,1]$ are such that $|x-y| \leq \frac{2}{n}$. Let $[0,1] = I_1 \sqcup I_2 \sqcup \ldots \sqcup I_n$ be a partition of $[0,1]$ into pairwise disjoint intervals of lengths $\frac{1}{n}$. For every i fix a point $x_i \in I_i$. As for every i we have

$$\int_{(x,y)\in I_i\times[0,1]} x\,\mathrm{d}\Psi(x,y) = \int_{(x,y)\in I_i\times[0,1]} y\,\mathrm{d}\Psi(x,y)\,,$$

$$\left| x_i\Psi(I_i\times[0,1]) - \int_{(x,y)\in I_i\times[0,1]} x\,\mathrm{d}\Psi(x,y)\right| \leq \frac{1}{n}\cdot\Psi(I_i\times[0,1])$$

and similarly

$$\left|\sum_{j=1}^{n} y_j\Psi(I_i\times I_j) - \int_{(x,y)\in I_i\times[0,1]} y\,\mathrm{d}\Psi(x,y)\right| \leq \frac{1}{n}\cdot\Psi(I_i\times[0,1])\,,$$

it follows that

$$\left| x_i - \sum_{j=1}^{n}\frac{\Psi(I_i\times I_j)}{\Psi(I_i\times[0,1])}y_j\right| \leq \frac{2}{n}\,.$$

[g]The boundary of a set $A \subseteq [0,1]^2$ is defined as the set of all points in the closure of A which are not interior points of A. Note that the interior points are considered only from point of view of the topological space $[0,1]^2$ (not $\mathbb{R}^2$). So for example, the boundary of the closed set $A = [0,\frac{1}{2}]\times[0,1]$ is $\{\frac{1}{2}\}\times[0,1]$ as all other points from A are interior points of A in $[0,1]^2$.

So by the convexity of f we have for every i that

$$f(x_i) \overset{\varepsilon}{\approx} f\left(\sum_{j=1}^{n} \frac{\Psi(I_i \times I_j)}{\Psi(I_i \times [0,1])} y_j\right) \leq \sum_{j=1}^{n} \frac{\Psi(I_i \times I_j)}{\Psi(I_i \times [0,1])} f(y_j) \ .$$

Therefore

$$\begin{aligned}
\int_{x\in[0,1]} f(x)\,\mathrm{d}\Lambda_1(x) \qquad &= \sum_{i=1}^{n}\int_{x\in I_i} f(x)\,\mathrm{d}\Lambda_1(x) \overset{\varepsilon\cdot\Psi([0,1]^2)}{\approx} \sum_{i=1}^{n} f(x_i)\Psi(I_i\times[0,1]) \\
&\overset{\varepsilon\cdot\Psi([0,1]^2)}{\approx} \sum_{i=1}^{n} f\left(\sum_{j=1}^{n}\frac{\Psi(I_i\times I_j)}{\Psi(I_i\times[0,1])}y_j\right)\Psi(I_i\times[0,1]) \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{n} f(y_j)\Psi(I_i\times I_j) \\
&= \sum_{j=1}^{n} f(y_j)\Psi([0,1]\times I_j) \\
&\overset{\varepsilon\cdot\Psi([0,1]^2)}{\approx} \int_{y\in[0,1]} f(y)\,\mathrm{d}\Lambda_2(y) \ .
\end{aligned}$$

As this is true for every $\varepsilon > 0$ we conclude that $\int_{[0,1]} f\,\mathrm{d}\Lambda_1 \leq \int_{[0,1]} f\,\mathrm{d}\Lambda_2$.

Now suppose that Λ_1 is strictly flatter than Λ_2. Then there is a continuous function $h\colon [0,1] \to \mathbb{R}$ such that $\int_{[0,1]} h\,\mathrm{d}\Lambda_1 \neq \int_{[0,1]} h\,\mathrm{d}\Lambda_2$. Recall that functions of the form $g_1 - g_2$, where both $g_1, g_2\colon [0,1] \to \mathbb{R}$ are continuous and convex, are uniformly dense in the space of all continuous functions $h\colon [0,1] \to \mathbb{R}$. Indeed, this follows from the Stone–Weierstrass theorem as every polynomial function $p\colon [0,1] \to \mathbb{R}$ can be written as $p = g_1 - g_2$ for g_1, g_2 continuous convex (just define $g_1(x) := p(x) + Kx^2$ and $g_2(x) := Kx^2$ where K is a sufficiently large constant). So there is a continuous convex function $g\colon [0,1] \to \mathbb{R}$ such that $\int_{[0,1]} g\,\mathrm{d}\Lambda_1 \neq \int_{[0,1]} g\,\mathrm{d}\Lambda_2$. But then the previous part of the proof gives us that $\int_{[0,1]} g\,\mathrm{d}\Lambda_1 < \int_{[0,1]} g\,\mathrm{d}\Lambda_2$. □

Let us now give an intuitively clear lemma.

Lemma 4.21. *Let $\Theta^A, \Delta^A, \Theta^B, \Delta^B$ be four finite measures on $[0,1]$. Suppose that Θ^A is strictly flatter than Δ^A and that Θ^B is at least as flat as Δ^B. Then the measure $\Theta^A + \Delta^A$ is strictly flatter than $\Theta^B + \Delta^B$.*

Proof. Let Ψ^A (resp. Ψ^B) be a witness for Θ^A (resp. Θ^B) being as flat as Δ^A (resp. Δ^B), as in Definition 4.10. Then $\Psi^A + \Psi^B$ shows that $\Theta^A + \Delta^A$ is at least as flat as $\Theta^B + \Delta^B$. It remains to show that the relation is actually strict.

Let $g\colon [0,1] \to \mathbb{R}$ be a continuous convex function such that $\int_{[0,1]} g\,\mathrm{d}\Theta^A < \int_{[0,1]} g\,\mathrm{d}\Delta^A$. Such a function exists by Lemma 4.20. Then another application of Lemma 4.20 gives that $\int_{[0,1]} g\,\mathrm{d}(\Theta^A + \Delta^A) < \int_{[0,1]} g\,\mathrm{d}(\Theta^B + \Delta^B)$, and so the measures $\Theta^A + \Delta^A$ and $\Theta^B + \Delta^B$ are not the same. □

The proof of Proposition 4.17 relies on the following lemma.

Lemma 4.22. *Let B be a separable atomless finite measure space with the measure β, and let $(f_n : B \to [0,1])_n$ and $f : B \to [0,1]$ be measurable functions such that $\lim_{n\to\infty} \int_A f_n(x)\ d\beta(x) = \int_A f(x)\ d\beta(x)$ for every measurable $A \subseteq B$. Suppose that Δ_n and Δ are the pushforward measures of f_n and f,*

$\Delta_n(L) := \beta\left(f_n^{-1}(L)\right)$, $\Delta(L) := \beta\left(f^{-1}(L)\right)$. *Suppose that the measures* Δ_n *weak* converge to a probability measure* Δ^*. *Then* Δ *is at least as flat as* Δ^*.

Proof. For every natural number n and every measurable subset A of $[0,1]^2$ we define

$$\Psi_n(A) = \beta\left(\{x \in B\colon\ (f(x), f_n(x)) \in A\}\right).$$

Clearly every Ψ_n is a measure on $[0,1]^2$. Let Ψ be some weak* accumulation point of the sequence $\Psi_1, \Psi_2, \Psi_3, \ldots$. Without loss of generality, we may assume that the sequence $\Psi_1, \Psi_2, \Psi_3, \ldots$ converges to Ψ. Let Z be the set consisting of all points $z \in (0,1)$ for which either $\Psi(\{z\} \times [0,1]) > 0$ or $\Psi([0,1] \times \{z\}) > 0$ or $\Delta^*(\{z\}) > 0$. Then Z is at most countable. So if $\mathcal{I}$ is the system of all intervals $I \subseteq [0,1]$ whose endpoints do not belong to Z then $\mathcal{I}$ is closed under taking finite intersections and it generates the sigma-algebra of all Borel subsets of $[0,1]$. Moreover, whenever $I, J \in \mathcal{I}$ then the boundary of $I \times J$ is of Ψ-measure 0 and of Δ^*-measure 0. Denote by Ψ^x and Ψ^y the marginals of Ψ on the first and on the second coordinate, respectively. Then for every $I \in \mathcal{I}$ we have

$$\begin{aligned}
\Psi^x(I) &= \Psi(I \times [0,1]) = \lim_{n\to\infty} \Psi_n(I \times [0,1]) \\
&= \lim_{n\to\infty} \beta\left(\{x \in B\colon\ (f(x), f_n(x)) \in I \times [0,1]\}\right) \\
&= \lim_{n\to\infty} \Delta(I) = \Delta(I).
\end{aligned}$$

As this is true for every $I \in \mathcal{I}$, it clearly follows that $\Psi^x = \Delta$. On the other hand, for every $I \in \mathcal{I}$ we have that

$$\begin{aligned}
\Psi^y(I) &= \Psi([0,1] \times I) = \lim_{n\to\infty} \Psi_n([0,1] \times I) \\
&= \lim_{n\to\infty} \beta\left(\{x \in B\colon\ (f(x), f_n(x)) \in [0,1] \times I\}\right) \\
&= \lim_{n\to\infty} \Delta_n(I) \\
\boxed{\Delta_n \xrightarrow{w^*} \Delta^*} \quad &= \Delta^*(I).
\end{aligned}$$

So, again we have $\Psi^y = \Delta^*$. To finish the proof, it remains to show that Ψ satisfies (4.13). That is, we need to show that

$$\int_{(x,y)\in C\times[0,1]} x \,\mathrm{d}\Psi(x,y) = \int_{(x,y)\in C\times[0,1]} y \,\mathrm{d}\Psi(x,y) \tag{4.16}$$

for every Borel measurable subset C of $[0,1]$. Again, it is enough to show (4.16) only for every $C \in \mathcal{I}$. So fix $C \in \mathcal{I}$, $\varepsilon > 0$ and find some partition $\{I_1, I_2, \ldots, I_m\}$ of the interval $[0,1]$ into intervals from $\mathcal{I}$ of lengths smaller than ε. Suppose first that the interval C is of length smaller than ε. Fix some points $x_0 \in C$ and $y_j \in I_j$ (for every j). Then we have

$$\begin{aligned}
\int_{(x,y)\in C\times[0,1]} y \,\mathrm{d}\Psi(x,y) &= \sum_{j=1}^{m} \int_{(x,y)\in C\times I_j} y \,\mathrm{d}\Psi(x,y) \\
&\overset{\varepsilon\cdot\Psi(C\times[0,1])}{\approx} \sum_{j=1}^{m} y_j \Psi(C\times I_j) = \sum_{j=1}^{m} y_j \lim_{n\to\infty} \Psi_n(C\times I_j) \\
&= \sum_{j=1}^{m} y_j \lim_{n\to\infty} \beta\left(\{x\in B\colon\ (f(x), f_n(x)) \in C\times I_j\}\right) \\
&= \sum_{j=1}^{m} \lim_{n\to\infty} \int_{x\in f^{-1}(C)\cap f_n^{-1}(I_j)} y_j \,\mathrm{d}\beta \\
&\overset{\varepsilon\cdot\beta(f^{-1}(C))}{\approx} \lim_{n\to\infty} \int_{x\in f^{-1}(C)} f_n(x)\,\mathrm{d}\beta = \int_{x\in f^{-1}(C)} f(x)\,\mathrm{d}\beta \\
&\overset{\varepsilon\cdot\beta(f^{-1}(C))}{\approx} \int_{x\in f^{-1}(C)} x_0 \,\mathrm{d}\beta = x_0\cdot\beta\left(f^{-1}(C)\right) = x_0\cdot\Delta(C) = x_0\cdot\Psi^x(C) \\
&= \int_{(x,y)\in C\times[0,1]} x_0 \,\mathrm{d}\Psi(x,y) \overset{\varepsilon\cdot\Psi(C\times[0,1])}{\approx} \int_{(x,y)\in C\times[0,1]} x \,\mathrm{d}\Psi(x,y),
\end{aligned}$$

so

$$\int_{(x,y)\in C\times[0,1]} y \,\mathrm{d}\Psi(x,y) \overset{2\varepsilon\cdot(\Psi(C\times[0,1])+\beta(f^{-1}(C)))}{\approx} \int_{(x,y)\in C\times[0,1]} x \,\mathrm{d}\Psi(x,y).$$

In general, the interval C can be decomposed into subintervals belonging to $\mathcal{I}$ such that each of them is of length smaller than ε, and by additivity of integration we conclude that

$$\int_{(x,y)\in C\times[0,1]} y \,\mathrm{d}\Psi(x,y) \overset{2\varepsilon\cdot(\Psi([0,1]^2)+\beta(B))}{\approx} \int_{(x,y)\in C\times[0,1]} x \,\mathrm{d}\Psi(x,y).$$

As this is true for every $\varepsilon > 0$, we have verified (4.16). $\square$

We are now ready to prove Proposition 4.17.

Proof of Proposition 4.17, non-strict part. Suppose that we have two graphons $U, W : \Omega^2 \to [0,1]$, where $U \preceq W$. Then there exist measure preserving bijections $\pi_1, \pi_2, \pi_3, \ldots$ on Ω so that

$$W^{\pi_n} \xrightarrow{\mathrm{w}^*} U\,. \tag{4.17}$$

Now, we can apply Lemma 4.22 with $B := \Omega^2$, $f_n := W^{\pi_n}$, $\Delta_n := \mathbf{\Phi}_{W^{\pi_n}} = \mathbf{\Phi}_W$, $f := U$, and $\Delta := \mathbf{\Phi}_U$. The lemma gives that $\mathbf{\Phi}_U$ is at least as flat as $\mathbf{\Phi}_W$.

Observe that (4.17) implies that for the degree functions $\deg_U : \Omega \to [0,1]$ and $\deg_{W^{\pi_n}} : \Omega \to [0,1]$ we have $\int_A \deg_{W^{\pi_n}} \xrightarrow{\mathrm{w}^*} \int_A \deg_U$ for every measurable $A \subseteq \Omega$. Now, we can apply Lemma 4.22 with $B := \Omega$, $f_n := \deg_{W^{\pi_n}}$, $\Delta_n := \mathbf{Y}_{W^{\pi_n}} = \mathbf{Y}_W$, $f := \deg_U$, and $\Delta := \mathbf{Y}_U$. The lemma gives that $\mathbf{Y}_U$ is at least as flat as $\mathbf{Y}_W$. $\square$

Proof of Proposition 4.17, strictly flatter part. Next, suppose that $U \prec W$. Then for the measure preserving bijections $\pi_1, \pi_2, \pi_3, \ldots$ as above, we have

$$W^{\pi_n} \overset{\|\cdot\|_1}{\not\longrightarrow} U\,. \tag{4.18}$$

Indeed, suppose that (4.18) is not true, that is, $W^{\pi_n} \stackrel{\|\cdot\|_1}{\not\longrightarrow} U$. Then for measure preserving bijections $\psi_n := (\pi_n)^{-1}$ we have $U^{\psi_n} \stackrel{\|\cdot\|_1}{\longrightarrow} W$, and in particular $W \preceq U$. This is a contradiction to the fact that $U \prec W$. Now, Lemma 2.4 implies that there exists an interval $J \subseteq [0,1]$ and a «strictly bigger» interval J^+ such that

$$\nu^{\otimes 2}\left(U^{-1}(J) \setminus (W^{\pi_n})^{-1}(J^+)\right) \not\longrightarrow 0\ .$$

By passing to a subsequence, let us assume that $\nu^{\otimes 2}\left(U^{-1}(J) \setminus (W^{\pi_n})^{-1}(J^+)\right) > \varepsilon$ for each n and for some $\varepsilon > 0$. We shall now apply Lemma 4.22 twice. To this end, we take $X := U^{-1}(J)$. Furthermore, we write the measures $\mathbf{\Phi}_U$ and $\mathbf{\Phi}_{W^{\pi_n}}$ as $\mathbf{\Phi}_U = \Phi_U^X + \Phi_U^{\Omega^2\setminus X}$ and $\mathbf{\Phi}_{W^{\pi_n}} = \Phi_{W^{\pi_n}}^X + \Phi_{W^{\pi_n}}^{\Omega^2\setminus X}$, where $\Phi_U^X(L) := \nu^{\otimes 2}\left(X \cap U^{-1}(L)\right)$, $\Phi_U^{\Omega^2\setminus X}(L) := \nu^{\otimes 2}\left(U^{-1}(L) \setminus X\right)$, and the measures $\Phi_{W^{\pi_n}}^X$ and $\Phi_{W^{\pi_n}}^{\Omega^2\setminus X}$ are defined analogously. Let $\left(\Delta^X, \Delta^{\Omega^2\setminus X}\right)$ be an arbitrary accumulation point of the sequence $\left(\Phi_{W^{\pi_n}}^X, \Phi_{W^{\pi_n}}^{\Omega^2\setminus X}\right)_n$ of pairs of measures with respect to the product of weak* topologies on measures on $[0,1]^2$. Crucially, note that $\mathbf{\Phi}_W = \Delta^X + \Delta^{\Omega^2\setminus X}$.

- We first apply Lemma 4.22 with $B := X$, $f_n := (W^{\pi_n})_{\restriction X}$, $\Delta_n := \Phi_{W^{\pi_n}}^X$, $f := U_{\restriction X}$, and $\Delta := \Phi_U^X$. The lemma gives that Φ_U^X is at least as flat as Δ^X. Recall that the support of the individual measures $\Phi_{W^{\pi_n}}^X$ uniformly exceeds the interval J^+. From this, we conclude that the support of their weak* accumulation point Δ^X does not lie in J. On the other hand, observe that the support of Φ_U^X lies inside J. Thus $\Phi_U^X \neq \Delta^X$. We conclude that Φ_U^X is strictly flatter than Δ^X.
- Next, we apply Lemma 4.22 with $B := \Omega^2 \setminus X$, $f_n := (W^{\pi_n})_{\restriction \Omega^2\setminus X}$, $\Delta_n := \Phi_{W^{\pi_n}}^{\Omega^2\setminus X}$, and $f := U_{\restriction \Omega^2\setminus X}$, $\Delta := \Phi_U^{\Omega^2\setminus X}$. The lemma gives that $\Phi_U^{\Omega^2\setminus X}$ is at least as flat as $\Delta^{\Omega^2\setminus X}$.

The proof now follows by Lemma 4.21 to $\mathbf{\Phi}_W = \Delta^X + \Delta^{\Omega^2\setminus X}$. □

Let us finish this section with an auxiliary result which will be applied in Section 8. The result states that probability measures supported on $\{0,1\}$ are maximal with respect to the flat order.

Lemma 4.23. *Suppose that Λ_1 and Λ_2 are two probability measures on $[0,1]$ such that Λ_1 is strictly flatter than Λ_2. Then Λ_1 is not supported on $\{0,1\}$.*

Proof. The proof is obvious. □

4.5. **Relationship between envelopes, the cut distance, and range frequencies.** Our last result states that two envelopes are equal if and only if the corresponding graphons are weakly isomorphic. This result relies on Theorem 3.5 which will be proven in Section 5, and is used later in the proof of Corollary 6.2, which puts into relation the cut distance and the Vietoris hyperspace $K(\mathcal{W}_0)$.

Corollary 4.24. *Let $U, W \in \mathcal{W}_0$. The following are equivalent:*

- $\langle U \rangle = \langle W \rangle$,
- $\delta_\square(U, W) = 0$.

Moreover, if $U \preceq W$ then the conditions above are equivalent to

- $\mathbf{\Phi}_U = \mathbf{\Phi}_W$.

Proof. Assume that $\langle U\rangle = \langle W\rangle$. Then the conditions in Theorem 3.5(b) are satisfied because we have $\langle U\rangle = \mathbf{LIM}_{w*}(U, U, U \ldots) = \mathbf{ACC}_{w*}(U, U, U, \ldots)$. By the «furthermore» part of Theorem 3.5, the constant sequence $U, U, U, \ldots$ converges in $\delta_\square$ to a maximal element of $\langle U\rangle$. By the assumption $\langle U\rangle = \langle W\rangle$, the graphon W is a maximal element of $\langle U\rangle$. The only possibility under which a constant sequence can converge in the cut distance to another graphon is when their cut distance is 0, that is $\delta_\square(U, W) = 0$.

The opposite implication is Lemma 4.2(d).

We now turn to the «moreover» part. Suppose that $\langle U\rangle = \langle W\rangle$. By Proposition 4.17 we have that $\mathbf{\Phi}_U$ is at least as flat as $\mathbf{\Phi}_W$ and vice-versa. By Lemma 4.20 this means that $\int f\,\mathrm{d}\mathbf{\Phi}_W = \int f\,\mathrm{d}\mathbf{\Phi}_U$ for every continuous convex function $f\colon [0,1] \to \mathbb{R}$. Applying the Stone–Weierstrass theorem in the same way as in the proof of Lemma 4.20, this is true for every continuous (not necessarily convex) function $f\colon [0,1] \to \mathbb{R}$. Therefore, $\mathbf{\Phi}_U = \mathbf{\Phi}_W$.

Assume finally that $U \preceq W$ and $\mathbf{\Phi}_U = \mathbf{\Phi}_W$. Then $\mathbf{\Phi}_U$ is not strictly flatter than $\mathbf{\Phi}_W$ and therefore by Proposition 4.17, it is not the case that $U \prec W$. This means that $\langle U\rangle = \langle W\rangle$. ☐

5. Proof of Theorem 3.5

The main idea of the proof of the implication (b)⇒(a) is the following. Let W be a $\preceq$-maximal element in $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Such a W is guaranteed to exist by Lemma 4.9. Then the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ converges to W in the cut distance. To make this argument precise we first recall several definitions and results from [8].

Lemma 5.1 (Claim 1 and Claim 2 in [8])**.** *Let $W, \Gamma_1, \Gamma_2, ... \in \mathcal{W}_0$ be graphons defined on $\Omega = [0,1]$ and assume that $\Gamma_n \xrightarrow{w^*} W$. Take some sequence $B_1, B_2, ... \subseteq [0,1]$ of measurable sets and a subsequence $(n_k)_k$ such that $1_{B_{n_k}} \xrightarrow{w^*} s$ and ${}_{B_{n_k}}\Gamma_{n_k} \xrightarrow{w^*} \tilde{W}$ (with the notation from (4.3)) for some function $s : [0,1] \to [0,1]$ and some graphon $\tilde{W}$. We define*

$$\psi(x) = \int_0^x s(y)\,\mathrm{d}y \quad \text{and} \quad \varphi(x) = \psi(1) + \int_0^x (1 - s(y))\,\mathrm{d}y. \tag{5.1}$$

Then for almost every $(x,y) \in [0,1]^2$ we have

$$\begin{aligned} W(x,y) =& \tilde{W}(\psi(x), \psi(y))s(x)s(y) + \tilde{W}(\psi(x), \varphi(y))s(x)(1 - s(y)) \\ &+ \tilde{W}(\varphi(x), \psi(y))(1 - s(x))s(y) + \tilde{W}(\varphi(x), \varphi(y))(1 - s(x))(1 - s(y)). \end{aligned} \tag{5.2}$$

Moreover, if W is not a cut-norm accumulation point of the sequence $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ and the sets $B_1, B_2, ... \subseteq [0,1]$ are chosen to witness this fact, i.e., such that $\int_{B_n \times B_n}(\Gamma_n - W) > \varepsilon$ or $\int_{B_n \times B_n}(\Gamma_n - W) < -\varepsilon$ for some $\varepsilon > 0$ (which does not depend on n), the convex combination (5.2) is proper[h] on a set of positive $\nu^{\otimes 2}$ measure.

Proof of Theorem 3.5, (b)⇒(a). Suppose that $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ and assume that $W \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is a maximal element as given in Lemma 4.9. We may also assume that $\Gamma_n \xrightarrow{w^*} W$. We claim that this already implies that $\Gamma_n \xrightarrow{\|\cdot\|_\square} W$. Suppose not. Then by passing to a subsequence we may assume that there is an $\varepsilon > 0$ and a sequence $B_1, B_2, ...$ of Borel subsets of $[0,1]$ such that $\int_{B_n \times B_n}(\Gamma_n - W) > \varepsilon$ (or $\int_{B_n \times B_n}(\Gamma_n - W) < -\varepsilon$ which can be handled similarly). We will use versions ${}_{B_n}\Gamma_n$ of the graphons Γ_n that are defined via (4.3). We take a function $s : [0,1] \to [0,1]$ and a graphon $\tilde{W}$ such that $1_{B_n} \xrightarrow{w^*} s$ and

[h]that is, at least two summands on the right-hand side are positive

${}_{B_n}\Gamma_n \xrightarrow{\text{w}^*} \tilde{W}$. If 1_{B_n} or ${}_{B_n}\Gamma_n$ are not weak* convergent (that is, s or $\tilde{W}$ do not exist), then we pass to a suitable convergent subsequence (which we still index by $1, 2, \ldots$).

Claim. We have $\langle W \rangle = \langle \tilde{W} \rangle$.

Proof. For each $n \in \mathbb{N}$, let $\mathcal{I}_n = \{I_{n,0}, I_{n,1}, ..., I_{n,2^n-1}\}$ be a partition of $[0,1]$ into 2^n many equimeasurable intervals, i.e., $I_{n,k} = [\frac{k}{2^n}, \frac{k+1}{2^n})$. Define a measure preserving almost-bijection φ_n by

$$\varphi_n(x) = \begin{cases} \frac{k}{2^n} + x - \psi\left(\frac{k}{2^n}\right) & \text{if } x \in \left[\psi\left(\frac{k}{2^n}\right), \psi\left(\frac{k+1}{2^n}\right)\right) \text{ and} \\ \frac{k}{2^n} + \psi\left(\frac{k+1}{2^n}\right) - \psi\left(\frac{k}{2^n}\right) + x - \varphi\left(\frac{k}{2^n}\right) & \text{if } x \in \left[\varphi\left(\frac{k}{2^n}\right), \varphi\left(\frac{k+1}{2^n}\right)\right), \end{cases}$$

where ψ and φ are defined by (5.1). Define $\tilde{W}_n(x,y) = \tilde{W}\left(\varphi_n^{-1}(x), \varphi_n^{-1}(y)\right)$. We claim that $\tilde{W}_n \xrightarrow{\text{w}^*} W$. First note that if $m \in \mathbb{N}$ and $m \geq n$ and $k, l < 2^n$, then

$$\int_{I_{n,k} \times I_{n,l}} \tilde{W}_m = \sum_{i=2^{m-n}\cdot k}^{2^{m-n}\cdot(k+1)-1} \; \sum_{j=2^{m-n}\cdot l}^{2^{m-n}\cdot(l+1)-1} \int_{I_{m,i} \times I_{m,j}} \tilde{W}_m. \tag{5.3}$$

For each $n \in \mathbb{N}$ we have

$$\begin{aligned}
\int_{I_{n,k} \times I_{n,l}} \tilde{W}_n &= \int_{\psi(\frac{k}{2^n})}^{\psi(\frac{k+1}{2^n})} \int_{\psi(\frac{l}{2^n})}^{\psi(\frac{l+1}{2^n})} \tilde{W} + \int_{\psi(\frac{k}{2^n})}^{\psi(\frac{k+1}{2^n})} \int_{\varphi(\frac{l}{2^n})}^{\varphi(\frac{l+1}{2^n})} \tilde{W} \\
&+ \int_{\varphi(\frac{k}{2^n})}^{\varphi(\frac{k+1}{2^n})} \int_{\psi(\frac{l}{2^n})}^{\psi(\frac{l+1}{2^n})} \tilde{W} + \int_{\varphi(\frac{k}{2^n})}^{\varphi(\frac{k+1}{2^n})} \int_{\varphi(\frac{l}{2^n})}^{\varphi(\frac{l+1}{2^n})} \tilde{W} \\
&= \int_{I_{n,k} \times I_{n,l}} \tilde{W}(\psi(x), \psi(y)) s(x) s(y) + \int_{I_{n,k} \times I_{n,l}} \tilde{W}(\psi(x), \varphi(y)) s(x)(1 - s(y)) \\
&+ \int_{I_{n,k} \times I_{n,l}} \tilde{W}(\varphi(x), \psi(y))(1 - s(x)) s(y) + \int_{I_{n,k} \times I_{n,l}} \tilde{W}(\varphi(x), \varphi(y))(1 - s(x))(1 - s(y)) \\
&= \int_{I_{n,k} \times I_{n,l}} W.
\end{aligned}$$

Plugging this back to (5.3) gives

$$\int_{I_{n,k} \times I_{n,l}} \tilde{W}_m = \int_{I_{n,k} \times I_{n,l}} W\,.$$

This implies that for every $n \in \mathbb{N}$ and $k, l < 2^n$, we have $\int_{I_{n,k} \times I_{n,l}} \tilde{W}_m \to \int_{I_{n,k} \times I_{n,l}} W$ as $m \to \infty$. Consequently $\tilde{W}_m \xrightarrow{\text{w}^*} W$ and $W \in \langle \tilde{W} \rangle$. On the other hand, $\tilde{W} \in \langle W \rangle$ since $\tilde{W} \in \mathbf{LIM}_{\text{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, and the maximality of W in $\mathbf{LIM}_{\text{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. $\square$

Now Proposition 4.17 together with Lemma 4.20 give us that $\int f \; \mathrm{d}\mathbf{\Phi}_{\tilde{W}} = \int f \; \mathrm{d}\mathbf{\Phi}_W$ for every continuous convex function. So to get the desired contradiction it suffices to prove the following claim.

Claim. Let $f : [0,1] \to [0,1]$ be a strictly convex function. Then $\int f \; \mathrm{d}\mathbf{\Phi}_{\tilde{W}} > \int f \; \mathrm{d}\mathbf{\Phi}_W$.

Proof. Recall that the convex combination in (5.2) is proper on a set of positive $\nu^{\otimes 2}$ measure. Therefore the strict convexity of f gives us

$$\begin{aligned}
\int_0^1 f(x)\,\mathrm{d}\mathbf{\Phi}_W &= \int_0^1\int_0^1 f(W(x,y)) \\
&< \int_0^1\int_0^1 f\left(\tilde{W}(\psi(x),\psi(y))s(x)s(y)\right) + \int_0^1\int_0^1 f\left(\tilde{W}(\psi(x),\varphi(y))s(x)(1-s(y))\right) \\
&\quad + \int_0^1\int_0^1 f\left(\tilde{W}(\varphi(x),\psi(y))(1-s(x))s(y)\right) \\
&\quad + \int_0^1\int_0^1 f\left(\tilde{W}(\varphi(x),\varphi(y))(1-s(x))(1-s(y))\right) \\
&= \int_0^{\psi(1)}\int_0^{\psi(1)} f\left(\tilde{W}(x,y)\right) + \int_0^{\psi(1)}\int_{\psi(1)}^1 f\left(\tilde{W}(x,y)\right) \\
&\quad + \int_{\psi(1)}^1\int_0^{\psi(1)} f\left(\tilde{W}(x,y)\right) + \int_{\psi(1)}^1\int_{\psi(1)}^1 f\left(\tilde{W}(x,y)\right) \\
&= \int_0^1\int_0^1 f(\tilde{W}(x,y)) = \int_0^1 f(x)\,\mathrm{d}\mathbf{\Phi}_{\tilde{W}}\,.
\end{aligned}$$

□

Let $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ be a sequence of graphons which is Cauchy with respect to the cut distance. By Theorem 3.3 there is a subsequence $\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots$ such that $\mathbf{LIM}_{w*}(\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots) = \mathbf{ACC}_{w*}(\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots)$. By the proof of implication (b)⇒(a) this subsequence converges to some W in the cut distance. As the original sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ is Cauchy with respect to the cut distance it follows that it converges to W as well. We may suppose that $\Gamma_n \xrightarrow{\|\cdot\|_\square} W$. To conclude that $\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots) = \mathbf{ACC}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ consider any $\Gamma_{n_k}^{\varphi_k} \xrightarrow{w^*} U$. We claim that then also $W^{\varphi_k} \xrightarrow{w^*} U$. To see this fix some Borel set $A\subseteq[0,1]$ then

$$\left|\int_{A\times A}(W^{\varphi_k}-U)\right| \le \left|\int_{\varphi_k(A)\times\varphi_k(A)}(W-\Gamma_{n_k})\right| + \left|\int_{A\times A}(\Gamma_{n_k}^{\varphi_k}-U)\right| \to 0$$

where the first term tends to 0 by $\|\cdot\|_\square$-convergence and the second by the weak* convergence. Then we use the same trick again:

$$\left|\int_{A\times A}(\Gamma_n^{\varphi_n}-U)\right| \le \left|\int_{\varphi_n(A)\times\varphi_n(A)}(\Gamma_n-W)\right| + \left|\int_{A\times A}(W^{\varphi_n}-U)\right| \to 0\,.$$

This gives us $\Gamma_n^{\varphi_n} \xrightarrow{w^*} U$ which means that $U\in\mathbf{LIM}_{w*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. □

6. Relating the hyperspace $K(\mathcal{W}_0)$ and the cut distance

Lemma *3.1(b)* says that envelopes are members of the hyperspace $K(\mathcal{W}_0)$. First, we provide a proof of an extension of Theorem 3.5, stated in Theorem 6.1 below. In addition to the original statement, we include a characterization of cut distance convergence in terms of the hyperspace $K(\mathcal{W}_0)$, and also describe the limit graphon.

Theorem 6.1. *Let $W,\Gamma_1,\Gamma_2,\Gamma_3,\ldots\in\mathcal{W}_0$. The following are equivalent:*

(a) $\Gamma_n \xrightarrow{\delta_\square} W$,
(b) $\langle\Gamma_n\rangle \to \langle W\rangle$ *in* $K(\mathcal{W}_0)$,

(c) $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots) = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$ *and* $W \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$ *is the* $\preceq$*-maximal element of* $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$.

Proof. Condition (c) is equivalent to Condition (a) by Theorem 3.5. Before we prove that (b) is equivalent to (c) we recall that a basic open neighborhood of $K \in K(\mathcal{W}_0)$ is given by some finite open cover $\mathcal{U}$ of K, namely for every such cover $\mathcal{U}$ we define the open neighborhood of K as $\mathcal{O}_{\mathcal{U}} = \{L : L \subseteq \bigcup_{O \in \mathcal{U}} O \text{ and } \forall O \in \mathcal{U}\ L \cap O \neq \emptyset\}$.

(c) $\implies$ (b). First note that $\langle W \rangle = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$. Suppose that $\langle \Gamma_n \rangle \not\to \langle W \rangle$ in $K(\mathcal{W}_0)$. There must be a basic open neighborhood $\mathcal{O}_{\mathcal{U}}$ of $\langle W \rangle$ and a subsequence $\langle \Gamma_{k_1} \rangle, \langle \Gamma_{k_2} \rangle, \dots$ such that either $\langle \Gamma_{k_j} \rangle \not\subseteq \bigcup_{O \in \mathcal{U}} O$ for every $j \in \mathbb{N}$ or there is $O \in \mathcal{U}$ such that $\langle \Gamma_{k_j} \rangle \cap O = \emptyset$ for every $j \in \mathbb{N}$. If the first possibility happens, then clearly $\mathbf{ACC}_{w*}\left(\Gamma_{k_1}, \Gamma_{k_2}, \dots\right) \cap (\mathcal{W}_0 \setminus \bigcup_{O \in \mathcal{U}} O) \neq \emptyset$, which is a contradiction with $\mathbf{ACC}_{w*}\left(\Gamma_{k_1}, \Gamma_{k_2}, \dots\right) = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots) \subseteq \bigcup_{O \in \mathcal{U}} O$. If the second possibility occurs, for some $O \in \mathcal{U}$ then clearly $\mathbf{ACC}_{w*}\left(\Gamma_{k_1}, \Gamma_{k_2}, \dots\right) \cap O = \emptyset$. That is also a contradiction once $\mathbf{ACC}_{w*}\left(\Gamma_{k_1}, \Gamma_{k_2}, \dots\right) = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots) = \langle W \rangle$ and $\langle W \rangle \cap O \neq \emptyset$ because $\langle W \rangle \in \mathcal{O}_{\mathcal{U}}$.

(b) $\implies$ (c). Let $U \in \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$. We claim that $U \in \langle W \rangle$. Assume it is not true and take any open set $O \subseteq \mathcal{W}_0$ in the weak* topology such that $\langle W \rangle \subseteq O$ and U is not in the closure of O. Then by the convergence in $K(\mathcal{W}_0)$ we may find some $m \in \mathbb{N}$ such that for every $k \geq m$ we have $\langle \Gamma_k \rangle \subseteq O$. This implies that $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$ is a subset of the closure of O and that is a contradiction because we assumed that U is not in that closure. Therefore, we have just proved that $\mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots) \subseteq \langle W \rangle$.

Let $U \in \langle W \rangle$, then for every open set $O \subseteq \mathcal{W}_0$ in the weak* topology such that $U \in O$ we can find $m \in \mathbb{N}$ such that for every $k \geq m$ we have $\langle \Gamma_k \rangle \cap O \neq \emptyset$. Since $\mathcal{W}_0$ with the weak* topology is a metric space we may find a sequence of versions such that $\Gamma_n^{\varphi_n} \xrightarrow{w^*} U$. Therefore, we have $\langle W \rangle \subseteq \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$, and so $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots) = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$. By the first part of the argument it follows that $\langle W \rangle = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$ and thus W is the $\preceq$-maximal element of $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \dots)$. $\square$

We can now formulate Corollary 6.2, which is the final statement of this section. It allows us to transfer the space $\left(\widetilde{\mathcal{W}_0}, \delta_\square\right)$ into the hyperspace $K(\mathcal{W}_0)$. This transference will be useful later.

Prior to giving the statement, observe that for the envelope map $\langle \cdot \rangle$ defined on $\mathcal{W}_0$ we have $\langle W_1 \rangle = \langle W_2 \rangle$ for weakly isomorphic graphons W_1 and W_2. That allows us to define $\langle \cdot \rangle$ even on the factor-space $\widetilde{\mathcal{W}_0}$.

Corollary 6.2. *The envelope map* $\langle \cdot \rangle : \left(\widetilde{\mathcal{W}_0}, \delta_\square\right) \to K(\mathcal{W}_0)$ *is a continuous injection. We have that* $\left(\widetilde{\mathcal{W}_0}, \delta_\square\right)$ *is homeomorphic to some closed subspace* X *of* $K(\mathcal{W}_0)$*. Moreover, the metric* $\delta_\square$ *is equivalent*[i] *to the pullback* χ *of the hyperspace metric on* $K(\mathcal{W}_0)$ *defined in (2.3), that is,*

$$\chi\left([\![U]\!], [\![W]\!]\right) = \max\left\{\sup_{\varphi}\left\{\inf_{\psi}\{d_{w^*}(U^\varphi, W^\psi)\}\right\}, \sup_{\varphi}\left\{\inf_{\psi}\{d_{w^*}(U^\psi, W^\varphi)\}\right\}\right\}, \quad [\![U]\!], [\![W]\!] \in \widetilde{\mathcal{W}_0},$$

where φ *and* ψ *range through all measure preserving bijections on* Ω*.*[j]

[i]Recall that two metrics on a topological space are *equivalent* if they give the same topology.

[j]Note that the definition of $\chi\left([\![U]\!], [\![W]\!]\right)$ does not depend on the particular representatives U and W.

Finally, $\left(\widetilde{\mathcal{W}_0},\delta_\square\right)$ *is compact.*

Proof. The map $\langle\cdot\rangle$ is well-defined and injective by Corollary 4.24 and it is continuous by Theorem 6.1. The set $X=\left\langle\widetilde{\mathcal{W}_0}\right\rangle$ is closed in the Vietoris topology. Indeed, suppose that $\langle\Gamma_n\rangle\to L$ in $K(\mathcal{W}_0)$ for some $L\in K(\mathcal{W}_0)$. Then it follows from [13, Exercise 4.21, Exercise 4.23] that $L=\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,...)=\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,...)$, see the definition of topological limit before [13, Exercise 4.23]. By Theorem 3.5 we know that there is some $W\in\mathcal{W}_0$ such that $\Gamma_n\xrightarrow{\delta_\square}W$. By the continuity of $\langle.\rangle$ we also have $\langle\Gamma_n\rangle\to\langle W\rangle$ in $K(\mathcal{W}_0)$.

$K(\mathcal{W}_0)$ is compact by Fact 2.7 and Remark 2.8. Since X is a closed subset of $K(\mathcal{W}_0)$, it is also compact. By Theorem 6.1, we have that the inverse map $\langle\cdot\rangle^{-1}$ is also continuous. Therefore $\widetilde{\mathcal{W}_0}$ is homeomorphic to X, and hence compact. Therefore, $\delta_\square$ and χ give the same compact topology on $\widetilde{\mathcal{W}_0}$. $\square$

7. Relation to multiway cuts

In this section, we compare our view with [5]. The main result of [5] is a statement that several properties of graph(on) sequences are equivalent to cut distance convergence. These properties include so-called right-convergence and convergence of free energies. As we shall see, however, the one that relates to our approach involves so-called multiway cuts.

Let us give the key definition (see p. 178 in [5]). First, for two real matrices M and N of the same dimensions, their L_1*-distance* is defined as $\|M-N\|_1:=\sum_{i,j}\left|M_{ij}-N_{ij}\right|$. Suppose that $\mathbf{a}=\left(a_1,\ldots,a_q\right)$ is a vector with non-negative entries that sum up to 1. Given a graphon $W:\Omega^2\to[0,1]$, we write $\mathcal{S}_\mathbf{a}(W)$ for the set of all $q\times q$ matrices M that can be obtained in the following way. Partition suitably $\Omega=\Omega_1\sqcup\ldots\sqcup\Omega_q$ so that $\nu(\Omega_i)=a_i$. Then for each $i,j\in[q]$, define $M_{ij}:=\int_{\Omega_i}\int_{\Omega_j}W$. So, $\mathcal{S}_\mathbf{a}(W)\subseteq[0,1]^{q\times q}$. In particular, given two graphons U and W, we can talk about the *Hausdorff distance with respect to* ℓ_1 of the sets $\mathcal{S}_\mathbf{a}(U)$ and $\mathcal{S}_\mathbf{a}(W)$,

$$d_1^{\mathrm{Hf}}\left(\mathcal{S}_\mathbf{a}(U),\mathcal{S}_\mathbf{a}(W)\right):=\max\left\{\sup_{M\in\mathcal{S}_\mathbf{a}(U)}\inf_{N\in\mathcal{S}_\mathbf{a}(W)}\|M-N\|_1,\sup_{M\in\mathcal{S}_\mathbf{a}(W)}\inf_{N\in\mathcal{S}_\mathbf{a}(U)}\|M-N\|_1\right\}.$$

Then one of the main equivalences of [5], given there in Theorem 3.5 and stated here in a slightly tailored form[k], is the following.

Theorem 7.1 ([5])**.** *Let* $\Gamma_1,\Gamma_2,\Gamma_3,\ldots\in\mathcal{W}_0$*. The following are equivalent:*

(a) *The sequence* $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ *is Cauchy with respect to the cut distance* $\delta_\square$*,*
(b) *For each* $\mathbf{a}$ *as above, the sequence* $\mathcal{S}_\mathbf{a}(\Gamma_1),\mathcal{S}_\mathbf{a}(\Gamma_2),\mathcal{S}_\mathbf{a}(\Gamma_3),\ldots$ *is Cauchy with respect to* d_1^{Hf}*.*

To close the circle, we prove that the second conditions of Theorem 3.5 and of Theorem 7.1 are equivalent. We shall use the Hausdorff metric $d_{\mathrm{w}^*}^{\mathrm{Hf}}$ on $K(\mathcal{W}_0)$ derived as in (2.3) from the weak* metric defined, say, by (2.1) .

Proposition 7.2. *Let* $\Gamma_1,\Gamma_2,\Gamma_3,\ldots\in\mathcal{W}_0$*. The following are equivalent:*

(a) $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)=\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$*,*
(b) *The sequence* $\langle\Gamma_1\rangle,\langle\Gamma_2\rangle,\langle\Gamma_3\rangle,\ldots$ *is Cauchy with respect to* $d_{\mathrm{w}^*}^{\mathrm{Hf}}$*,*

Also note that this statement holds for any metric d_{w^*} compatible with the weak* topology, not only the one given in (2.1).

[k]More precisely, Theorem 3.5 of [5] does not talk about $\mathcal{S}_\mathbf{a}(\cdot)$, but about a slightly different object $\mathcal{S}_q(\cdot)$. However, equivalence of these two approaches easily follows from Lemma 4.5 of [5].

(c) For each $\mathbf{a}$ *as above, the sequence* $\mathcal{S}_{\mathbf{a}}(\Gamma_1), \mathcal{S}_{\mathbf{a}}(\Gamma_2), \mathcal{S}_{\mathbf{a}}(\Gamma_3), \ldots$ *is Cauchy with respect to* d_1^{Hf}*.*

Proof. First we show that (c) implies (a). Suppose that $\Omega = [0,1]$ and let $\mathcal{I}_k = \{I_{k,1}, \ldots, I_{k,k}\}$ be the canonical partition of $[0,1]$ into equimeasurable intervals, i.e., $\lambda(I_{k,i}) = \frac{1}{k}$ and $I_{k,i}$ precedes $I_{k,i+1}$. If $\mathcal{P} = \{P_1, \ldots, P_k\}$ is another equitable partition of $[0,1]$, then we denote as $\pi_{\mathcal{P}}$ any fixed measure preserving bijection such that $\pi_{\mathcal{P}}(P_i) = I_{k,i}$ for every $i \in [k]$. Let $W \in \mathbf{ACC}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. We may suppose that there is a subsequence $\{n_j\}_{j\in\mathbb{N}}$ such that $\Gamma_{n_j} \xrightarrow{\mathrm{w}^*} W$. Fix $k \in \mathbb{N}$. It is easy to see that $\left(\Gamma_{n_j}^{\ltimes \mathcal{I}_k}\right) \xrightarrow{\mathrm{w}^*} W^{\ltimes \mathcal{I}_k}$. Define $M_{n_j}(r,s) = \int_{I_{k,r}\times I_{k,s}} \Gamma_{n_j}$ and $N_k(r,s) = \int_{I_{k,r}\times I_{k,s}} W$ for every $r,s \in [k]$. Then we have $M_{n_j} \in \mathcal{S}_{\mathbf{a}}\left(\Gamma_{n_j}\right)$ for every $j \in \mathbb{N}$ and $N_k \in \mathcal{S}_{\mathbf{a}}(W)$ where $\mathbf{a} = \left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$. Moreover, $M_{n_j} \xrightarrow{\|\cdot\|_1} N_k$. The assumption (c) allows to find an equitable partition $\mathcal{P}_n = \{P_i\}_{\in[k]}$ for every $n \notin \{n_j : j \in \mathbb{N}\}$ such that if we define $M_n(r,s) = \int_{P_r\times P_s} \Gamma_n$ for every $r,s \in [k]$, then we have $M_n \xrightarrow{\|\cdot\|_1} N_k$. One can easily verify that for $W_n := \left((\Gamma_n)^{\ltimes \mathcal{P}_n}\right)^{\pi_{\mathcal{P}_n}}$ we have $W_n = \left((\Gamma_n)^{\pi_{\mathcal{P}_n}}\right)^{\ltimes \mathcal{I}_k}$ and therefore $W_n \in \langle \Gamma_n \rangle$ by Lemma4.1. This implies that $W_n \xrightarrow{\mathrm{w}^*} W^{\ltimes \mathcal{I}_k}$ because $W_n \restriction I_{k,r} \times I_{k,s} = k^{-2} \cdot M_n(r,s)$ holds for every $n \in \mathbb{N}$ and $r,s \in [k]$. Hence, $W^{\ltimes \mathcal{I}_k} \in \mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Finally, since $W^{\ltimes \mathcal{I}_k} \xrightarrow{\mathrm{w}^*} W$ and $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ is weak* closed by Lemma3.1 (a), we conclude that $W \in \mathbf{LIM}_{\mathrm{w}*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$.

The direction from (a) to (b) follows from Corollary 6.2. It remains to show that (b) implies (c). Let $\mathbf{a} \in [0,1]^q$ be a vector whose entries sum up to 1, $\epsilon > 0$ and $\mathcal{P} = \{P_i\}_{i=1}^q$ be a partition of Ω with the property that $\nu(P_i) = a_i$. Recall that we have fixed a collection $\{A_r\}_{r\in\mathbb{N}}$ of measurable subsets of $[0,1]$ that is dense in the measure algebra and that defines the Hausdorff metric $d_{\mathrm{w}^*}^{\mathrm{Hf}}$ on $K(\mathcal{W}_0)$ by (2.1). It follows that there is a sequence $\{A_{r_i}\}_{i=1}^q$ such that $\sum_{i=1}^q \lambda(P_i \triangle A_{r_i}) < \epsilon$. Let $W, U \in \mathcal{W}$ be such that $d_{\mathrm{w}^*}^{\mathrm{Hf}}(\langle U \rangle, \langle W \rangle) < \frac{\epsilon}{2^{2r_q}}$.

Let $M \in \mathcal{S}_{\mathbf{a}}(W)$. Then there are measure preserving bijections π, ϕ such that $M(i,j) = \int_{P_i\times P_j} W^\phi = \int_{\phi^{-1}(P_i)\times\phi^{-1}(P_j)} W$ and $d_{w^*}\left(\left(W^\phi\right)^{\ltimes\mathcal{P}}, U^\pi\right) < \frac{\epsilon}{2^{2r_q}}$. Define $N \in \mathcal{S}_{\mathbf{a}}(U)$ as $N(i,j) = \int_{P_i\times P_j} U^\pi = \int_{\pi^{-1}(P_i)\times\pi^{-1}(P_j)} U$. This gives

$$
\begin{aligned}
\|M - N\|_1 &= \sum_{i,j=1}^{q} \left| \int_{P_i\times P_j} \left(W^\phi\right)^{\ltimes\mathcal{P}} - U^\pi \right| \le \sum_{i,j=1}^{q} \left| \int_{A_{r_i}\times A_{r_j}} \left(W^\phi\right)^{\ltimes\mathcal{P}} - U^\pi \right| + \epsilon \\
&\le 2^{2r_q} \sum_{i,j=1}^{q} 2^{-(r_i+r_j)} \left| \int_{A_{r_i}\times A_{r_j}} \left(W^\phi\right)^{\ltimes\mathcal{P}} - U^\pi \right| + \epsilon \le 2^{2r_q} d_{\mathrm{w}^*}^{\mathrm{Hf}}(\langle U \rangle, \langle W \rangle) + \epsilon \\
&\le 2\epsilon.
\end{aligned}
$$

This shows that if $d_{\mathrm{w}^*}^{\mathrm{Hf}}(\langle U \rangle, \langle W \rangle) < \frac{\epsilon}{2^{2r_q}}$, then $d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)) \le 2\epsilon$. Consequently, (b) implies (c). □

8. Properties of the structuredness (quasi)order

Above, we obtained properties of the structuredness (quasi)order $\preceq$ that were needed for our abstract proof of Theorem 6.1. In this section, we establish further properties of $\preceq$. In Lemma 8.1 we prove that $\preceq$ is actually a closed order on $\widetilde{\mathcal{W}}_0$. In Corollary 8.3, we prove that

$\preceq$-increasing/decreasing chains of graphons are cut distance convergent. In Corollary 8.4, we characterize the elements of $K(\mathcal{W}_0)$ that are envelopes of graphons. Finally, in Proposition 8.4, we characterize $\preceq$-minimal and $\preceq$-maximal elements. This last-mentioned result is just starting point of investigating the structure of the poset $\preceq$, which we leave open.

Let us first prove that $\preceq$ is actually an order modulo weak isomorphism.

Lemma 8.1. *The relation $\preceq$ on the space $\widetilde{\mathcal{W}_0}$ is an order, and as a subset of $\widetilde{\mathcal{W}_0} \times \widetilde{\mathcal{W}_0}$ it is closed.*

Proof. Since by Corollary 6.2 the space $\left(\widetilde{\mathcal{W}_0}, \delta_\square\right)$ is homeomorphic to some closed subspace of $K(\mathcal{W}_0)$ and the relation $\preceq$ is interpreted as $\subseteq$ on $K(\mathcal{W}_0)$ it is enough to verify the properties for the relation $\subseteq$ on $K(\mathcal{W}_0)$. But both properties are trivially satisfied for the relation $\subseteq$. $\square$

Next, we turn our attention to finding upper and lower bounds with respect to $\preceq$. Let us first give an auxiliary result, which is then utilized in Corollaries 8.3 and 8.4.

Proposition 8.2. *(a) Suppose that $P \subseteq \mathcal{W}_0$ is upper-directed in the structuredness order, i.e., for every $U_0, U_1 \in P$ there is $V \in P$ such that $U_0, U_1 \preceq V$. Then there is a graphon $W \in \mathcal{W}_0$ such that W is the supremum of P with respect to $\preceq$.*

(b) Suppose that $P \subseteq \mathcal{W}_0$ is down-directed in the structuredness order, i.e., for every $U_0, U_1 \in P$ there is $V \in P$ such that $V \preceq U_0, U_1$. Then there is a graphon $W \in \mathcal{W}_0$ such that W is the infimum of P with respect to $\preceq$.

Proof. Let us focus on (a). First of all consider the set $\langle P \rangle = \{\langle U \rangle : U \in P\}$. This set is upper-directed with respect to $\subseteq$ in $K(\mathcal{W}_0)$. Let K be the weak* closure of $\bigcup_{U \in P} \langle U \rangle$. Clearly, $K \in K(\mathcal{W}_0)$. Further, K is the supremum of $\langle P \rangle$ with respect to $\subseteq$ on $K(\mathcal{W}_0)$. To finish the proof, we only need to show that there exists $W \in \mathcal{W}_0$ such that $K = \langle W \rangle$. Consider a countable set $P_0 \subseteq P$ such that $\langle P_0 \rangle$ is dense in $\langle P \rangle$. This can be done since $K(\mathcal{W}_0)$ is separable metrizable by Fact 2.7. Take some enumeration $U_1, U_2, ...$ of P_0. Define inductively an increasing chain $\Gamma_1, \Gamma_2, ... \in P$ such that for every $n \in \mathbb{N}$ we have that $U_1, ..., U_n, \Gamma_{n-1} \preceq \Gamma_n$. This can be done since P is upper-directed. Since $\Gamma_1 \preceq \Gamma_2 \preceq \Gamma_3 \preceq \ldots$, we have $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots) = \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Indeed, whenever we take $\Gamma \in \mathbf{ACC}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, say $\Gamma_{n_1}^{\pi_{n_1}}, \Gamma_{n_2}^{\pi_{n_2}}, \Gamma_{n_3}^{\pi_{n_3}}, \ldots \xrightarrow{\|\cdot\|_\square} \Gamma$, then for each index i in the interval (n_k, n_{k+1}), we can use that $\Gamma_{n_k} \preceq \Gamma_i$ to approximate $\Gamma_{n_k}^{\pi_{n_k}}$ by some version $\Gamma_i^{\pi_i}$ of Γ_i (with a vanishing error as $i \to \infty$). With the gaps (n_k, n_{k+1}) filled-in this way, we have $\Gamma_{n_1}^{\pi_{n_1}}, \Gamma_{1+n_1}^{\pi_{1+n_1}}, \Gamma_{2+n_1}^{\pi_{2+n_1}}, \ldots \xrightarrow{\|\cdot\|_\square} \Gamma$, and consequently $\Gamma \in \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, as we wanted. Moreover $K = \mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$ because $\langle P_0 \rangle$ is dense in $\langle P \rangle$. By Lemma 4.9, we can pick a $\preceq$-maximal element W of $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$. Now, we have that $K = \langle W \rangle$.

The proof of (b) is similar. The only difference is that the desired infimum is of the form $K = \bigcap_{U \in P} \langle U \rangle$ and we inductively build a decreasing sequence. $\square$

Along very similar lines, we can prove that $\preceq$-increasing/decreasing chains of graphons are cut distance convergent.

Corollary 8.3. *(a) Suppose that $W_1 \preceq W_2 \preceq W_3 \preceq \ldots$ is a sequence of graphons. Then this sequence is cut distance convergent.*

(b) Suppose that $W_1 \succeq W_2 \succeq W_3 \succeq \ldots$ is a sequence of graphons. Then this sequence is cut distance convergent.

Proof. Suppose that $W_1 \preceq W_2 \preceq W_3 \preceq \ldots$. Then we have $\mathbf{LIM}_{w*}(W_1, W_2, W_3, \ldots) = \mathbf{ACC}_{w*}(W_1, W_2, W_3, \ldots)$. By Lemma 4.9, we can pick a $\preceq$-maximal element W of $\mathbf{LIM}_{w*}(W_1, W_2, W_3, \ldots)$. Now, by Theorem 6.1 we have that $W_n \xrightarrow{\delta_\square} W$.

The proof for a decreasing sequence is the same. □

Next, we characterize the elements of $K(\mathcal{W}_0)$ that are envelopes of graphons.

Corollary 8.4. *Let $K \in K(\mathcal{W}_0)$. Then there exists $W \in \mathcal{W}_0$ such that $K = \langle W \rangle$ if and only if K is upper-directed (for every $U_0, U_1 \in K$ there is $V \in K$ such that $U_0, U_1 \preceq V$) and downwards closed (for every $U \in K$ and $V \preceq U$ we have $V \in K$).*

Proof. Suppose first that $K = \langle W \rangle$. Then for every $U_0, U_1 \in K$, we have $U_0, U_1 \preceq W$. That is, K is upper-directed. Suppose next that $U \in K$ and $V \preceq U$. Let $\epsilon > 0$ be arbitrary. Since $V \preceq U$, there exists a version U^π of U such that $d_{w^*}(V, U^\pi) < \frac{\epsilon}{2}$. Since $U \in K$, we have $U \preceq W$. Thus, also $U^\pi \preceq W$. Therefore, there exists a version W^θ of W such that $d_{w^*}\left(U^\pi, W^\theta\right) < \frac{\epsilon}{2}$. We conclude that $d_{w^*}\left(V, W^\theta\right) < \epsilon$. Since ϵ was arbitrarily small and since K is weak* closed, we conclude that $V \in K$.

Let us now turn to the other implication. As K is upper-directed then by Proposition 8.2 we can take its supremum W. Because K is downwards closed we have $\langle W \rangle = K$. □

Last, we characterize the minimal and maximal elements of the structuredness order; the latter part being suggested to us by László Miklós Lovász.

Proposition 8.5. *The minimal elements of the structuredness order are exactly constant graphons, and for every graphon W there is a minimal graphon $W_{\min} \preceq W$. The maximal elements of the structuredness order are exactly 0-1 valued graphons, and for every graphon W there is a maximal graphon $W_{\max} \succeq W$.*

Proof. The first part follows directly from the fact that an envelope of any graphon contains a constant graphon (see Lemma 4.2(b)), and also the graphon itself. For the second part we at first prove that only 0-1 valued graphons can be maximal.

Suppose that W is a graphon such that its value is neither 0 nor 1 on a set of positive measure. This implies that there is an $\varepsilon > 0$ such that W has values between ε and $1-\varepsilon$ on a set of positive measure. Now consider a map $\varphi\colon [0,1] \to [0,1]$ such that $\varphi(x) = 2x$ for $0 \le x \le \frac{1}{2}$ and $\varphi(x) = 2x - 1$ for $\frac{1}{2} < x \le 1$. The graphon W^φ contains four copies of W scaled by a factor of one half. Let $B \subseteq [0,1]^2$ be the set, on which W^φ takes values between ε and $1-\varepsilon$. Let $\widetilde{W}$ be a graphon such that $\widetilde{W} = W^\varphi + \varepsilon$ for $x, y \in \left([0,\frac{1}{2}]^2 \cup [\frac{1}{2},1]^2\right) \cap B$, $\widetilde{W} = W^\varphi - \varepsilon$ for $x, y \in \left([0,\frac{1}{2}] \times [\frac{1}{2},1] \cup [\frac{1}{2},1] \times [0,\frac{1}{2}]\right) \cap B$ and $\widetilde{W} = W$ otherwise. The values of $\widetilde{W}$ are bounded by 0 and 1 and W and $\widetilde{W}$ are not weakly isomorphic (compare $\mathrm{INT}_f(W)$ and $\mathrm{INT}_f(\widetilde{W})$ for any strictly convex function f). Moreover, we claim that $W \in \left\langle \widetilde{W} \right\rangle$. To see this, one can construct a sequence of measure preserving almost-bijections $\psi_1, \psi_2, \ldots$, defined as $\psi_n(x) = \frac{\lfloor 2nx \rfloor}{2n} + x$ for $0 \le x \le \frac{1}{2}$ and $\psi_n(x) = \frac{\lfloor 2nx \rfloor - 2n + 1}{2n} + x$ for $\frac{1}{2} \le x \le 1$, that interlace the two intervals $[0,\frac{1}{2}]$ and $[\frac{1}{2},1]$ and thus serve as an approximation of φ. The fact that $\widetilde{W}_1^{\psi_1^{-1}}, \widetilde{W}_2^{\psi_2^{-1}}, \ldots \xrightarrow{w^*} W$ can be seen directly from the definition of weak* convergence.

Next, we prove that all 0-1 graphons are maximal. Indeed, let W be a 0-1 valued graphon, and suppose that there exists some graphon U such that $U \succ W$. Then for the measures

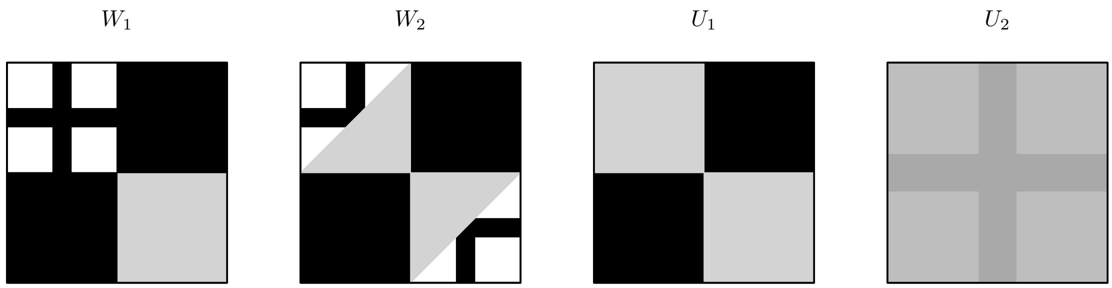

FIGURE 8.1. The four graphons W_1, W_2, U_1, U_2 witnessing that an intersection of two envelopes is not necessarily an envelope.

$\mathbf{\Phi}_U$ and $\mathbf{\Phi}_W$ we have that $\mathbf{\Phi}_W$ is strictly flatter than $\mathbf{\Phi}_U$ by Proposition 4.17. This contradicts Lemma 4.23.

Finally, let W be an arbitrary graphon. Consider the set $\mathcal{P}$ of all graphons $P \succeq W$. Then every chain in $\mathcal{P}$ has a supremum in the structuredness order by Proposition 8.2(a). Therefore, we can apply Zorn's lemma to conclude that there is a maximal graphon $W_{\max} \succeq W$. $\square$

Remark 8.6. *Let us take $W \equiv \frac{1}{2}$ and a sequence W_n of graphons corresponding to Erdős–Rényi random graphs $\mathbb{G}(n, \frac{1}{2})$. By Proposition 8.5, W is a $\prec$-minimal elements, while each W_n is a $\prec$-maximal element. Yet, it is well-known that $W_n \xrightarrow{\delta_\square} W$ almost surely. This example shows that even a cut distance convergent sequence can be fairly «separated» from the limit point with respect to the structurdness order. (Note also, that while W is a $\prec$-minimal elements, and each W_n is a $\prec$-maximal element, W_n and W are typically incomparable in the structuredness order, since the density of W_n is typically not exactly $\frac{1}{2}$.)*

In the example below we show that the structuredness order does not have meet and joins in general.

Example 8.7. We shall construct graphons W_1, W_2, U_1, U_2 such that we have $U_1, U_2 \preceq W_1, W_2$, but there is no graphon V such that $U_1, U_2 \preceq V \preceq W_1, W_2$. For a fixed $\varepsilon > 0$ we define the four graphons (on the unit square, denoting the Lebesgue measure on $[0, 1]$ as ν) as follows (the definitions should be clear from Figure 8.1).

(1) Define $W_1(x, y) = 1$ if and only if $(x, y) \in [\frac{1}{2}, 1] \times [0, \frac{1}{2}] \;\cup\; [0, \frac{1}{2}] \times [\frac{1}{2}, 1] \;\cup\; [\frac{1}{4} - \frac{\varepsilon}{2}, \frac{1}{4} + \frac{\varepsilon}{2}] \times [0, 1] \;\cup\; [0, 1] \times [\frac{1}{4} - \frac{\varepsilon}{2}, \frac{1}{4} + \frac{\varepsilon}{2}]$. Moreover, $W_1 = 4(\varepsilon - \varepsilon^2)$ for $(x, y) \in [\frac{1}{2}, 1]^2$ and is zero otherwise.
(2) Define W_2 as W_1 but switch its values on the triangle with vertices $[0, \frac{1}{2}], [\frac{1}{2}, \frac{1}{2}], [\frac{1}{2}, 0]$ with the triangle with vertices $[\frac{1}{2}, 1], [1, 1], [1, \frac{1}{2}]$. Note that $\int_{[0,\frac{1}{2}]^2} W_1 = \int_{[\frac{1}{2},1]^2} W_1 = \int_{[0,\frac{1}{2}]^2} W_2 = \int_{[\frac{1}{2},1]^2} W_2$.
(3) Define $U_1 = 4(\varepsilon - \varepsilon^2)$ for $(x, y) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2$ and 1 otherwise. Note that $U_1 \preceq W_1, W_2$ (this follows by applying Lemma 4.2(b) twice just for the top-left and bottom-right subgraphons of the graphons W_1 and W_2).
(4) Finally, define U_2 to be such that

$$U_2(x, y) = \frac{1}{4}\left(W_1\left(\frac{x}{2}, \frac{y}{2}\right) + W_1\left(\frac{x+1}{2}, \frac{y}{2}\right) + W_1\left(\frac{x}{2}, \frac{y+1}{2}\right) + W_1\left(\frac{x+1}{2}, \frac{y+1}{2}\right)\right).$$

We have again $U_2 \preceq W_1, W_2$ (consider a sequence of bijections interlacing the two intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, as in the proof of Proposition 8.5).

We are now ready to get to the main two properties of this example.

Claim (Claim A). If V is a graphon such that $U_1 \preceq V$, then V contains a set J such that V is 1 almost everywhere on $J \times ([0,1] \setminus J)$.

Proof. The argument is similar to the proof of [7, Lemma 3] and [12, Lemma 2.3]. Observe that in U_1, the values on the rectangle $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$ are 1 everywhere. Since there exists a sequence of measure preserving transformations $\pi_1, \pi_2, \ldots$ such that $V^{\pi_n} \xrightarrow{w^*} U_1$, taking $J_n := \pi_n\left(\left[0, \frac{1}{2}\right]\right)$ yields sets of measure $\frac{1}{2}$ with the property that

$$\lim_n \int_{J_n \times ([0,1] \setminus J_n)} V = \int_{[0,\frac{1}{2}] \times [\frac{1}{2},1]} U_1 = \frac{1}{4} . \tag{8.1}$$

Let us pass to a subsequence so that the sequence of indicators $\left(\mathbf{1}_{J_n}\right)_n$ converges weak*, say to a function $g : [0,1] \to [0,1]$. Note that then the sequence $\left(\mathbf{1}_{[0,1] \setminus J_n}\right)_n$ converges weak* to $1 - g$.

We claim that V is 1 almost everywhere on $(\operatorname{supp} g) \times (\operatorname{supp}(1-g))$. Before proving this, let us explain, why this gives the desired set J. To this end, note that since $\nu(J_n) = \frac{1}{2}$, weak* convergence implies $\int g = \frac{1}{2}$. In particular, the measure of the support of g is at least $\frac{1}{2}$ while at the same time the measure of points where $g \equiv 1$ is at most $\frac{1}{2}$. So, we take J to be an arbitrary set of measure exactly $\frac{1}{2}$ contained in the former set and containing the latter set. It is clear that all the required properties are satisfied.

So, let us prove that V is 1 almost everywhere on $(\operatorname{supp} g) \times (\operatorname{supp}(1-g))$. It suffices to prove that for each $a > 0$, the Lebesgue measure of those pairs $(x,y) \in g^{-1}((a,1]) \times g^{-1}([0,1-a))$ for which $V(x,y) < 1 - a$ is zero. So, suppose that this fails for some $a > 0$. By basic properties of measure, we can then find two sets $X \subseteq g^{-1}((a,1])$, $Y \subseteq g^{-1}([0,1-a))$ of positive measure such that

$$\nu^{\otimes 2}\left(\{(x,y) \in X \times Y : V(x,y) \geq 1 - a\}\right) < \frac{a^3}{100} \cdot \nu(X)\nu(Y) . \tag{8.2}$$

Now, observe that since g is a weak* limit of $\left(\mathbf{1}_{J_n}\right)_n$ and $X \subseteq g^{-1}((a,1])$, we have that for all sufficiently large n, $\nu(X \cap J_n) > a \cdot \nu(X)$. By a similar reasoning, $\nu\left(Y \cap ([0,1] \setminus J_n)\right) > a \cdot \nu(Y)$ for all sufficiently large n. To say this in words, a substantial (non-vanishing) part of the set J_n is inside the set X and a substantial part of the set $[0,1] \setminus J_n$ is inside the set Y, and by (8.2), V is very far from being close to 1 on $X \times Y$. This contradicts (8.1), which is saying that V must be very close to 1 on most of $J_n \times ([0,1] \setminus J_n)$. □

Claim (Claim B). There does not exist any graphon V such that $U_1, U_2 \preceq V \preceq W_1, W_2$.

Proof. Assume that there is a graphon V such that $U_1, U_2 \preceq V \preceq W_1, W_2$. Let the set J be given from Claim A. Without loss of generality assume that $J = [0, \frac{1}{2}]$. Now, we turn our attention to the graphons W_i, $i = 1, 2$. From the fact that there is a sequence of measure preserving transformations $\varphi_1, \varphi_2, \ldots$ such that $W_1^{\varphi_1}, W_1^{\varphi_2}, \ldots \xrightarrow{w^*} V$ and, thus, $\lim_{n\to\infty} \int_{[0,\frac{1}{2}] \times [\frac{1}{2},1]} W_1^{\varphi_n} = \int_{[0,\frac{1}{2}] \times [\frac{1}{2},1]} V$, we obtain that for any $\delta > 0$ there is n sufficiently large such that either $\nu\left(\varphi_n\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) \leq \delta$ or $\nu\left(\varphi_n\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) \geq \frac{1}{2} - \delta$. Indeed, suppose that $\frac{1}{2} - \delta > \nu\left(\varphi_n\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) > \delta$

and observe that the density of $W_1^{\varphi_n}$ is equal to $4(\varepsilon-\varepsilon^2)$ on a subset of $[0,\frac{1}{2}]\times[\frac{1}{2},1]$ of measure at least δ^2. Now it suffices to recall that $\int_{[0,\frac{1}{2}]\times[\frac{1}{2},1]} W_1^{\varphi_n} \to \nu\left([0,\frac{1}{2}]\times[\frac{1}{2},1]\right)=\frac{1}{4}$ which implies that the assumption is false for large enough n. From the fact that either $\nu\left(\varphi_n\left([0,\frac{1}{2}]\right)\cap[0,\frac{1}{2}]\right)\le\delta$ or $\nu\left(\varphi_n\left([0,\frac{1}{2}]\right)\cap[0,\frac{1}{2}]\right)\ge\frac{1}{2}-\delta$ we conclude that actually (after passing to a subsequence) some versions of $W_1\cap[0,\frac{1}{2}]^2, W_1\cap[0,\frac{1}{2}]^2,\ldots$ converge weak* to either $V\cap[0,\frac{1}{2}]^2$ or $V\cap[\frac{1}{2},1]^2$ (redefine φ_n on a set of small measure such that it maps $[0,\frac{1}{2}]$ either onto $[0,\frac{1}{2}]$ or onto $[\frac{1}{2},1]$ to get the required measure preserving bijections). Without loss of generality assume that some versions of $W_1\cap[\frac{1}{2},1]^2, W_1\cap[\frac{1}{2},1]^2,\ldots$ converge weak* to $V\cap[\frac{1}{2},1]^2$. Similarly, we get that there are versions of $W_2\cap[0,\frac{1}{2}]^2, W_2\cap[0,\frac{1}{2}]^2,\ldots$ converging weak* to $V\cap[0,\frac{1}{2}]^2$ (the other case when the versions $W_2\cap[\frac{1}{2},1]^2, W_2\cap[\frac{1}{2},1]^2,\ldots$ converge weak* to $V\cap[0,\frac{1}{2}]^2$ can be treated in the same way).

Notice that

$$\int_{[0,1]\times[\frac{1}{2}-\varepsilon,\frac{1}{2}+\varepsilon]} U_2=(2\varepsilon)\cdot 1\cdot\frac{3}{4}+o(\varepsilon)=\frac{3}{2}\varepsilon+o(\varepsilon)\ . \tag{8.3}$$

On the other hand,

$$\begin{aligned}
\sup_{\nu(C)=2\varepsilon}\int_{[0,1]\times C} V &= \sup_{A\subseteq[0,\frac{1}{2}],B\subseteq[\frac{1}{2},1],\nu(A\cup B)=2\varepsilon}\left(\int_{[0,1]\times A} V+\int_{[0,1]\times B} V\right)\\
&=2\varepsilon\cdot\frac{1}{2}+\sup_{A\subseteq[0,\frac{1}{2}],B\subseteq[\frac{1}{2},1],\nu(A\cup B)=2\varepsilon}\left(\int_{[0,\frac{1}{2}]\times A} V+\int_{[\frac{1}{2},1]\times B} V\right)\\
&\le\varepsilon+\sup_{A\subseteq[0,\frac{1}{2}],B\subseteq[\frac{1}{2},1],\nu(A\cup B)=2\varepsilon}\left(\int_{[0,\frac{1}{2}]\times A} W_2+\int_{[\frac{1}{2},1]\times B} W_1\right)\\
&=\varepsilon+\left(\frac{1}{4}\varepsilon+o(\varepsilon)\right)+o(\varepsilon)\\
&=\frac{5}{4}\varepsilon+o(\varepsilon), \qquad (8.4)
\end{aligned}$$

which, for ε small enough, is smaller than the appropriate value for U_2 appearing in (8.3). Now, we can conclude that $V\not\succeq U_2$. Indeed, suppose that $V\succeq U_2$. Then for every $a>0$ there is a version V^π of V such that $d_{\mathrm{w}^*}(V^\pi,U_2)<a$. In particular, there exists a version V^π such that $\left|\int_{[0,1]\times[\frac{1}{2}-\varepsilon,\frac{1}{2}+\varepsilon]} V^\pi-U_2\right|<\frac{1}{100}$. Taking $C=\pi^{-1}\left([\frac{1}{2}-\varepsilon,\frac{1}{2}+\varepsilon]\right)$, (8.4) contradicts (8.3). □

9. Conclusion

Inspired by previous work [8], we created a comprehensive theory of approaching the cut distance convergence via the weak* topology. The main results, Theorem 3.5 and Theorem 3.3, say that in each sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ there exists a subsequence $\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots$ such that $\mathbf{LIM}_{\mathrm{w}*}(\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots)=\mathbf{ACC}_{\mathrm{w}*}(\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots)$, and that the latter property is equivalent to the cut distance convergence of the said subsequence. It follows from the equivalence established in Proposition 7.2 that Theorem 3.5 can be regarded as the functional-analytic formulation of Theorem 7.1, which was originally established in [5]. So, while each of the above mentioned theorems could have been obtained from previously existing results, we believe that the weak* approach we introduce provides an important perspective on the theory of graph limits. For example, Corollary 8.3 offers families of convergent graphon sequences, which are very natural with this perspective but would be even difficult to define previously. It turns out that many of these concepts extend (in a nontrivial way) to hypergraphons, and this is currently work in progress.

Our initial motivation was to build a theory rather than to solve any particular problem, quite contrary to the usual perception of combinatorics (see [11]). That said, we already see that this abstract theory has its fruits on the problem-solving side: in [9] we use it to prove that a connected graph is weakly norming if and only if it is step Sidorenko, and that if a graph is norming then it is step forcing (see [9, Remark 3.20] for an explanation of the role of the the weak* theory in these results). We see a potential for more applications in the theory of hypergraphons.

The weak* approach may provide alternative proofs of various generalizations of Theorem 1.1, such as Banach space valued graphons [15], representations of limits of sparse graphs using measures [16], and an L^p-approach to limits of sparse graphs [3]. To indicate the plausibility of this approach in the setting L^p-limits (which seems the most feasible one), we recall that the corresponding notion of weak* convergence in this setting is based on tests against all L^q-functions (where $\frac{1}{p}+\frac{1}{q}=1$). This would of course change the definition of $\mathbf{LIM}_{w*}(\Gamma_1, \Gamma_2, \Gamma_3, \ldots)$, and all related definitions. Working out this program in detail is currently in progress.

Acknowledgments

We would like to thank László Miklós Lovász, Daniel Král and Oleg Pikhurko for their useful suggestions, and Frederik Garbe and Ondřej Kůrka for comments on a preliminary version of this paper. We thank a referee for his or her comments.

References

[1] T. Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probab. Surv.*, 5:80–145, 2008.

[2] P. Billingsley. *Probability and Measure*, volume 60 of *Wiley series in probability and mathematical statistics*. John Wiley & sons, New York, NY., 1995.

[3] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Trans. Amer. Math. Soc.*, 372(5):3019–3062, 2019.

[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] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012.

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

[7] M. Doležal and J. Hladký. Matching polytons. *Electron. J. Combin.*, 26(4):Paper 4.38, 33, 2019.

[8] J. Doležal, M. Hladký. Cut-norm and entropy minimization over weak* limits. *J. Combin. Theory Ser. B*, 137:232–263, 2019.

[9] M. Doležal, J. Grebík, J. Hladký, I. Rocha, and V. Rozhoň. Cut distance identifying graphon parameters over weak* limits. arXiv:1809.03797v4.

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

[11] W. T. Gowers. The two cultures of mathematics. In *Mathematics: frontiers and perspectives*, pages 65–78. Amer. Math. Soc., Providence, RI, 2000.

[12] J. Hladký, P. Hu, and D. Piguet. Komlós's tiling theorem via graphon covers. *J. Graph Theory*, 90(1):24–45, 2019.

[13] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[14] D. Král', T. Martins, P. P. Pach, and M. Wrochna. The step Sidorenko property and non-norming edge-transitive graphs. *J. Combin. Theory Ser. A*, 162:34–54, 2019.

[15] D. Kunszenti-Kovács, L. Lovász, and B. Szegedy. Multigraph limits, unbounded kernels, and banach space decorated graphs. arXiv:1406.7846.

[16] D. Kunszenti-Kovács, L. Lovász, and B. Szegedy. Measures on the square as sparse graph limits. *J. Combin. Theory Ser. B*, 138:1–40, 2019.

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

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

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

[20] G. Regts and A. Schrijver. Compact orbit spaces in Hilbert spaces and limits of edge-colouring models. *European J. Combin.*, 52(part B):389–395, 2016.

Doležal: INSTITUTE OF MATHEMATICS, CZECH ACADEMY OF SCIENCES. ŽITNÁ 25, 110 00, PRAHA, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985840, *Grebík:* MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY, CV4 7AL, UK. *This work was done while affiliated with:* INSTITUTE OF MATHEMATICS, CZECH ACADEMY OF SCIENCES. ŽITNÁ 25, 110 00, PRAHA, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985840, *Hladký:* INSTITUTE OF MATHEMATICS, CZECH ACADEMY OF SCIENCES. ŽITNÁ 25, 110 00, PRAHA, CZECHIA. *Part of this work was done while affiliated with:* INSTITUT FÜR GEOMETRIE, TU DRESDEN, 01062 DRESDEN, GERMANY, *Rocha*: INSTITUTE OF COMPUTER SCIENCE, CZECH ACADEMY OF SCIENCES. POD VODÁRENSKOU VĚŽÍ 2, 182 07, PRAGUE, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985807., *Rozhoň:* ETH ZÜRICH, SWITZERLAND. *This work was done while affiliated with:* INSTITUTE OF COMPUTER SCIENCE, CZECH ACADEMY OF SCIENCES. POD VODÁRENSKOU VĚŽÍ 2, 182 07, PRAGUE, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985807.

E-mail address: `doleza1@math.cas.cz`, `greboshrabos@seznam.cz`, `honzahladky@gmail.com`, `israelrocha@gmail.com`, `rozhonv@ethz.ch`