INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Cut distance identifying graphon parameters over weak* limits

Martin Doležal
Jan Grebík
Jan Hladký
Israel Rocha
Václav Rozhoň

Preprint No. 49-2018

PRAHA 2018

CUT DISTANCE IDENTIFYING GRAPHON PARAMETERS OVER WEAK* LIMITS

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

ABSTRACT. The theory of graphons comes with the so-called cut distance. The cut distance is finer than the weak* topology. Doležal and Hladký [arXiv:1705.09160] showed, that given a sequence of graphons, a cut distance accumulation graphon can be pinpointed in the set of weak* accumulation points as minimizers of the entropy. Motivated by this, we study graphon parameters with the property that their minimizers or maximizers identify cut distance accumulation points over the set of weak* accumulation points. We call such parameters *cut distance identifying*.

Of particular importance are cut distance identifying parameters coming from subgraph densities, $t(H,\cdot)$. It turns out that this concept is closely related to graph norms. In particular, we prove that a connected graph H is step Sidorenko (a concept very similar to $t(H,\cdot)$ being cut distance identifying) if and only if it is weakly norming. This answers a question of Král', Martins, Pach and Wrochna [arXiv:1802.05007].

Further, we study convexity properties of cut distance identifying graphon parameters, and find a way to identify cut distance limits using spectra of graphons.

1. INTRODUCTION

The theory of graphons, initiated in [2, 19] and covered in depth in [18], provides a powerful formalism for handling large graphs that are dense, i.e., they contain a positive proportion of edges. In this paper, we study the relation between the cut norm and the weak* topology on the space of graphons through various graphon parameters. Let us give basic definitions needed to explain our motivation and results.

We write $\mathcal{W}_0$ for the space of all *graphons*, i.e., all symmetric measurable functions from Ω^2 to $[0,1]$. Here as well as in the rest of the paper, Ω is an arbitrary separable atomless probability space with probability measure ν. Given a graphon W and a measure preserving bijection $\varphi:\Omega\to\Omega$, we define a *version of* W by

$$W^{\varphi}(x,y)=W(\varphi(x),\varphi(y))\;. \tag{1.1}$$

Let us recall that the *cut norm* is defined by[a]

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

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

Hladký was supported by the Alexander von Humboldt Foundation. *Rocha* and *Rozhoň* were supported by the Czech Science Foundation, grant number GJ16-07822Y. *Doležal* and *Hladký* were supported by the Czech Science Foundation, grant number GJ18-01472Y. *Grebík* was supported by the Czech Science Foundation, grant number 17-33849L..

[a]All the sets and functions below are tacitly assumed to be measurable.

Given two graphons U and W we define in (1.2) their *cut norm distance* and in (1.3) their *cut distance*,

$$d_{\square}(U,W) := \|U-W\|_{\square} \text{ , and} \tag{1.2}$$

$$\delta_{\square}(U,W) := \inf_{\varphi:\Omega\to\Omega \text{ m.p.b.}} d_{\square}(U,W^{\varphi}) \ . \tag{1.3}$$

Recall that the key property of the space $\mathcal{W}_0$, which makes the theory such a powerful in applications in extremal graph theory, random graphs, property testing, and other areas, is its compactness with respect to the cut distance. The result was first proven by Lovász and Szegedy [19] using the regularity lemma,[b] and then by Elek and Szegedy [11] using ultrafilter techniques, by Austin [1] and Diaconis and Janson [8] using the theory of exchangeable random graphs, and finally by Doležal and Hladký [9] and by Doležal, Grebík, Hladký, Rocha and Rozhoň [10] in a way explained below. For our later purposes, it is more convenient to state the result in terms of the cut norm distance.

Theorem 1.1. *For every sequence $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ of graphons there is a subsequence $\Gamma_{n_1},\Gamma_{n_2},\Gamma_{n_3},\ldots$, measure preserving bijections $\pi_{n_1},\pi_{n_2},\pi_{n_3},\ldots:\Omega\to\Omega$ and a graphon Γ such that $d_{\square}\left(\Gamma_{n_i}^{\pi_{n_i}},\Gamma\right)\to 0$.*

Let us now explain the approach from [9] and from [10], which is based on the weak* topology. Recall that a sequence of graphons $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ *converges weak** to a graphon W if for every $S,T\subset\Omega$ we have

$$\lim_{n\to\infty}\int_{S\times T}\Gamma_n-\int_{S\times T}W=0 \ .$$

From this, we get that the weak* topology is weaker than the topology generated by $d_{\square}$, of which the former can be viewed as a certain uniformization.

So, the idea in [9] and [10], on a high level, is to look on the set $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ of all weak* accumulation points of sequences,

$$\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)=\bigcup_{\pi_1\pi_2,\pi_3,\ldots:\Omega\to\Omega \text{ m.p.b.}} \text{weak* accumulation points of } \Gamma_1^{\pi_1},\Gamma^{\pi_2}{}_2,\Gamma_3^{\pi_3},\ldots \ .$$

and locate in the set $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ one graphon Γ that is an accumulation point not only with respect to the weak* topology but also with respect to the cut norm distance. In [9], this was done by choosing Γ as a maximizer[c] of an operator $\mathrm{INT}_f(\cdot)$, defined for a continuous strictly convex function $f:[0,1]\to\mathbb{R}$ by

$$\mathrm{INT}_f(W) := \int_x\int_y f(W(x,y)) \ . \tag{1.4}$$

In [10], we then approached Theorem 1.1 by more abstract means. Namely, we showed that Γ can be chosen as the element with the maximum «envelope» in $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$. We recall the notion of envelopes only in Section 2.5. For now, it suffices to say that each envelope is a subset of $L^{\infty}(\Omega^2)$ and the notion maximality is with respect to the set inclusion. In the current paper, we return to the program from [9]. We provide a comprehensive study of graphon parameters where the maximization problem over $\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)$ pinpoints

[b]see also [20] and [21] for variants of this approach

[c]In fact, the supremum of $\left\{\mathrm{INT}_f(W) : W\in\mathbf{ACC}_{\mathrm{w}*}(\Gamma_1,\Gamma_2,\Gamma_3,\ldots)\right\}$ need not be attained (see [9, Section 7.4]), so the rigorous treatment needs to be a bit more technical. Similarly, we simplify the presentation of the approach from [10] below. The correct way is shown in Theorems 3.3 and 3.4.

cut distance accumulation points. We call such parameters «cut distance identifying» and «cut distance compatible» (definitions are given in Section 3.1). In Section 3.1 we sketch that each cut distance identifying parameter can indeed be used to prove Theorem 1.1. As we explain in Section 3.1.1, the defining properties of cut distance identifying parameters in particular imply that they can be used for characterization of graph quasi-random sequences, in the spirit of the Chung–Graham–Wilson Theorem. As we show, the two most prominent parameters in the Chung–Graham–Wilson Theorem, the 4-cycle density and the spectrum of the adjacency matrix, indeed possess these stronger properties and can be used as cut distance identifying parameters.

In Section 3.2 we reprove the result of Doležal and Hladký and show that the assumption of f being continuous (1.4) is not really needed. This result is a short application of our concept of so-called «range frequencies» which we previously introduced in [10] (this notion is recalled in Section 2.5). In particular, our current approach gives as a shorter proof of the results from [9], even when the necessary theory from [10] is counted.

In Section 3.4, we prove that when the spectrum is turned appropriately into a graphon parameter (using the «spectral quasiorder» which we define in Section 2.2.3), we indeed get a cut distance identifying graphon parameter.

Last, but most importantly, we study in Section 3.5 cut distance identifying and cut distance compatible graph parameters coming of the form $t(H, \cdot)$, that is, densities of a fixed graph H. As we show, this is tightly linked with concepts from graph norms (with notions such norming graphs and (weakly) Hölder graphs). It is also tightly related to concepts of graphs with the «step Sidorenko property» and the «step forcing property» introduced recently by Král', Martins, Pach and Wrochna [15]. In fact, it follows from Proposition 3.1 that a graph H has the step Sidorenko property if and only if $t(H, \cdot)$ is cut distance compatible and an analogous equivalence between the step forcing property and cut distance identifying parameters $t(H, \cdot)$ would follow from Conjecture 3.2. In Theorem 3.19 we prove if for a connected graph H we have that $t(H, \cdot)$ is cut distance compatible, then H is weakly Hölder (the opposite implication was already known). This in particular answers a question of Král', Martins, Pach and Wrochna [15, page 20].

2. Preliminaries

In this section we recall some notation, standard facts about graphons, and some results from [10] which we will build on in this paper. Parts of this section are borrowed from [10].

2.1. **General notation.** We write $\overset{\varepsilon}{\approx}$ for equality up to ε. For example, $1 \overset{0.2}{\approx} 1.1 \overset{0.2}{\approx} 1.3$. 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 $A_0 \subset A$ and $B_0 \subset B$ so that $f_{\restriction A \setminus A_0}$ is a bijection between $A \setminus A_0$ and $B \setminus B_0$. Note that in (1.3), we could have worked with measure preserving almost-bijections φ instead.

2.2. **Graphon basics.** Our notation is mostly standard, following [18]. Let us fix a separable measure space Ω with a probability measure ν. Let $\mathcal{W}$ denote the space of *kernels*, i.e. all bounded symmetric measurable real functions defined on Ω^2. We always work modulo differences on null-sets. We write $\mathcal{W}_0 \subset \mathcal{W}$ for the space of all *graphons*, that is, symmetric measurable functions from Ω^2 to $[0,1]$, and $\mathcal{W}^+ \subset \mathcal{W}$ for the space of all bounded symmetric measurable functions from Ω^2 to $[0,+\infty)$. We write $\nu^{\otimes k}$ for the product measure on Ω^k.

For $p \in [0,1]$, we write $\mathcal{G}_p = \left\{W \in \mathcal{W}_0 : \int_x \int_y W(x,y) = p\right\}$ for all graphons with edge density p.

Remark 2.1. *It is a classical fact that there is a measure preserving almost-bijection between each two separable atomless probability spaces. So, while most of the time we shall work with graphons on Ω^2, a graphon defined on a square of any other probability space as above can be represented (even though not in a unique way) on Ω^2.*

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 $(x,y) \in \Omega^2$.

2.2.1. *Subgraph densities.* As usual, given a finite graph H on the vertex set $\{v_1, v_2, \dots, v_n\}$ and a graphon W, we write

$$t(H,W) := \int_{x_1\in\Omega}\int_{x_2\in\Omega}\cdots\int_{x_n\in\Omega}\prod_{v_iv_j\in E(H)} W(x_i,x_j) \tag{2.1}$$

for the *density* of H in W. Note that (2.1) extends to all $W \in \mathcal{W}$. We call the quantity $t(P_2, W) = \int_x \int_y W(x,y)$ 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 $x \in \Omega$, $\deg_W(x) = p$.

We will need to generalize homomorphism densities to decorated graphs, as is done in [18, p. 120]. A *$\mathcal{W}$-decorated graph* is a finite simple graph H on the vertex set $\{v_1, v_2, \dots, v_n\}$ in which each edge $v_iv_j \in E(H)$ is labelled by an element $W_{v_iv_j} \in \mathcal{W}$. We denote such a $\mathcal{W}$-decorated graph by (H,w), where $w = \left(W_{v_iv_j}\right)_{v_iv_j\in E(H)}$. For such a $\mathcal{W}$-decorated graph (H,w) we define

$$t(H,w) = \int_{x_1\in\Omega}\int_{x_2\in\Omega}\cdots\int_{x_n\in\Omega}\prod_{v_iv_j\in E(H)} W_{v_iv_j}(x_i,x_j)\ .$$

2.2.2. *Tensor product.* Finally, we will need the definition of the tensor product of two graphons. Suppose that $U, V : \Omega^2 \to [0,1]$ are two graphons. We define their *tensor product* as a $[0,1]$-valued function $U \otimes V : \left(\Omega^2\right)^2 \to [0,1]$ by $(U\otimes V)\left((x_1,x_2),(y_1,y_2)\right) = U(x_1,y_1)V(x_2,y_2)$.

Using Remark 2.1, we can think of $U \otimes V$ as a graphon in $\mathcal{W}_0$. Note that for every graph H we have

$$t(H, U\otimes V) = \int_{\Omega^{2|H|}}\prod_{v_iv_j\in E(H)} U(x_i,x_j)\prod_{v_iv_j\in E(H)} V(x_{|H|+i}, x_{|H|+j}) = t(H,U)\cdot t(H,V)\ .$$

One can deal with the generalised homomorphism density for decorations on a fixed finite graph H (where the tensor product $w_1 \otimes w_2$ is defined coordinatewise) in the same way and get that

$$t(H, w_1\otimes w_2) = t(H,w_1)\cdot t(H,w_2)\ . \tag{2.2}$$

2.2.3. *Spectrum and the spectral quasiorder.* We recall basic spectral theory for graphons, details and proofs can be found in [18, §7.5]. We shall work with the real Hilbert space $L^2(\Omega)$, inner product on which is denoted by $\langle\cdot,\cdot\rangle$. Given a graphon $W:\Omega^2\to[0,1]$, we can associate to it an operator $T_W: L^2(\Omega)\to L^2(\Omega)$,

$$(T_W f)(x) := \int_y W(x,y)f(y)\ .$$

T_W is a Hilbert–Schmidt operator, and hence has a discrete spectrum of finitely or countably many non-zero eigenvalues (with possible multiplicities). All these eigenvalues are real, bounded in modulus by 1, and their only possible accumulation point is 0. For a given graphon W denote its eigenvalues by

$$\lambda_1^+(W)\ge\lambda_2^+(W)\ge\lambda_3^+(W)\ge\ldots\ge 0\ ,$$
$$\lambda_1^-(W)\le\lambda_2^-(W)\le\lambda_3^-(W)\le\ldots\le 0\ .$$

(We pad zeros if the spectrum has only finitely many positive or negative eigenvalues.)

We now introduce the notion of spectral quasiorder (this definition has not appeared in other literature). We write $W\overset{S}{\preceq}U$ if $\lambda_i^+(W)\le\lambda_i^+(U)$ and $\lambda_i^-(W)\ge\lambda_i^-(U)$ for all $i=1,2,3,\ldots$. Further we write $W\overset{S}{\prec}U$ if $W\overset{S}{\preceq}U$ and at least one of the above inequalities is strict. Then $\overset{S}{\preceq}$ is a quasiorder on $\mathcal{W}_0$, which we call the *spectral quasiorder*.

Recall that the eigenspaces are pairwise orthogonal. Recall also that (see e.g. [18, p. 124])

$$\|W\|_2^2=\sum_i\lambda_i^+(W)^2+\sum_i\lambda_i^-(W)^2\ . \tag{2.3}$$

In Section 3.5 we shall use the following formula connecting eigenvalues and cycle densities. For any graphon W for any $k\ge 3$, we have by [18, eq. (7.22), (7.23)],

$$t(C_k,W)=\sum_i\lambda_i^+(W)^k+\sum_i\lambda_i^-(W)^k\ . \tag{2.4}$$

2.2.4. *The stepping operator.* Suppose that $W:\Omega^2\to[0,1]^2$ 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 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 $\nu(\Omega_j)=0$.

In [18], a stepping is denoted by $\Gamma_{\mathcal{P}}$ rather than $\Gamma^{\Join\mathcal{P}}$. The following easy technical result taken from [10, Lemma 2.5] will be used.

Lemma 2.3. *For every graphon $\Gamma:\Omega^2\to[0,1]$ and every $\varepsilon>0$ there exists a finite partition $\mathcal{P}$ of Ω such that $\left\|\Gamma-\Gamma^{\Join\mathcal{P}}\right\|_1<\varepsilon$.*

We call $\Gamma^{\Join\mathcal{P}}$ with properties as in Lemma 2.3 an *averaged L^1-approximation of Γ by a step-graphon for precision ε.*

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 and $U^{\ltimes\mathcal{P}} = W$.

2.3. **Norms defined by graphs.** In this section we briefly recall how subgraph densities $t(H,\cdot)$ induce norms on the space of graphons. More details can be found in [18, §14.1].

We now introduce the seminorming and weakly norming graphs and graphs with the (weak) Hölder property, concepts first introduced in [12]. We say that a graph H is *(semi)norming*, if the function $\|W\|_H := t(H,W)^{1/e(H)}$ is a (semi)norm on $\mathcal{W}$. This means that we require that $t(H,\cdot)$ is subadditive and homogeneous (i.e., $t(H,cW) = c \cdot t(H,W)$), and in the case of norming graphs we moreover assume that there is not a kernel U that is not identically zero, but $t(H,U) = 0$. Similarly, we say that a graph H is *weakly norming*, if the function $\|W\|_H := t(H,|W|)^{1/e(H)}$ is a seminorm on $\mathcal{W}$. Note that by adding the absolute values we change our attention to the space $\mathcal{W}_0$ of graphons. Also note that in this case every seminorm is also a norm, since if U is a graphon that is not zero almost everywhere, it is bounded from zero on a rectangle of positive measure, and therefore $t(H,U) > 0$.

Since the homomorphism density is homogeneous, the only nontrivial requirement is the triangle inequality for the homomorphism density defined on the space of kernels $\mathcal{W}$, or the space of graphons $\mathcal{W}_0$, respectively. In other words we ask that for each $W_1, W_2 \in \mathcal{W}$, or for each $W_1, W_2 \in \mathcal{W}_0$, we have

$$\|W_1 + W_2\|_H \le \|W_1\|_H + \|W_2\|_H \tag{2.5}$$

Complete bipartite graphs (in particular, stars), complete balanced bipartite graphs without a perfect matching , even cycles, and Hamming cubes are the known examples of weakly norming graphs.

A graph H has the *Hölder property*, if for every $\mathcal{W}$-decoration $w = (W_e)_{e\in E(H)}$ of H we have

$$t(H,w)^{e(H)} \le \prod_{e\in E(H)} t(H,W_e)\ . \tag{2.6}$$

The graph H has the *weak Hölder property*, if (2.6) holds for every $\mathcal{W}_0$-decoration w of H.

Remark 2.4. *Due to the homogeneity of (2.6), we could have defined the weak the Hölder property by testing over all $\mathcal{W}^+$-decorations w of H . For the same reason, it is enough to test the Hölder property only over all $\mathcal{W}$-decorations w of H that satisfy $t(H,W_e) = 1$ for every W_e.*

2.4. **Topologies on $\mathcal{W}_0$.** There are several natural topologies on $\mathcal{W}_0$. The $\|\cdot\|_\infty$ topology inherited from the normed space $L^\infty(\Omega^2)$, the $\|\cdot\|_1$ topology inherited from the normed space $L^1(\Omega^2)$, the topology given by the $\|\cdot\|_\square$ norm, the weak* topology inherited from the weak* topology of the dual Banach space $L^\infty(\Omega^2)$, and the weak topology inherited from the weak topology of the Banach space $L^1(\Omega^2)$. Note that $\mathcal{W}_0$ is closed in both $L^1(\Omega^2)$ and $L^\infty(\Omega^2)$. We write $d_1(\cdot,\cdot)$ for the distance derived from the $\|\cdot\|_1$ norm and $d_\infty(\cdot,\cdot)$ for the distance derived from the $\|\cdot\|_\infty$ norm. The weak* topology of the dual Banach space $L^\infty(\Omega^2)$ is generated by elements of its predual $L^1(\Omega^2)$. That means that the weak* topology on $L^\infty(\Omega^2)$ is the smallest topology on $L^\infty(\Omega^2)$ such that all functionals of the form $g \in L^\infty(\Omega^2) \mapsto \int_{\Omega^2} fg$, where $f \in L^1(\Omega^2)$ is fixed, are continuous. Recall that by the Banach–Alaoglu theorem, $\mathcal{W}_0$ equipped with the weak* topology is compact. Recall also that the weak* topology on $\mathcal{W}_0$ is metrizable. We shall denote by $d_{w^*}(\cdot,\cdot)$ any metric compatible with this topology. For example, we can take some countable family $\{A_n\}_{n\in\mathbb{N}}$ of measurable subsets of Ω which forms dense set in the sigma-algebra of Ω, and define $d_{w^*}(U,W) := \sum_{n,k\in\mathbb{N}} 2^{-(n+m)} \left|\int_{A_n\times A_k}(U-W)\right|$.

2.5. **Envelopes, the structuredness order, and the pushforward measures $\mathbf{\Phi}_W$ and $\mathbf{Y}_W$.** Here, we recall the key concepts from [10].

For every graphon $W \in \mathcal{W}_0$ we define the set $\langle W\rangle$ as the set of all weak* limit points of sequences of versions of W. That is, a graphon $U \in \mathcal{W}_0$ belongs to $\langle W\rangle$ if and only if there are measure preserving bijections $\pi_1, \pi_2, \pi_3, \ldots$ of Ω such that the sequence $W^{\pi_1}, W^{\pi_2}, W^{\pi_3}, \ldots$ converges to U in the weak* topology. We call the set $\langle W\rangle$ the *envelope* of W.

We say that a graphon U is *at most as structured as a graphon* W if $\langle U\rangle \subset \langle W\rangle$. We write $U \preceq W$ in this case. We write $U \prec W$ if $U \preceq W$ but it does not hold that $W \preceq U$.

It follows directly from the definition of the weak* topology that the edge density of a weak* limit of a sequence of graphons equals to the the limit of the edge densities of the graphons in the sequence. Thus, we obtain the following.

Fact 2.5. *If two graphons have different edge densities then they are incomparable in the structuredness order.*

Given a graphon $W : \Omega^2 \to [0,1]$, we can define a pushforward probability measure on $[0,1]$ by

$$\mathbf{\Phi}_W(A) := \nu^{\otimes 2}\left(W^{-1}(A)\right) , \tag{2.7}$$

for every measurable set $A \subset [0,1]$. The measure $\mathbf{\Phi}_W$ gives us the distribution of the values of W. In [10], $\mathbf{\Phi}_W$ is called the *range frequencies of* W. Similarly, we can take the pushforward measure of the degrees, which is called the *degree frequencies of* W,

$$\mathbf{Y}_W(A) := \nu\left(\deg_W^{-1}(A)\right) , \tag{2.8}$$

for every measurable set $A \subset [0,1]$. The measures $\mathbf{\Phi}_W$ and $\mathbf{Y}_W$ provide substantial information about the graphon W. It is therefore natural to ask how these measures relate with respect to the structuredness order. To this end the following «flatness relation» on measures is introduced.

Definition 2.6. 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 $D \subset [0,1]$ we have

$$\int_{(x,y)\in D\times[0,1]} x \,\mathrm{d}\Psi = \int_{(x,y)\in D\times[0,1]} y \,\mathrm{d}\Psi . \tag{2.9}$$

We say that Λ_1 is *strictly flatter* than Λ_2 if Λ_1 is at least as flat as Λ_2 and $\Lambda_1 \neq \Lambda_2$.

We can now state the main result of Section 4.2 of [10].

Proposition 2.7. *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 $\mathbf{\Phi}_W$.*

3. Cut distance identifying graphon parameters

3.1. **Basics.** In the paper [10], we based our treatment of the cut distance on $\mathbf{ACC}_{w*}(W_1, W_2, W_3, \ldots)$ and $\mathbf{LIM}_{w*}(W_1, W_2, W_3, \ldots)$, which are sets of functions. In contrast, the key objects in [9] are the sets of numerical values

$$\left\{\mathrm{INT}_f(W) : W \in \mathbf{ACC}_{w*}(W_1, W_2, W_3, \ldots)\right\} \text{ and } \left\{\mathrm{INT}_f(W) : W \in \mathbf{LIM}_{w*}(W_1, W_2, W_3, \ldots)\right\},$$

with notation taken from (1.4). In this section, we introduce an abstract framework to approaching the cut distance via similar optimization problems. Our key definitions of cut distance identifying graphon parameters and cut distance compatible graphon parameters use $\mathbb{R}^n$ together with lexicographical ordering and Euclidean metric, and $\mathbb{R}^\mathbb{N}$ together with lexicographical ordering which we denote just $\leq$. Further, we use the following metric on $\mathbb{R}^\mathbb{N}$ (or $\mathbb{R}^n$). For $u, v \in \mathbb{R}^\mathbb{N}$ (or for $u, v \in \mathbb{R}^n$) we define $d_{\mathrm{lex}}(u, v)$ to be the smallest number $\varepsilon \geq 0$ such that $|u_i - v_i| \leq \varepsilon$ for every $1 \leq i < \frac{1}{\varepsilon}$. Note that the metric d_{lex} gives us the topology of pointwise convergence on $\mathbb{R}^\mathbb{N}$ (or $\mathbb{R}^n$).

By a *graphon parameter* we mean any function that either $\theta : \mathcal{W}_0 \to \mathbb{R}$, $\theta : \mathcal{W}_0 \to \mathbb{R}^n$ (for some $n \in \mathbb{N}$), or $\theta : \mathcal{W}_0 \to \mathbb{R}^\mathbb{N}$, 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 θ is a *cut distance identifying graphon parameter* if we have that $W_1 \prec W_2$ implies $\theta(W_1) < \theta(W_2)$ (here, by $<$ we understand the usual Euclidean order on $\mathbb{R}$ in case $\theta : \mathcal{W}_0 \to \mathbb{R}$ and the lexicographic order in case $\theta : \mathcal{W}_0 \to \mathbb{R}^n$ or $\theta : \mathcal{W}_0 \to \mathbb{R}^\mathbb{N}$). We say that a graphon parameter θ is a *cut distance compatible graphon parameter* if we have that $W_1 \preceq W_2$ implies $\theta(W_1) \leq \theta(W_2)$.

The following proposition provides a useful criterion for cut distance compatible graphon parameters. In this criterion, we restrict ourselves to L^1-continuous parameters (which is not a big restriction really).

Proposition 3.1. *Suppose that θ is a graphon parameter that is continuous with respect to the L^1 norm. Then θ is cut distance compatible if and only if for each graphon $W : \Omega^2 \to [0,1]$ and each finite partition $\mathcal{P}$ of Ω we have $\theta\left(W^{\Join \mathcal{P}}\right) \leq \theta(W)$.*

Proof. The $\Rightarrow$ direction is obvious. As for the reverse direction, suppose that θ is not cut distance compatible. That is, there exist two graphons $U \preceq W$ so that $\theta(U) > \theta(W)$. Since θ is L^1-continuous at U we can use Lemma 2.3 to find a finite partition $\mathcal{Q}$ such that

$$\theta\left(U^{\Join \mathcal{Q}}\right) > \theta(W). \tag{3.1}$$

As $U \preceq W$, there exist measure preserving bijections $\pi_1, \pi_2, \pi_3, \ldots$ so that $W^{\pi_n} \xrightarrow{w^*} U$. In particular, the sequence $\left((W^{\pi_n})^{\Join \mathcal{Q}}\right)_n$ converges to $U^{\Join \mathcal{Q}}$ in L_1. Thus the L_1-continuity of θ at $U^{\Join \mathcal{Q}}$ gives us that for some n, $\theta\left((W^{\pi_n})^{\Join \mathcal{Q}}\right)$ is nearly as big as $\theta\left(U^{\Join \mathcal{Q}}\right)$. In particular, using (3.1) we have that $\theta\left((W^{\pi_n})^{\Join \mathcal{Q}}\right) > \theta(W)$. We let π_n act on the partition $\mathcal{Q}$, $\mathcal{P} := \pi_n(\mathcal{Q})$. Obviously, $(W^{\pi_n})^{\Join \mathcal{Q}}$ is a version of $W^{\Join \mathcal{P}}$, and thus $\theta\left(W^{\Join \mathcal{P}}\right) = \theta\left((W^{\pi_n})^{\Join \mathcal{Q}}\right) > \theta(W)$, as was needed.

It is natural to believe that there is a similar characterization for cut distance identifying parameters. We were however unable to prove it, so we leave it as a conjecture. □

Conjecture 3.2. *Suppose that θ is a graphon parameter that is continuous with respect to the L^1 norm. Then θ is cut distance identifying if and only if for each graphon W and each finite partition $\mathcal{P}$ of Ω for which $W^{\Join\mathcal{P}} \neq W$ we have $\theta\left(W^{\Join\mathcal{P}}\right) < \theta\left(W\right)$.*

Note that the $\Rightarrow$ direction is obvious as in Proposition 3.1.

Cut distance identifying graphon parameters can be used to prove compactness of the graphon space. This is stated in the next two theorems.

Theorem 3.3. *Let $\Gamma_1, \Gamma_2, \Gamma_3, \ldots$ be a sequence of graphons. Suppose that θ is a cut distance compatible graphon parameter. Then there exists a subsequence $\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots$ such that $\mathbf{ACC}_{\mathrm{w}*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right)$ contains an element Γ with $\theta\left(\Gamma\right) = \sup\left\{\theta(W) : W \in \mathbf{ACC}_{\mathrm{w}*}\left(\Gamma_{n_1}, \Gamma_{n_2}, \Gamma_{n_3}, \ldots\right)\right\}$.*

Proof. This follows immediately from [10, Theorem 3.3] and [10, Lemma 4.7]. [d] □

Theorem 3.4. *Let $W_1, W_2, W_3, \ldots$ be a sequence of graphons. Suppose that θ is a cut distance identifying graphon parameter. Suppose that $\Gamma \in \mathbf{LIM}_{\mathrm{w}*}\left(W_1, W_2, W_3, \ldots\right)$ is such that*

$$\theta\left(\Gamma\right) = \sup\left\{\theta(W) : W \in \mathbf{ACC}_{\mathrm{w}*}\left(W_1, W_2, W_3, \ldots\right)\right\}.$$

Then $W_1, W_2, W_3, \ldots$ converges to Γ in the cut distance.

Proof. As a first step, we show that $\langle\Gamma\rangle = \mathbf{ACC}_{\mathrm{w}*}\left(W_1, W_2, \ldots\right) = \mathbf{LIM}_{\mathrm{w}*}\left(W_1, W_2, \ldots\right)$. Let $U \in \mathbf{ACC}_{\mathrm{w}*}\left(W_1, W_2, \ldots\right)$. By Theorem 3.3 from [10] we can find a subsequence $W_{n_1}, W_{n_2}, \ldots$ such that $\mathbf{LIM}_{\mathrm{w}*}\left(W_{n_1}, W_{n_2}, \ldots\right) = \mathbf{ACC}_{\mathrm{w}*}\left(W_{n_1}, W_{n_2}, \ldots\right)$ and $U \in \mathbf{LIM}_{\mathrm{w}*}\left(W_{n_1}, W_{n_2}, \ldots\right)$. Note that $\Gamma \in \mathbf{ACC}_{\mathrm{w}*}\left(W_{n_1}, W_{n_2}, \ldots\right)$. Using Lemma 4.7 from [10], we can find a maximal element $W \in \mathbf{LIM}_{\mathrm{w}*}\left(W_{n_1}, W_{n_2}, \ldots\right)$ with respect to the structuredness order. It follows from the definition of the structuredness order that $\Gamma \preceq W$ and therefore $\theta\left(\Gamma\right) \leq \theta\left(W\right)$. Using our assumption on Γ and the fact that θ is a cut distance identifying graphon parameter, we must have $\langle\Gamma\rangle = \langle W\rangle$. This implies that $U \in \langle W\rangle = \langle\Gamma\rangle \subseteq \mathbf{LIM}_{\mathrm{w}*}\left(W_1, W_2, \ldots\right)$.

We may suppose that $W_n \xrightarrow{\mathrm{w}^*} \Gamma$. To show that in fact $W_n \xrightarrow{\delta_\square} \Gamma$, we can mimic the proof of Theorem 3.5 (b) $\Longrightarrow$ (a) from [10]. □

So, while the concepts of cut distance identifying graphon parameters do not bring any new tools compared to the structuredness order, knowing that a particular parameter is cut distance identifying allows calculations that are often more direct than working with the structuredness order.

3.1.1. *Relation to quasi-randomness.* Recall that dense quasi-random finite graphs correspond to constant graphons. Thus, the key question in the area of quasi-randomness is which graphon parameters can be used to characterize constant graphons.[e]

The Chung–Graham–Wilson Theorem [3], a version of which we state below, provides the most classical parameters whose minimizers in $\mathcal{G}_p$ are is constant-p graphon.

---

[d]Let us note that an alternative direct proof of Theorem 3.3 can be repeated mutatis mutandis from Lemma 13 in [9]. This latter proof is more elementary and does not need transfinite induction or any appeal to the Vietoris topology.

[e]Strictly speaking, only parameters that are continuous with respect to the cut distance are relevant for characterizing sequences of quasi-random graphs. Indeed, the assumption of continuity is used to transfer between finite graphs and their limits. The two main parameters we treat below — subgraph densities $t(H, \cdot)$ and spectrum — are indeed well-known to be cut distance continuous (see Theorems 11.52 and 11.53 in [18]). The parameter $\mathrm{INT}_f(\cdot)$ is not cut distance continuous, and hence does not admit such a transference.

Theorem 3.5. *Let $p \in [0,1]$. Then the constant-p graphon is the only graphon U in the family $\mathcal{G}_p$ satisfying any of the following conditions.*

(a) *We have $t(C_{2\ell}, U) \leq p^{2\ell}$ for a fixed $\ell \in \{2,3,4,\ldots\}$.*
(b) *The largest eigenvalue of U is at most p and all other eigenvalues are zero.*

Such characterizations of quasi-randomness fit very nicely our framework of cut distance identifying graphon parameters. Indeed, constant graphons are exactly the minimal elements in the structuredness order; we refer to [10, Proposition 7.5] for an easy proof. Thus, each cut distance identifying graphon parameter can be used to characterize constant graphons.

In the opposite direction, we show in Sections 3.4 and 3.5 that the graphon parameters considered in Theorem 3.5 are actually cut distance identifying. Such a strengthening is not automatic (even for reasonable graphon parameters); for example the parameter $t(C_4^+, \cdot)$ (here, C_4^+ is a 4-cycle with a pendant edge) is shown in [15, Section 2] to be minimized on constant graphons but not to be cut distance identifying.[f]

3.2. **Revising the parameter** $\mathrm{INT}_f(\cdot)$**.** Recall that in [9], the parameter $\mathrm{INT}_f(\cdot)$ (for a strictly convex continuous function $f:[0,1]\to\mathbb{R}$) was used to identify cut distance limits of graphons (thus providing a new proof of Theorem 1.1). One of the key steps in [9] was to show that a certain refinement of a graphon leads to an increase of $\mathrm{INT}_f(\cdot)$. While not approached this way in [9], this hints that $\mathrm{INT}_f(\cdot)$ is cut distance identifying. We prove this statement in the current section, as a quick application of the results from [10, Section 4.2]. Also, here we show that the requirement of continuity of f was just an artifact of the proof in [9].

Theorem 3.6. (a) *Suppose that $f:[0,1]\to\mathbb{R}$ is a convex function. Then $\mathrm{INT}_f(\cdot)$ is cut distance compatible.*
(b) *Suppose that $f:[0,1]\to\mathbb{R}$ is a strictly convex function. Then $\mathrm{INT}_f(\cdot)$ is cut distance identifying.*

Proof of Part (a). Suppose that $f:[0,1]\to\mathbb{R}$ is a convex function. Recall that every convex function admits left and right derivatives which are both increasing functions. The key is to observe that for a graphon Γ, we have $\mathrm{INT}_f(\Gamma) = \int_{x\in[0,1]} f(x)\,\mathrm{d}\mathbf{\Phi}_\Gamma$, where $\mathbf{\Phi}_\Gamma$ is defined by (2.7). So suppose that $U \preceq W$. By Proposition 2.7, we have that $\mathbf{\Phi}_U$ is at least as flat as $\mathbf{\Phi}_W$. Let Λ be a measure on $[0,1]^2$ as in Definition 2.6 that witnesses this fact. If Λ is carried by the diagonal of $[0,1]^2$ then $\mathbf{\Phi}_U = \mathbf{\Phi}_W$. In that case $U \not\prec W$ by Proposition 2.7. In other words, $\langle U\rangle = \langle W\rangle$. By Corollary 4.22 of [10], we have $\delta_\square(U,W)=0$. Since θ is a graphon parameter, we conclude that $\theta(U)=\theta(W)$. It remains to consider the case when Λ is not carried by the diagonal. Then there are intervals $[a,b],[c,d]\subseteq[0,1]$ with $\Lambda([a,b]\times[c,d])>0$ and $b<c$ (the other case when $d<a$ is similar). For every $y\in[c,d]$ we have

$$f(y) \geq f'_+(b)\cdot y + \left(f(b) - f'_+(b)\cdot b\right) . \tag{3.2}$$

Fix $\varepsilon>0$ arbitrarily and note that f is continuous on the open interval $(0,1)$ by convexity, thus the points 0 and 1 are the only possible points of discontinuity of f. So for every $x\in(0,1)$ there is an interval $J_x\subset(0,1)$ containing x such that every two values of f on J_x differ by at most ε. Take a covering of $(0,1)$ consisting of at most countably many such intervals, add the singletons $\{0\}$ and $\{1\}$, and then refine the resulting family to a countable disjoint covering

[f]See Remark 3.20 for a more general result.

$\{J_1, J_2, \ldots\}$ of $[0,1]$. Then for every i and for every $x \in J_i$ we have $|f(x) - f(x_i)| \leq \varepsilon$ where x_i is the $\mathbf{\Phi}_U$-mean value of x on J_i, i.e. (by (2.9))

$$x_i = \frac{1}{\mathbf{\Phi}_U(J_i)} \int_{J_i} x \,\mathrm{d}\mathbf{\Phi}_U = \frac{1}{\Lambda(J_i \times [0,1])} \int_{J_i \times [0,1]} x \,\mathrm{d}\Lambda = \frac{1}{\Lambda(J_i \times [0,1])} \int_{J_i \times [0,1]} y \,\mathrm{d}\Lambda \tag{3.3}$$

(if for some i we have $\Phi_U(J_i) = 0$ then we can define x_i to be an arbitrary element of J_i). We may moreover assume that for every i either $J_i \subseteq [a,b]$ or $J_i \cap [a,b] = \emptyset$, then $x_i \in [a,b]$ whenever $J_i \subseteq [a,b]$. Note that convexity of f together with equation (3.2) imply that

$$f(y) \geq f'_+(x_i) \cdot y + (f(x_i) - f'_+(x_i) \cdot x_i) \tag{3.4}$$

for every $y \in [c,d]$ and every i with $J_i \subseteq [a,b]$.

We have

$$\begin{aligned}
\mathrm{INT}_f(U) &= \int_{x \in [0,1]} f(x) \,\mathrm{d}\mathbf{\Phi}_U \\
&= \sum_i \int_{x \in J_i} f(x) \,\mathrm{d}\mathbf{\Phi}_U \\
&\overset{\varepsilon}{\approx} \sum_i f(x_i) \mathbf{\Phi}_U(J_i) \\
&= \sum_i f(x_i) \Lambda(J_i \times [0,1]) \\
\boxed{\text{Jensen's inequality and (3.3)}} \quad &\leq \sum_i \int_{(x,y) \in J_i \times [0,1]} f(y) \,\mathrm{d}\Lambda \\
&= \int_{(x,y) \in [0,1]^2} f(y) \,\mathrm{d}\Lambda \\
&= \int_{y \in [0,1]} f(y) \,\mathrm{d}\mathbf{\Phi}_W \\
&= \mathrm{INT}_f(W) \ .
\end{aligned}$$

As this is true for every $\varepsilon > 0$ we conclude that $\mathrm{INT}_f(U) \leq \mathrm{INT}_f(W)$.

Proof of Part (b): Now suppose that $f : [0,1] \to \mathbb{R}$ is strictly convex and that $U \prec W$ (then Φ_U is strictly flatter than Φ_W, and so the witnessing measure Λ cannot be carried by the diagonal of $[0,1]^2$). In that case both one-sided derivatives of f are strictly increasing, and so equation (3.2) can be strengthen to

$$f(y) \geq f'_+(b) \cdot y + (f(b) - f'_+(b) \cdot b) + \delta \tag{3.5}$$

for $y \in [c,d]$, for some $\delta > 0$. Equation (3.4) then also holds in the stronger form

$$f(y) \geq f'_+(x_i) \cdot y + (f(x_i) - f'_+(x_i) \cdot x_i) + \delta \tag{3.6}$$

for every $y \in [c,d]$ and every i with $J_i \subseteq [a,b]$. We show that then the application of Jensen's inequality above ensures that $\mathrm{INT}_f(U) < \mathrm{INT}_f(W)$. To this end it suffices to show that there is a constant $K > 0$ not depending on ε such that

$$\sum_{i:\ J_i \subseteq [a,b]} f(x_i) \Lambda(J_i \times [0,1]) \leq \sum_{i:\ J_i \subseteq [a,b]} \int_{(x,y) \in J_i \times [0,1]} f(y) \,\mathrm{d}\Lambda - K \ .$$

For every i denote $g_i(y) := f'_+(x_i)\cdot y + (f(x_i) - f'_+(x_i)\cdot x_i)$. Then we have

$$\begin{aligned}
&\sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times[0,1]} f(y)\,\mathrm{d}\Lambda \\
=\ &\sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times[c,d]} f(y)\,\mathrm{d}\Lambda + \sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times([0,1]\setminus[c,d])} f(y)\,\mathrm{d}\Lambda \\
\overset{\text{(3.6) and convexity}}{\geq}\ &\sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times[c,d]} (g_i(y)+\delta)\,\mathrm{d}\Lambda + \sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times([0,1]\setminus[c,d])} g_i(y)\,\mathrm{d}\Lambda \\
=\ &\sum_{i\colon J_i\subseteq[a,b]}\int_{(x,y)\in J_i\times[0,1]} g_i(y)\,\mathrm{d}\Lambda + \delta\cdot\Lambda([a,b]\times[c,d]) \\
\overset{(3.3)}{=}\ &\sum_{i\colon J_i\subseteq[a,b]} f(x_i)\Lambda(J_i\times[0,1]) + \delta\cdot\Lambda([a,b]\times[c,d])\ .
\end{aligned}$$

So it suffices to set $K := \delta\cdot\Lambda([a,b]\times[c,d])$. □

For a later reference, let us apply Theorem 3.6 to the strictly convex function $x\mapsto x^2$, for which $\mathrm{INT}_{x\mapsto x^2}(\cdot) = \|\cdot\|_2^2$.

Corollary 3.7. *Suppose that U and W are two graphons with $U\prec W$. Then $\|U\|_2 < \|W\|_2$.*

3.3. **Convex graphon parameters.** In Definition 3.8 we introduce convex graphon parameters. In Theorem 3.9 we prove that such parameters are cut distance compatible if they are also L_1-continuous. In Remark 3.11 we observe that the opposite implication is not true.

Definition 3.8. A graphon parameter $f:\mathcal{W}_0\to\mathbb{R}$ is *convex* if for every $\alpha_1,\alpha_2,\alpha_3,\ldots,\alpha_k\in[0,1]$ with $\sum_i\alpha_i = 1$ and every graphons $W, W_1, W_2,\ldots,W_k\in\mathcal{W}_0$ with $W=\sum_i\alpha_iW_i$ we have $f(W)\leq\sum_i\alpha_i f(W_i)$.

Theorem 3.9. *Let $f:\mathcal{W}_0\to\mathbb{R}$, $f:\mathcal{W}_0\to\mathbb{R}^n$, or $f:\mathcal{W}_0\to\mathbb{R}^{\mathbb{N}}$ be a graphon parameter that is convex and continuous in L_1. Then f is cut distance compatible.*

Theorem 3.9 can be used to give another proof of the first part of Theorem 3.6 under the additional assumption that the convex function $f:[0,1]\to\mathbb{R}$ is continuous. (Note that this is not a very much restrictive assumption as the only possible discontinuities of any convex function defined on a closed interval are the endpoints of the interval.) Indeed, the continuity of f easily implies that the graphon parameter INT_f is continuous in L^1, and the the convexity of INT_f is also clear.

The rest of this section is devoted to the proof of Theorem 3.9. The crux of the proof is the following lemma.

Lemma 3.10. *Let U, W be two graphons such that $U\prec W$. Then for any $\varepsilon>0$ there exist measure preserving bijections $\pi_1,\pi_2,\ldots,\pi_n$ and a convex combination $\sum_{i=1}^n\alpha_iW^{\pi_i}$ such that*

$$\left\|\sum_{i=1}^n\alpha_iW^{\pi_i} - U\right\|_1 < \varepsilon.$$

Proof. Find measure preserving bijections $\pi_1,\pi_2,\pi_3,\ldots$ of Ω such that $W^{\pi_i}\xrightarrow{\mathrm{w}^*}U$. As the weak topology on $\mathcal{W}_0$ is weaker than the weak* topology, we have $W^{\pi_i}\xrightarrow{w}U$ as well. Therefore $U\in\overline{\mathrm{conv}\{W^{\pi_1},W^{\pi_2},W^{\pi_3},\ldots\}}^{w}$. But in any Banach space, the weak closure of any convex set coincides with its norm closure. So $U \in \overline{\text{conv}\{W^{\pi_1}, W^{\pi_2}, W^{\pi_3}, \dots\}}^{L^1}$ and the result follows. □

Now we are ready to prove Theorem 3.9.

Proof of Theorem 3.9. First, suppose that f is convex. Suppose that $U, W : \Omega^2 \to [0,1]$ are arbitrary graphons such that $U \prec W$. Suppose that $\varepsilon > 0$ is arbitrary. Let $\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}$ be the convex combination given by Lemma 3.10 for U, W and error ε. Then, we have

$$\begin{aligned} f(U) &= f\left(\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}\right) + \left(f(U) - f\left(\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}\right)\right) \\ \boxed{\text{convexity}} \quad &\le \sum_{i=1}^{n(\varepsilon)} \alpha_i f\left(W^{\pi_{\varepsilon,i}}\right) + \left(f(U) - f\left(\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}\right)\right) \\ &= f(W) + \left(f(U) - f\left(\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}\right)\right) . \end{aligned} \tag{3.7}$$

Now, as ε goes to 0, the graphon $\sum_{i=1}^{n(\varepsilon)} \alpha_i W^{\pi_{\varepsilon,i}}$ goes to U in $L^1(\Omega^2)$. Thus, the L^1-continuity of f tells us that the last term in (3.7) vanishes, and thus $f(U) \le f(W)$. Thus f is cut distance compatible. □

Remark 3.11. *In this example we first construct two graphons U and V such that V is a convex combination of versions of U but $V \not\preceq U$. We then use this to construct a cut distance compatible graphon parameter f^* that is not convex.The graphons U and V are shown in Figure 3.1. The graphon U is defined as $U(x,y) = 1$ if and only if $(x,y) \in [0, \frac{1}{2}]^2$ and $U(x,y) = 0$ otherwise, while $V(x,y) = \frac{1}{2}$ if and only if $(x,y) \in [0,\frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2$ and $V(x,y) = 0$ otherwise. If we set $\varphi(x) = 1 - x$, then clearly $V = \frac{U + U^\varphi}{2}$. Let us now argue that $V \not\preceq U$. For any measure preserving bijection π we have $\int_{[0,\frac{1}{2}]\times[\frac{1}{2},1]} U^\pi = \nu\left(\pi\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) \cdot \nu\left(\pi\left([0,\frac{1}{2}]\right) \cap [\frac{1}{2},1]\right)$. Thus, for any sequence of measure preserving bijections $\pi_1, \pi_2, \dots$ such that (after passing to a subsequence if necessary) we have that either $\nu\left(\pi_n\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) \to 0$ or $\nu\left(\pi_n\left([0,\frac{1}{2}]\right) \cap [0,\frac{1}{2}]\right) \to \frac{1}{2}$. We conclude that $U^{\pi_1}, U^{\pi_2}, \dots \overset{w^*}{\not\longrightarrow} V$.*

Now, take any cut distance compatible parameter f and suppose that it is convex. In particular, we have that $\frac{1}{2}f(U) + \frac{1}{2}f(U^\varphi) \ge f(V)$ for the two graphons U and V defined above. We can now define

$$f^*(W) = f(W) + \left(\frac{1}{2}f(U) + \frac{1}{2}f(U^\varphi) - f(V) + 1\right)$$

for each graphon W such that $W \succeq V$ and

$$f^*(W) = f(W)$$

otherwise. The graphon parameter f^* is clearly cut distance compatible, but no longer convex, since

$$f^*(V) = \frac{1}{2}f(U) + \frac{1}{2}f(U^\varphi) + 1 > \frac{1}{2}f^*(U) + \frac{1}{2}f^*(U^\varphi) .$$

This example works even if we restrict ourselves to graphons lying in an envelope of a certain fixed graphon W, since if we set $W(x,y) = 1$ if and only if $(x,y) \in [0,\frac{1}{4}]^2 \cup [\frac{1}{4},\frac{1}{2}]^2$ and $W(x,y) = 0$ otherwise, and set $U' = \frac{U}{2}, V' = \frac{V}{2}$, then we have three graphons U', V', W such that $U', V' \preceq W$, $V' = \frac{U' + U'^\varphi}{2}$, but $V' \not\preceq U'$.

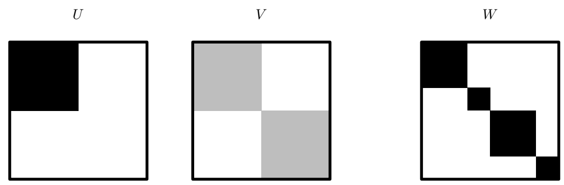

FIGURE 3.1. Graphons U and V from Example 3.11

The function f^* from Example 3.11 is, however, very unnatural since it is not continuous with respect to L_1. We leave it as an open problem, whether there is a continuous example.

Problem 3.12. Is there a function $f : \mathcal{W}_0 \to \mathbb{R}$ that is is not convex, but is continuous in L_1 and compatible with the structuredness order?

In the subsequent Section 3.5 we manage to partially answer this problem by showing that for homomorphism densities, which are an important class of functions defined on the space of graphons (and continuous in L_1), we can indeed reverse Theorem 3.9 and get that compatibility with structuredness order implies that the respective function is convex.

3.4. **Spectrum.** We will prove in Theorem 3.14 that the structuredness order is a suborder of the spectral quasiorder defined in Section 2.2.3. But first we need an easy lemma.

Lemma 3.13. *Let $(W_n)_n$ and U be graphons on Ω^2 such that $W_n \xrightarrow{\text{w}^*} U$. Let $u, v \in L^2(\Omega)$. Then we have $\langle W_n u, v\rangle \to \langle Uu, v\rangle$.*

Proof. Since step functions are dense in $L^2(\Omega)$, and since the forms $\langle W_n\cdot, \cdot\rangle$ and $\langle U\cdot, \cdot\rangle$ are obviously bilinear, it suffices to prove the statement for indicator functions of sets, $u = \mathbf{1}_A$, $v = \mathbf{1}_B$ (where $A, B \subset \Omega$). But in that case $\langle W_n u, v\rangle = \int_{A\times B} W_n$ and $\langle Uu, v\rangle = \int_{A\times B} U$. The statement follows since $W_n \xrightarrow{\text{w}^*} U$. □

We are now ready to prove the main result of this section. Let us note that the arguments that we use to prove this result also turned out to be useful in the setting of finitely forcible graphs; in particular Král', Lovász, Noel, and Sosnovec [14], use our arguments in a final step of their proof that for each graphon and each $\epsilon > 0$, there exists a forcible graphon that differs from the original one on a set of measure at most ϵ.

Theorem 3.14. *If $U \prec W$, then $U \overset{S}{\prec} W$.*

Proof. Consider the sequence of versions $W^{\pi_n} \xrightarrow{\text{w}^*} U$. Let $\lambda_1^+ \geq \lambda_2^+ \geq \lambda_3^+ \geq \ldots \geq 0$ be the positive eigenvalues of U with associated pairwise orthogonal unit eigenvectors $u_1, u_2, u_3, \ldots$, and let $\beta_1^+ \geq \beta_2^+ \geq \beta_3^+ \geq \ldots \geq 0$ be the positive eigenvalues of W. First, we will prove that for any given $\varepsilon > 0$ and k, we have $\beta_k^+ \geq \lambda_k^+ - \varepsilon$. By the maxmin characterization of eigenvalues, we have

$$\beta_k^+ = \max_{\substack{H \text{ subspace of } L^2(\Omega)\\ \dim(H)=k}} \min_{\substack{g\in H\\ \|g\|_2=1}} \langle Wg, g\rangle \ .$$

Fix the space $\widetilde{H} = \operatorname{span}\left\{u_1^{\pi_n^{-1}}, u_2^{\pi_n^{-1}}, \dots, u_k^{\pi_n^{-1}}\right\}$, where $u_i^{\pi_n^{-1}}(x) = u_i(\pi_n^{-1}(x))$. Then, we have

$$\beta_k^+ \geq \min_{\substack{g \in \widetilde{H} \\ \|g\|_2 = 1}} \langle Wg, g \rangle . \tag{3.8}$$

Furthermore, by Lemma 3.13 we can find n large enough so that for all $i, j \leq k$ we have

$$\left| \langle W^{\pi_n} u_i, u_j \rangle - \langle U u_i, u_j \rangle \right| < \frac{\varepsilon}{k^2}.$$

Now, for $g \in \widetilde{H}$ that realizes the minimum in (3.8), we can write its orthogonal decomposition as $g = \sum_{i=1}^k c_i u_i^{\pi_n^{-1}}$, where $\sum_{i=1}^k c_i^2 = 1$. Thus, we obtain

$$\begin{aligned}
\langle Wg, g \rangle &= \langle W^{\pi_n} g^{\pi_n}, g^{\pi_n} \rangle \\
&= \left\langle W^{\pi_n} \sum_{i=1}^k c_i u_i, \sum_{i=1}^k c_i u_i \right\rangle \\
&= \sum_{i,j=1}^k c_i c_j \langle W^{\pi_n} u_i, u_j \rangle \\
&> \sum_{i=1}^k c_i^2 \left(\langle U u_i, u_i \rangle - \frac{\varepsilon}{k^2} \right) + \sum_{\substack{i,j=1 \\ i \neq j}}^k c_i c_j \left(\langle U u_i, u_j \rangle - \frac{\varepsilon}{k^2} \right) \\
&= \sum_{i=1}^k c_i^2 \left(\lambda_i^+ - \frac{\varepsilon}{k^2} \right) - \sum_{\substack{i,j=1 \\ i \neq j}}^k c_i c_j \frac{\varepsilon}{k^2} \\
&\geq \lambda_k^+ - \varepsilon.
\end{aligned}$$

Thus, by equation (3.8) we have $\beta_k^+ \geq \lambda_k^+ - \varepsilon$.

A similar argument can be used for the negative eigenvalues $\lambda_1^- \leq \lambda_2^- \leq \lambda_3^- \leq \dots \leq 0$ of U and $\beta_1^- \leq \beta_2^- \leq \beta_3^- \leq \dots \leq 0$ of W to show that $\beta_k^- \leq \lambda_k^- + \varepsilon$. That implies $U \overset{S}{\preceq} W$. To show that the inequality is strict for at least one eigenvalue, assume by contradiction that the eigenvalues of U and W are all the same. Then a double application of (2.3) gives

$$\|W\|_2^2 = \sum \left(\beta_i^+\right)^2 + \sum \left(\beta_i^-\right)^2 = \sum \left(\lambda_i^+\right)^2 + \sum \left(\lambda_i^-\right)^2 = \|U\|_2^2 .$$

But this is a contradiction with Corollary 3.7. This finishes the proof. ☐

Corollary 3.15. *The graphon parameter $\theta(\cdot) : \mathcal{W}_0 \to \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$ defined by*[g]

$$\theta(\cdot) = \left(\left(\lambda_1^+(\cdot), \lambda_2^+(\cdot), , \dots \right), \left(\left|\lambda_1^-(\cdot)\right|, \left|\lambda_2^-(\cdot)\right|, , \dots \right) \right)$$

is cut distance identifying graphon.

[g]The codomain of this graphon parameter is not of the form $\mathbb{R}^n$ or $\mathbb{R}^{\mathbb{N}}$ as required in our definition in Section 3.1. The right way to extend the lexicographic order to this setting is to say that $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$ is *at less than or equal to* $(\mathbf{c}, \mathbf{d}) \in \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$ if $\mathbf{a}$ is less than or equal to $\mathbf{b}$ in the lexicographic order and $\mathbf{c}$ is less than or equal to $\mathbf{d}$ in the lexicographic order. Actually, our Theorem 3.14 is strong enough that the lexicographic order can be replaced by the pointwise order.

3.5. **Subgraph densities.** In this section, we address the following problem.

Problem 3.16. Characterize graphs H for which $t(H,\cdot):\mathcal{W}_0\to\mathbb{R}$ is a cut distance compatible (respectively a cut distance identifying) graphon parameter.

Observe that thanks to Proposition 3.1, for the case of compatible graphon parameters, Problem 3.16 reduces to characterizing graphs H for which we have

$$t\left(H,W^{\ltimes\mathcal{P}}\right)\le t\left(H,W\right)\text{ for each }W\in\mathcal{W}_0\text{ and each finite partition }\mathcal{P}. \tag{3.9}$$

Similarly, if true, our Conjecture 3.2 implies that for the case of identifying graphon parameters, Problem 3.16 reduces to characterizing graphs H for which we have

$$t\left(H,W^{\ltimes\mathcal{P}}\right)<t\left(H,W\right)\text{ for each }W\in\mathcal{W}_0\text{ and each finite partition }\mathcal{P}\text{ for which }W\neq W^{\ltimes\mathcal{P}}. \tag{3.10}$$

This is closely related to Sidorenko's conjecture (which was asked independently by Erdős and Simonovits, and by Sidorenko, [22, 23]) and the Forcing conjecture (first hinted in [24, Section 5]). Indeed, these conjectures — when stated in the language of graphons — ask to characterize graphs H for which we have

$$t\left(H,W^{\ltimes\{\Omega\}}\right)\le t\left(H,W\right)\text{ for each }W\in\mathcal{W}_0 \tag{3.11}$$

(Sidorenko's conjecture), and

$$t\left(H,W^{\ltimes\{\Omega\}}\right)<t\left(H,W\right)\text{ for each nonconstant }W\in\mathcal{W}_0 \tag{3.12}$$

(Forcing conjecture).

Recall that Sidorenko's conjecture asserts that H satisfies (3.11) if and only if H is bipartite. Similarly, the Forcing conjecture asserts that H satisfies (3.12) if and only if H is bipartite and contains a cycle. In both cases, the $\Rightarrow$ direction is easy. Let us recall that the reason why at least one cycle is required for the Forcing conjecture is that the density of any forest H in any p-regular graphon (whether constant-p, or not) is $p^{e(H)}$. The other direction in both conjectures is open, despite being known in many special cases, see [6, 17, 13, 4, 12, 16, 25, 5].

Because all the properties we investigate in this section strengthen (3.11), we are concerned only with bipartite graphs throughout. The only exception is Remark 3.21 which addresses a possible «converse» definition of cut distance identifying properties.

Graphs satisfying (3.9) were investigated in [15] where these graphs are said to *have the step Sidorenko property*. Similarly, graphs satisfying (3.10) are said to *have the step forcing property*. Clearly, these properties imply (3.11) and (3.12), respectively. These stronger «step» properties do not follow automatically from (3.11) and (3.12); in [15, Section 2] it is shown that the 4-cycle with a pendant edge C_4^+ has the Sidorenko property but not the step Sidorenko property. Thus, every graph having the step Sidorenko property must be bipartite and every graph having the step forcing property must be bipartite with a cycle. The focus of [15] was in providing negative examples. For example, it was shown in [15] that a Cartesian product of cycles does not have the step Sidorenko property, unless all the cycles have length 4.

The connection to our running Problem 3.16 comes from Proposition 14.13 of [18] which implies that each weakly norming graph has the step Sidorenko property (it also directly follows from Theorem 3.9).

Corollary 3.17. *For each weakly norming graph H the function $t(H, \cdot)$ is cut distance compatible.*

In this section we show that connected graphs with the step Sidorenko property are exactly the weakly norming graphs (thus answering a question of Král', Martins, Pach and Wrochna [15]). To this end, we at first need to recall several definitions, in particular the equivalent characterization of weakly norming graphs from [12].

Theorem 3.18 (Theorem 2.8 in [12])**.** *A graph is seminorming if and only if it has the Hölder property. It is weakly norming if and only if it has the weak Hölder property.*

We are now ready to show the converse of Corollary 3.17 for connected graphs.

Theorem 3.19. *Suppose that H is a connected graph. If the function $t(H, \cdot)$ is cut distance compatible (or, equivalently, if H has the step Sidorenko property), then H is weakly Hölder.*

Remark 3.20. *Two nontrivial necessary conditions are established for a graph H being weakly Hölder are established in* [12, Theorem 2.10]*. One of them basically says that H does not contain a subgraph dense than itself. The other condition says that if $V(H) = A_1 \sqcup A_2$ is a bipartition of H and $u, v \in A_i$ are two vertices from the same part, then $\deg(u) = \deg(v)$. Thus, Theorem 3.19 restricts quite substantially the class graphs having the step Sidorenko property, compared to the class of all bipartite graphs which are conjectured to have the Sidorenko property. In particular, we see directly that C_4^+ does not have the step Sidorenko property.*

Before showing the proof of Theorem 3.19, we summarize the situation for weakly norming graphs in the upper part of the Figure 3.2. The notion of weakly norming graphs was introduced in [12] together with the proof of equivalence with the notion of weakly Hölder graphs. The step Sidorenko property was introduced in [15]. The authors used Proposition 14.3 from [18] that guarantees that all weakly norming graphs have the step Sidorenko property. Our notion of cut distance compatible parameters is basically a rewording of the step Sidorenko property in the language of the structuredness order (see Proposition 3.1). Finally, we now present a proof that for connected graphs the notion of compatibility implies the weak Hölder property, thus closing a circle for connected graphs. We do not regard disconnected graphs. We now return to the proof of Theorem 3.19.

Proof of Theorem 3.19. Suppose that H has m edges and n vertices. We prove that H is weakly Hölder. By Theorem 3.18 we already know that all weakly Hölder graphs are exactly weakly norming graphs. We divide the proof of the theorem into two parts, at first we prove that $t(H, \cdot)$ is subadditive up to a constant loss, specifically, we show that

$$t(H,U)^{1/m} + t(H,V)^{1/m} \geq \frac{1}{4} \cdot t(H,U+V)^{1/m} . \tag{3.13}$$

Then we use this inequality to prove that H is weakly Hölder using the tensoring technique in the same way as it is used in the proof of Theorem 3.18 from [12].

Let U and V be two arbitrary graphons and let W_1 be a graphon containing a copy of U scaled by a factor of one half in its top-left corner (i.e., $W_1(x,y) = U(2x, 2y)$ for $(x,y) \in [0, \frac{1}{2}]^2$), a copy of V in its bottom-right corner (i.e., $W_1(x,y) = V\left(2(x-\frac{1}{2}), 2(y-\frac{1}{2})\right)$ for $(x,y) \in [\frac{1}{2}, 1]^2$), and zero otherwise (see Figure 3.3). Note that for the homomorphism density $t(H, W_1)$ we have

$$t(H, W_1) = \frac{t(H,U) + t(H,V)}{2^n} .$$

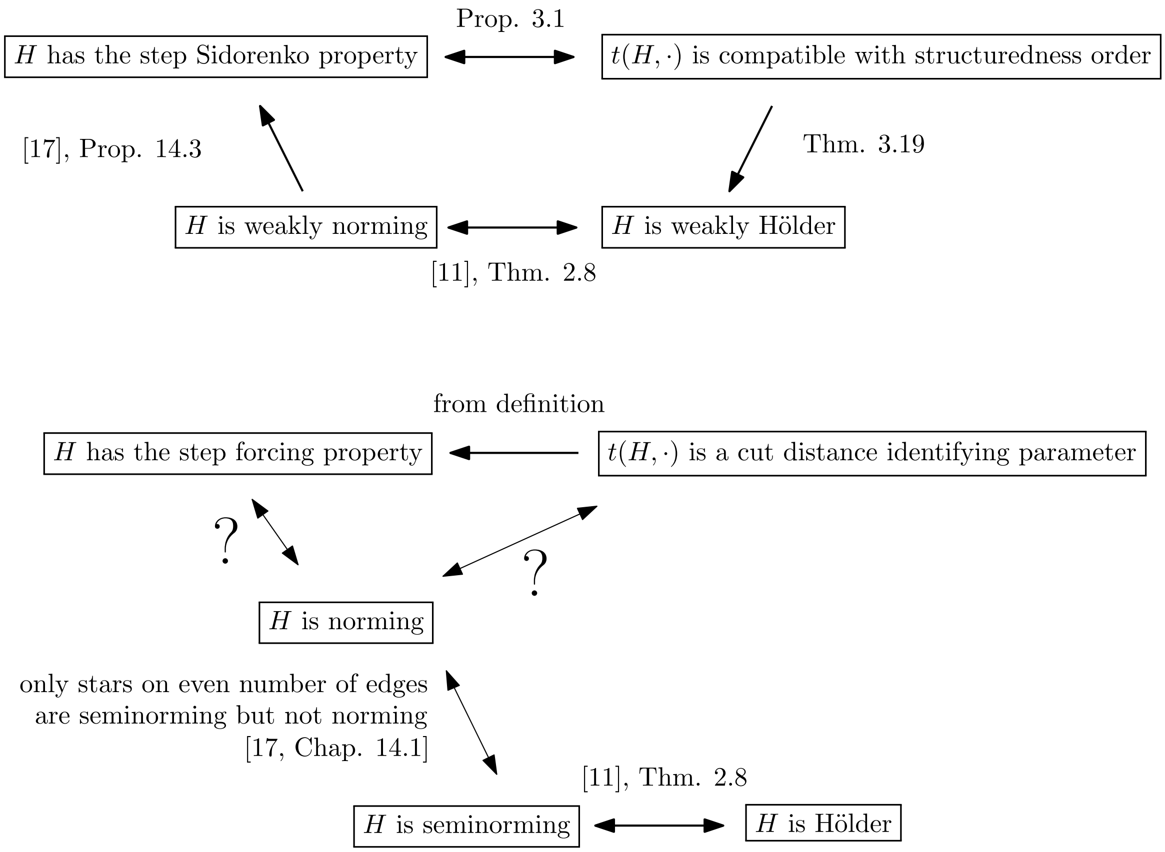

FIGURE 3.2. Diagram of notions used in this paper and their relations. Question marks suggest unproven relationships. Theorem 3.19 holds only for connected graphs.

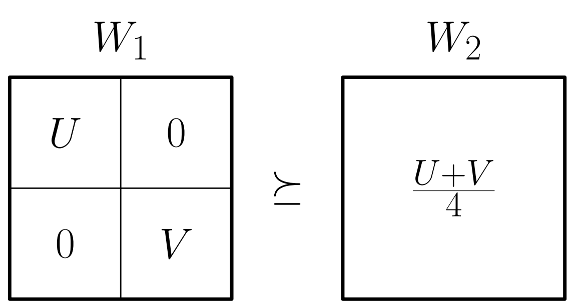

FIGURE 3.3. Graphons U, V, W_1 and W_2 from the proof of Theorem 3.19

This is because H is connected and, thus, homomorphisms that map nonzero number of vertices of H to $[0, \frac{1}{2}]$, and nonzero number of vertices to $[\frac{1}{2}, 1]$ do not contribute to the value of the integral $t(H, W_1)$. Now consider the graphon $W_2 = \frac{U+V}{4}$. Certainly, $W_1 \succeq W_2$, as can be certified by a sequence of measure preserving almost-bijections $\varphi_1, \varphi_2, \ldots$, defined as $\varphi_n(x) = \frac{\lfloor 2nx \rfloor}{2n} + x$ for $0 \le x \le \frac{1}{2}$ and $\varphi_n(x) = \frac{\lfloor 2nx \rfloor - 2n + 1}{2n} + x$ for $\frac{1}{2} < x \le 1$, that interlace the two intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ and thus the two copies U and V in W_1 together. We get that $t(H, W_1) \geq t(H, W_2)$. Observe that $t(H, W_2) = t\left(H, \frac{U+V}{4}\right) = \frac{t(H,U+V)}{4^m}$, hence we get

$$\frac{t(H,U) + t(H,V)}{2^n} \geq \frac{t(H,U+V)}{4^m} .$$

We are actually interested in the quantity $t(H,U)^{1/m}$, so we rewrite this as

$$(t(H,U) + t(H,V))^{1/m} \geq \frac{2^{n/m}}{4} \cdot t(H,U+V)^{1/m} \geq \frac{1}{4} \cdot t(H,U+V)^{1/m} .$$

Finally note that $t(H,U)^{1/m} + t(H,V)^{1/m} \geq (t(H,U) + t(H,V))^{1/m}$, as can be verified by raising the inequality to the m-th power. This yields the desired inequality (3.13).

Now we merely replicate the proof from [12] that all weakly norming graphs are weakly Hölder (see also [18], Theorem 14.1). At first note that the inequality (3.13) can be inductively generalised to yield that for a sequence of graphons $U_1, \ldots, U_\ell$ we have

$$\sum_{i=1}^{\ell} t(H, U_i)^{1/m} \geq \left(\frac{1}{4}\right)^{\ell-1} \cdot t\left(H, \sum_{i=1}^{\ell} U_i\right)^{1/m} . \tag{3.14}$$

Now let (H, w) be a $\mathcal{W}^+$-decoration of H. By Remark 2.4 we may assume that $t(H, W_e) = 1$ for every W_e. We want to prove that $t(H, w) \leq 1$, but at first we prove a weaker inequality $t(H, w) \leq 4^{m(m-1)} \cdot m^m$. Indeed, we have

$$t(H,w) \leq t\left(H, \sum_{e \in E(H)} W_e\right) \leq \left(4^{m-1} \cdot \sum_{e \in E(H)} t(H, W_e)^{1/m}\right)^m = 4^{m(m-1)} \cdot m^m , \tag{3.15}$$

where in the first inequality we replaced each W_e by $\sum_{e \in E(H)} W_e$, while the second inequality is due to the bound (3.14). Now suppose that we decorate each edge of H by $W_e^{\otimes k}$ for $k \geq 1$. As we observed in (2.2), we then have $t\left(H, w^{\otimes k}\right) = t(H, w)^k$ and $t\left(H, W_e^{\otimes k}\right) = t\left(H, W_e\right)^k = 1$. Thus the inequality (3.15) gives that $t(H, w)^k = t(H, w^{\otimes k}) \leq 4^{m(m-1)} \cdot m^m$, thus $t(H, w) \leq \left(4^{m(m-1)} \cdot m^m\right)^{1/k}$. Since this holds for any $k \geq 1$, we conclude that $t(H, w) \leq 1$. □

Remark 3.21. *Note that the definition of cut distance compatible (resp. identifying) parameters given at the beginning of Section 3.1 was somewhat arbitrary. That is, instead of requiring that $W_1 \preceq W_2$ implies $\theta(W_1) \leq \theta(W_2)$ (resp. that $W_1 \prec W_2$ implies $\theta(W_1) < \theta(W_2)$), we could have reversed the inequalities to $\theta(W_1) \geq \theta(W_2)$ (resp. $\theta(W_1) > \theta(W_2)$). There are only trivial examples of cut distance compatible parameters in this sense which are the graphs that are disjoint union of cliques on one and two vertices. For these graphs the homomorphism densities are either always constant one (if the graph is a disjoint union of vertices), or the edge density of the graph (otherwise). Since we know that $U \succeq V$ implies that the edge densities of the two graphons are the same (Fact 2.5), these examples are cut distance compatible parameters in both senses for a trivial reason, and, in particular, they are not cut distance identifying parameters in this reverse sense. To see that there are no other examples of cut distance compatible parameters in the reverse sense, consider the two following graphons: a graphon W_{clique} consisting of a clique of measure 0.5 ($W_{clique}(x,y) = 1$ if and only if $0 \leq x, y \leq \frac{1}{2}$ and $W_{clique}(x,y) = 0$ otherwise), and the constant graphon $W_{\mathrm{const}} \equiv \frac{1}{4}$. Now let H be a graph that is not a disjoint union of cliques of order one or two. Without loss of generality we assume that H*

does not contain any component consisting of a single vertex. Hence $2e(H) > v(H)$. *Now we have* $W_{\mathrm{const}} \preceq W_{\mathrm{clique}}$, *but* $t(H, W_{\mathrm{const}}) = \left(\frac{1}{4}\right)^{e(H)} < \left(\frac{1}{2}\right)^{v(H)} = t\left(H, W_{\mathrm{clique}}\right)$.

3.5.1. *Step Sidorenko versus step forcing.* We are not able to provide similar characterizations of graphs with the step forcing property and we leave it as a conjecture.

Conjecture 3.22. *A graph H is norming if and only if it is has the step forcing property and if and only if* $t(H,\cdot)$ *is a cut distance identifying parameter.*

It seems possible that the implication that norming graphs have the step forcing property can be proved using Theorem 2.16 from [12] about moduli of convexity of seminorming graphs.

We summary the relations between various properties of graphs related to norms in the Figure 3.2.

3.5.2. *Two positive results.* We conclude the treatment of Problem 3.16 by two positive results, namely that stars are step Sidorenko and that even cycles are step forcing. Propositions 3.23 and 3.24 in the case $\ell = 2$ are not new and follow from the results on weakly norming and Hölder graphs above. Yet, the short proofs given here nicely employ other parts of the theory established in this paper.

Proposition 3.23. *For each* $\ell \in \mathbb{N}$, *the graphon parameter* $t(K_{1,\ell}, \cdot) : \mathcal{W}_0 \to \mathbb{R}$ *is cut distance compatible.*

Proof. The key is to observe that for a graphon Γ, we have $t(K_{1,\ell}, \Gamma) = \int_{x\in[0,1]} x^\ell \mathrm{d}\mathbf{Y}_\Gamma$, where $\mathbf{Y}_\Gamma$ is defined by (2.8). So, suppose that $U \preceq W$. By Proposition 2.7, we have that $\mathbf{Y}_U$ is at least as flat as $\mathbf{Y}_W$. Let Λ be a measure on $[0,1]^2$ as in Definition 2.6 that witnesses this. We have

$$\begin{aligned} t(K_{1,\ell}, U) &= \int_{x\in[0,1]} x^\ell \mathrm{d}\mathbf{Y}_U \\ \boxed{\text{by Lemma 4.10 from [9]}} \quad &= \int_{x\in[0,1]} \left(\int_{y\in[0,1]} y \, \mathrm{d}\mathbf{Y}_W \right)^\ell \mathrm{d}\mathbf{Y}_U \\ \boxed{\text{Jensen's inequality}} \quad &\le \int_{x\in[0,1]} \int_{y\in[0,1]} y^\ell \, \mathrm{d}\mathbf{Y}_W \, \mathrm{d}\mathbf{Y}_U \\ &= \int_{y\in[0,1]} y^\ell \, \mathrm{d}\mathbf{Y}_W \\ &= t(K_{1,\ell}, W) \ . \end{aligned}$$

□

Proposition 3.24. *For each* $\ell \in \{2,3,4,\ldots\}$, *the graphon parameter* $t(C_{2\ell}, \cdot) : \mathcal{W}_0 \to \mathbb{R}$ *is cut distance identifying.*

Before giving a proof, let us note that Lemma 11 in [7] is equivalent to the case $\ell = 2$ of the proposition. However, the proof in [7] does not seem to generalize to higher ℓ, in which case Proposition 3.24 seems to be new.

Proof of Proposition 3.24. To prove the proposition, suppose that ℓ is fixed and $W_1 \prec W_2$ are two graphons. Theorem 3.14 tells us that $W_1 \overset{S}{\prec} W_2$. That is, the sum of the (2ℓ)-th powers of eigenvalues of W_1 is strictly smaller than that of W_2. The statement now follows from Equation (2.4). □

ACKNOWLEDGMENTS

We would like to thank Daniel Král for his useful suggestions. Most of this work was done while Hladký was hosted by Andreas Thom at the lively Institute of Geometry at TU Dresden.

REFERENCES

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

[2] 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.

[3] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[4] D. Conlon, J. Fox, and B. Sudakov. An approximate version of Sidorenko's conjecture. *Geom. Funct. Anal.*, 20(6):1354–1366, 2010.

[5] D. Conlon and J. Lee. Sidorenko's conjecture for blow-ups. arXiv:1809.01259.

[6] D. Conlon and J. Lee. Finite reflection groups and graph norms. *Adv. Math.*, 315:130–165, 2017.

[7] J. W. Cooper, D. Kral, and T. L. Martins. Finitely forcible graph limits are universal. Preprint (arXiv:1791.03846v3).

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

[9] J. Doležal, M. Hladký. Cut-norm and entropy minimization over weak* limits. arXiv:1705.09160.

[10] M. Doležal, J. Grebík, J. Hladký, I. Rocha, and V. Rozhoň. Relating the cut distance and the weak* topology for graphons. arXiv:1806.07368.

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

[12] H. Hatami. Graph norms and Sidorenko's conjecture. *Israel J. Math.*, 175:125–150, 2010.

[13] J. H. Kim, C. Lee, and J. Lee. Two approaches to Sidorenko's conjecture. *Trans. Amer. Math. Soc.*, 368(7):5057–5074, 2016.

[14] D. Král, L. M. Lovász, J. Noel, and J. Sosnovec. Finitely forcible graphons with an almost arbitrary structure. arXiv:1809.05973.

[15] D. Král', T. Martins, P. P. Pach, and M. Wrochna. The step Sidorenko property and non-norming edge-transitive graphs. arXiv:1802.05007.

[16] J. L. X. Li and B. Szegedy. On the logarithimic calculus and Sidorenko's conjecture. Preprint (arXiv:1107.1153).

[17] L. Lovász. Subgraph densities in signed graphons and the local Simonovits-Sidorenko conjecture. *Electron. J. Combin.*, 18(1):Paper 127, 21, 2011.

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

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

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

[21] 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.

[22] A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.*, 9(2):201–204, 1993.

[23] M. Simonovits. Extremal graph problems, degenerate extremal problems, and supersaturated graphs. In *Progress in graph theory (Waterloo, Ont., 1982)*, pages 419–437. Academic Press, Toronto, ON, 1984.

[24] J. Skokan and L. Thoma. Bipartite subgraphs and quasi-randomness. *Graphs Combin.*, 20(2):255–262, 2004.

[25] B. Szegedy. An information theoretic approach to Sidorenko's conjecture. Preprint (arXiv:1406.6738).

Doležal, Grebík, Hladký: INSTITUTE OF MATHEMATICS, CZECH ACADEMY OF SCIENCES. ŽITNÁ 25, 110 00, PRAHA, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985840, *Rocha, Rozhoň*: INSTITUTE OF COMPUTER SCIENCE, CZECH ACADEMY OF SCIENCES. POD VODÁRENSKOU VĚŽÍ 2, 182 07, PRAGUE, CZECHIA. WITH INSTITUTIONAL SUPPORT RVO:67985807.

Email address: `dolezal@math.cas.cz, greboshrabos@seznam.cz, honzahladky@gmail.com, israelrocha@gmail.com, vaclavrozhon@gmail.com`