HYPERGRAPH LIMITS: A REGULARITY APPROACH

YUFEI ZHAO

ABSTRACT. A sequence of k-uniform hypergraphs $H_1, H_2, \dots$ is convergent if the sequence of homomorphism densities $t(F, H_1), t(F, H_2), \dots$ converges for every k-uniform hypergraph F. For graphs, Lovász and Szegedy showed that every convergent sequence has a limit in the form of a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$. For hypergraphs, analogous limits $W\colon [0,1]^{2^k-2} \to [0,1]$ were constructed by Elek and Szegedy using ultraproducts. These limits had also been studied earlier by Hoover, Aldous, and Kallenberg in the setting of exchangeable random arrays.

In this paper, we give a new proof and construction of hypergraph limits. Our approach is inspired by the original approach of Lovász and Szegedy, with the key ingredient being a weak Frieze-Kannan type regularity lemma.

1. Introduction

One of the starting points in the theory of dense graph limits is the seminal paper by Lovász and Szegedy [13] where they constructed limit objects for convergent sequences of dense graphs. The subject has grown enormously since then with many exciting developments (see Lovász's recent monograph [12]).

For any two graphs F and G, let $\hom(F, G)$ denote the number of homomorphism from F to G, i.e., maps $V(F) \to V(G)$ that carry every edge of F to an edge of G. The homomorphism density $t(F, G)$ is defined to be the probability that a random map $V(F) \to V(G)$ is a homomorphism, i.e.,

$$t(F, G) := \frac{\hom(F, G)}{|V(G)|^{|V(F)|}}.$$

A sequence of graphs $G_1, G_2, \dots$ is called *convergent* if the sequence $t(F, G_1), t(F, G_2), \dots$ converges for every graph F. Convergent graph sequences were defined and studied in [4, 5]. The main result of Lovász and Szegedy [13] is that for every convergent graph sequence there is a limit object in the form of a *graphon*, which is a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$ (here *symmetric* means that $W(x, y) = W(y, x)$) such that $t(F, G_n) \to t(F, W)$ as $n \to \infty$ for all graphs F. Here $t(F, W)$ is defined by

$$t(F, W) := \int_{[0,1]^{V(F)}} \prod_{ij \in E(F)} W(x_i, x_j)\, dx_1 dx_2 \cdots dx_{|V(F)|}$$

The natural extension of these limits to hypergraphs was considered by Elek and Szegedy [7]. They constructed using ultraproducts an "ultralimit hypergraph" for any sequence of hypergraphs, and established a correspondence principle which enabled them to convert statements about finite hypergraphs, such as hypergraph regularity and removal lemmas [9, 15, 16], to measure-theoretic claims about ultralimit spaces. One of the consequences of their work is the existence of a limit object in the form of a measurable functions $W\colon [0,1]^{2^k-2} \to [0,1]$ for any convergent sequence of k-uniform hypergraphs.

These limit objects had actually appeared earlier in a different form, in the study of exchangeable random arrays, initiated by Hoover [10], Aldous [1], and Kallenberg [11] during the 1980s, building on the classic de Finnetti's theorem on exchangeable random variables. This connection is explained in the survey [3] by Austin, where he credits Tao [17] for initiating the link between exchangeable random variables and hypergraphs. These connections for graphs are also explained in the survey by Diaconis and Janson [6] as well as Aldous' ICM talk [2].

The purpose of this paper is to provide a new proof of the existence of hypergraph limits. Our approach is based on weak Frieze-Kannan [8] type regularity partitions, in line with mainstream perspectives on dense graph limits. The proof does not use any exchangeable random variables or ultraproducts, and the construction of the limit is subjectively more concrete than earlier proofs. Our proof is inspired by the original approach of Lovász and Szegedy [13], and the paper is self-contained other than an application of the Martingale Convergence Theorem.

1.1. **Convergence and limit object.** For any k-uniform hypergraphs F and H, let $\hom(F,H)$ denote the number of homomorphisms from F to H, i.e., maps $V(F) \to V(H)$ that carry every edge of F to an edge of H. Define $t(F,H) := \hom(F,H)/\left|V(H)\right|^{|V(F)|}$. This is the probability that a random map $V(F) \to V(H)$ is a homomorphism.

Definition 1.1 (Convergence)**.** A sequence of k-uniform hypergraphs $H_1, H_2, \dots$ is called *convergent* if the sequence $t(F,H_1), t(F,H_2), \dots$ converges for every k-uniform hypergraph F.

For any positive integer n, define $[n] := \{1,2,\dots,n\}$. For any set A, define $r(A)$ to be the collection of all nonempty subsets of A, and $r_<(A)$ to be collection of all nonempty proper subsets of A. More generally, let $r(A,m)$ denote the collection of all nonempty subsets of A of size at most m. So for instance, $r_<([k]) = r([k],k-1)$. We will also use the shorthand $r[k]$ and $r_<[k]$ to mean $r([k])$ and $r_<([k])$ respectively.

Any permutation σ of a set A induces a permutation on $r(A,m)$. We say that a function $W\colon [0,1]^{r([k],m)} \to [0,1]$ is *symmetric* if it remains invariant under any permutation of the coordinates induced by any permutation of $[k]$. For example, $W\colon [0,1]^{r_<[3]} \to [0,1]$ being symmetric means that

$$W(x_1,x_2,x_3,x_{12},x_{13},x_{23}) = W(x_{\sigma_1},x_{\sigma_2},x_{\sigma_3},x_{\sigma_1\sigma_2},x_{\sigma_1\sigma_3},x_{\sigma_2\sigma_3}) \tag{1}$$

for any permutation σ of $\{1,2,3\}$. Here we write x_i for $x_{\{i\}}$ and x_{ij} for $x_{\{i,j\}}$.

Definition 1.2. A *k-uniform hypergraphon* is a symmetric measurable function $W\colon [0,1]^{r_<([k])} \to [0,1]$.

Example 1.3. A 3-uniform hypergraphon is a measurable function $W\colon [0,1]^6 \to [0,1]$ satisfying the symmetry condition (1).

For any k-uniform hypergraph F and hypergraphon W, define the homomorphism density by

$$t(F,W) := \int_{[0,1]^{r(V(F),k-1)}} \prod_{A\in E(F)} W(\mathbf{x}_{r_<(A)})\, d\mathbf{x}$$

Our convention throughout the paper is that if $\mathbf{x} = (x_A : A \in \mathcal{A}) \in [0,1]^{\mathcal{A}}$ is a vector whose coordinates are indexed by some set system $\mathcal{A}$, and $\mathcal{B} \subseteq \mathcal{A}$ is a subcollection, then we write $\mathbf{x}_{\mathcal{B}} = (x_B : B \in \mathcal{B}) \in [0,1]^{\mathcal{B}}$ to mean the restriction of the vector to the coordinates indexed by $\mathcal{B}$.

Example 1.4. If $K_4^{(3)} = \{123,124,134,234\}$ is the complete 3-uniform hypergraph on 4 vertices and W is a 3-uniform hypergraphon, then

$$t(K_4^{(3)},W) = \int_{[0,1]^{10}} W(x_1,x_2,x_3,x_{12},x_{13},x_{23})W(x_1,x_2,x_4,x_{12},x_{14},x_{24})W(x_1,x_3,x_4,x_{13},x_{14},x_{34})\cdot$$
$$\cdot W(x_2,x_3,x_4,x_{23},x_{24},x_{34})\, dx_1dx_2dx_3dx_4dx_{12}dx_{13}dx_{14}dx_{23}dx_{24}dx_{34}.$$

Every k-uniform hypergraph H can be represented as a k-uniform hypergraphon W^H as follows: divide $[0,1]$ into $|V(H)|$ equal-length intervals $\{I_1, I_2, \ldots, I_{|V(H)|}\}$. For each $\mathbf{x} \in [0,1]^{r_<[k]}$ define

$$W^H(\mathbf{x}) := \begin{cases} 1 & \text{if } x_{\{i\}} \in I_{a_i} \text{ for } i = 1, \ldots, k \text{ and } \{a_1, a_2, \ldots, a_k\} \text{ is an edge of } H, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $W^H(\mathbf{x})$ depends only on the k coordinates of $\mathbf{x}$ corresponding to subsets of $[k]$ of size 1. It can be alternatively described as transforming the adjacency array of H into a $\{0,1\}$-valued step function and then adding $2^k - 2 - k$ extra free coordinates. Observe that $t(F, H) = t(F, W^H)$ for every k-uniform hypergraph F.

The main purpose of this paper is to give a new proof of the following result [7, Thm. 7] on the existence of hypergraph limits.

Theorem 1.5. *If $H_1, H_2, \ldots$ is a convergent sequence of k-uniform hypergraphs, then there exists a k-uniform hypergraphon W so that $t(F, H_n) \to t(F, W)$ as $n \to \infty$ for every k-uniform hypergraph F.*

1.2. **Why are there so many coordinates?** It may initially seem somewhat strange that we need 6 coordinates to describe the limit of 3-uniform hypergraphs, whereas every 3-uniform hypergraph can be described in terms of a 3-dimensional adjacency array. These extra dimensions do not arise for limits of graphs, but they are essential for hypergraphs. Here is a standard example illustrating why functions of the form $[0,1]^3 \to [0,1]$ cannot capture the richness of 3-uniform hypergraph limits. Take $G_n \sim \mathbb{G}(n, 1/2)$ to be a sequence of graphs on n vertices, where each edge is generated with probability $1/2$, and let H_n be the 3-uniform hypergraph whose edges are the triangles of G_n. Then with probability one, $t(F, H_n) \to 2^{-|\partial F|}$ for every 3-uniform hypergraph F, where ∂F is the collection of unordered pairs of vertices of F that are contained in some edge of F. The limit of H_n is different from, say, the constant hypergraphon $1/2$, which is the limit of a sequence of 3-uniform hypergraphs where every triple of vertices is taken to be an edge independently with probability $1/2$. To describe the limit of H_n, we need to incorporate the limit of G_n into the data, and this is achieved by the three extra coordinates. We know that the graph sequence G_n converges to the constant graphon with value $1/2$. To build the limit of H_n, we partition each of the last three coordinates, x_{12}, x_{13}, x_{23} into two intervals $[0,1/2] \cup (1/2,1]$, corresponding to the limit of G_n and the limit of its complement. The limiting hypergraphon has constant value 1 on $[0,1]^3 \times [0,1/2]^3$ (as the edges of H_n are supported on G_n) and 0 elsewhere. Intuitively, the first three coordinates encode the vertex types, the last three coordinates encode the vertex-pair types. This hypergraphon is $\{0,1\}$-valued since it is deterministic once the vertex and vertex-pairs types are set. If we modify the sequence H_n so that each triangle of G_n is included as an edge of H_n with some probability p independently, then the limiting hypergraphon would be constant p on $[0,1]^3 \times [0,1/2]^3$ and 0 elsewhere.

For k-uniform hypergraphs, we can similarly impose some structure at each level, corresponding to j-element subsets of vertices, for every $1 \le j \le k$. This is why we need a coordinate for every proper subset of $[k]$ to describe hypergraph limits.

1.3. **Random hypergraph model.** To further illustrate the involvement of the $2^k - 2$ coordinates in a hypergraphon, let us review the associated random hypergraph model.

Recall that if $W\colon [0,1]^2 \to [0,1]$ is a graphon, then we have the following natural random graph model $\mathbb{G}(n, W)$ on n vertices: choose i.i.d. uniform $x_1, x_2, \ldots, x_n \in [0,1]$, and let there be an edge between vertices i and j with probability $W(x_i, x_j)$ independently. It was shown [13, Cor. 2.6] using Azuma's inequality that $\mathbb{G}(n, W)$ converges to the limit W almost surely.

Similarly, a k-uniform hypergraphon W gives a natural model $\mathbb{G}(n, W)$ of a random k-uniform hypergraph on n vertices: choose a uniformly random $\mathbf{x} \in [0,1]^{r([n],k-1)}$ and add the edge $B = \{i_1, \ldots, i_k\} \subseteq [n]$ with probability $W(\mathbf{x}_{r_<(B)})$ independently. Essentially the same proof for graphs extend over to show [7, Thm. 11] that $\mathbb{G}(n, W)$ converges to W in the sense of Theorem 1.5, as $n \to \infty$ with probability one. Observe that the random hypergraphs H_n of triangles in $\mathbb{G}(n, 1/2)$ discussed earlier is a special case of this model.

1.4. **Analytic version and compactness.** It will be convenient to prove an analytic version of Theorem 1.5. We say that a sequence of k-uniform hypergraphons $W_1, W_2, \dots$ is *convergent* if the sequence $t(F, W_1), t(F, W_2), \dots$ converges for every k-uniform hypergraph F.

Theorem 1.6. *If $W_1, W_2, \dots$ is a convergent sequence of k-uniform hypergraphons, then there exists a k-uniform hypergraphon $\widetilde{W}$ so that $t(F, W_n) \to t(F, \widetilde{W})$ as $n \to \infty$ for every k-uniform hypergraph F.*

In this case we say that W_n converges to $\widetilde{W}$. Here is an equivalent formulation of the theorem.

Theorem 1.7. *Every sequence $W_1, W_2, \dots$ of k-uniform hypergraphons contains a subsequence that converges to some k-uniform hypergraphon $\widetilde{W}$.*

Theorem 1.7 implies Theorem 1.6 trivially since we can just take the limit $\widetilde{W}$ produced by Theorem 1.7. The converse is true because $[0, 1]^{\mathbb{N}}$ is sequentially compact, so we can restrict (W_n) to some subsequence (W_{n_i}) so that $t(F, W_{n_i})$ converges as $i \to \infty$ for every F.

We shall prove Theorem 1.7 with respect to another notion of convergence based on regular partitions, which implies the convergence of homomorphism densities. The partition-based convergence gives some structural insight into the convergence of hypergraphs.

There is a neat interpretation of Theorem 1.7 in terms of compactness, discovered by Lovász and Szegedy [14] in the case of graphons. Let $\mathcal{W}_0^{(k)}$ denote the set of k-uniform hypergraphons. Give $\mathcal{W}_0^{(k)}$ the weakest topology for which the functions $t(F, \cdot)$ are continuous for every k-uniform hypergraph F. Identify W with W' if $t(F, W) = t(F, W')$ for every k-uniform hypergraph F. Call this topology the *left-convergence topology* of $\mathcal{W}_0^{(k)}$.

Corollary 1.8. *The space $\mathcal{W}_0^{(k)}$ with the left-convergence topology is compact.*

Proof. The space is metrizable with the metric $\delta(W, W') = \sum_{i \geq 1} 2^{-i} |t(F_i, W) - t(F_i, W')|$ where (F_i) is some enumeration of all isomorphism classes of k-uniform hypergraphs. We know that compactness is equivalent to sequential compactness in metric spaces, and Theorem 1.7 shows that the space is sequentially compact. ■

When $k = 2$, Lovász and Szegedy [14] showed that $\mathcal{W}_0^{(2)}$ is compact under the cut metric topology, and Borgs, Chayes, Lovász, Sós, and Vesztergombi [4] showed that the cut metric topology is equivalent to the left-convergence topology. Lovász and Szegedy interpreted the compactness with respect to the cut metric as an analytic form of the regularity lemma, and they showed that the compactness of the space of graphons implies strong versions of the regularity lemma. Unfortunately, for $k \geq 3$, we do not know of a useful extension of the cut metric to hypergraphs (and there may be some reasons to believe that such a natural metric might be too much to ask for). This is one of the main obstacles in working with convergence of hypergraphs. It would be nice to have a simple and useful description of distance between hypergraphs which agrees with the topology induced by homomorphism densities.

1.5. **Organization.** In §2 we review the Lovász-Szegedy construction of graph limits. In §3 we give an informal sketch of the proof of the existence of 3-uniform hypergraph limits. Most of the ideas, minus the technical hairiness, are contained in §3. The proof of the main result is contained in §4–6. §4 collects some of the notation used in the proof. §5 contains the regularity and counting lemmas central to the proof. In §6 we introduce branching partitions and formulate the notion of partitionable convergence, which implies, via counting lemmas, the convergence of homomorphism densities. We then prove the existence of limits with respect to partitionable convergence.

2. Limits of graphons

For any symmetric measurable function $W\colon [0,1]^2 \to \mathbb{R}$, the cut norm is defined by

$$\|W\|_\square := \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} W(x,y)\,dxdy \right|, \tag{2}$$

where S and T range over all measurable subsets of $[0,1]$. We have the identity

$$\|W\|_\square = \sup_{u,v\colon [0,1]\to[0,1]} \left| \int W(x,y)u(x)v(y)\,dxdy \right| \tag{3}$$

where u and v range over all measurable functions $[0,1] \to [0,1]$. Indeed, since the integral in (3) is linear in both u and v, one can restrict to $\{0,1\}$-valued u and v, thereby reducing to (2).

Recall that a graphon is a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$. For any measure preserving bijection $\phi\colon [0,1] \to [0,1]$ and any graphon W, define W^ϕ by $W^\phi(x,y) = W(\phi(x),\phi(y))$. We define the cut distance between graphons by

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

where the infimum is taken over all measure preserving bijections $\phi\colon [0,1] \to [0,1]$. The cut distance can be defined for pairs of graphs by considering their associated graphons. Graphs that are close in cut distance are also close in homomorphism densities, by the following counting lemma.

Lemma 2.1 (Counting lemma)**.** *For any graphons U and W and any graph F, we have*

$$|t(F,U) - t(F,W)| \le e(F)\,\|U - W\|_\square$$

where $e(F)$ is the number of edges of F.

We illustrate the proof through the example $F = K_3$.

$$\begin{aligned}
&t(K_3,U) - t(K_3,W)\\
&= \int_{[0,1]^3} (U(x,y)U(x,z)U(y,z) - W(x,y)W(x,z)W(y,z))\,dxdydz\\
&= \int_{[0,1]^3} (U(x,y) - W(x,y))W(x,z)W(y,z)\,dxdydz\\
&\quad + \int_{[0,1]^3} U(x,y)(U(x,z) - W(x,z))W(y,z)\,dxdydz\\
&\quad + \int_{[0,1]^3} U(x,y)U(x,z)(U(y,z) - W(y,z))\,dxdydz
\end{aligned}$$

Each of the three terms in the final sum is bounded in absolute value by $\|U - W\|_\square$. For example, for the first term, for every fixed value of z, the integral has the form (3), and so it is bounded in absolute value by $\|U - W\|_\square$, and the same bound holds after integrating z by the triangle inequality. It follows that $|t(K_3,U) - t(K_3,W)| \le 3\,\|U - W\|_\square$.

For any graphon W and any partition $\mathcal{Q}$ of $[0,1]$ into a finite collection of measurable subsets, let $W_\mathcal{Q}$ be a graphon which is the step function obtained from W by replacing its value at $(x,y) \in Q_i \times Q_j$ by the average value of W on $Q_i \times Q_j$, for any $Q_i, Q_j \in \mathcal{Q}$, (if either Q_i or Q_j has measure zero, then assign value 0 on $Q_i \times Q_j$). For graphs, think of $\mathcal{Q}$ as a partition of the vertex set, and $W_\mathcal{Q}$ as recording the edge densities between pairs of vertex subsets.

A key tool in the construction of graph limits is the following weak regularity lemma due to Frieze and Kannan [8] (see also [14, Lem 3.1]). It can be proved by an L^2-energy increment argument.

Lemma 2.2 (Weak regularity lemma). *For every $\varepsilon > 0$ and every symmetric measurable function $W\colon [0,1]^2 \to [0,1]$, there is some partition $\mathcal{Q}$ of $[0,1]$ into at most $2^{2/\varepsilon^2}$ parts such that $\|W - W_{\mathcal{Q}}\|_\square \le \varepsilon$.*

Lovász and Szegedy [14] showed that with respect to the cut metric, after identifying graphons with cut distance zero, the space of all graphons is compact. Equivalently:

Theorem 2.3 (Lovász and Szegedy [14]). *Every sequence $W_1, W_2, \dots$ of graphons contains a subsequence converging to some graphon $\widetilde{W}$ in cut distance.*

Let us recall the idea of the proof of Theorem 2.3. Let $\varepsilon > 0$. We apply the weak regularity lemma to approximate every W_n by some $(W_n)_{\mathcal{Q}_n}$. By replacing each W_n by some $W_n^{\phi_n}$ for some measure preserving bijection ϕ_n, we may assume that the partition $\mathcal{Q}_n$ divides $[0,1]$ into intervals. Take a subsequence so that the lengths of the intervals converge, and the values of $(W_n)_{\mathcal{Q}_n}$ inside the boxes induced by the partition also converge, i.e., the value inside the (i,j)-th box of $(W_n)_{\mathcal{Q}_n}$ converges to some value as $n \to \infty$ (may be different limits for different (i,j)). Then in this subsequence, $(W_n)_{\mathcal{Q}_n}$ converges pointwise almost everywhere to some limit $\widetilde{U}_1$, which is also a step function.

Now repeat the same procedure with a smaller $\varepsilon' < \varepsilon$. We obtain new partitions $\mathcal{Q}'_n$ which are refinements of previous partitions. Call the resulting limit $\widetilde{U}_2$. Note that steps of $(W_n)_{\mathcal{Q}'_n}$ are refinements of the steps of $(W_n)_{\mathcal{Q}_n}$, and the values of the latter can be obtained from the former by averaging over each step. Thus a similar relation holds for $\widetilde{U}_2$ and $\widetilde{U}_1$.

Now we repeat this procedure for a sequence of ε_k tending to zero. We obtain a sequence $\widetilde{U}_1, \widetilde{U}_2, \dots$ of step functions so that each $\widetilde{U}_s$ can be obtained from $\widetilde{U}_{s+1}$ by average over each step. It follows that if (X,Y) is a uniform random point in $[0,1]^2$, then the sequence $(\widetilde{U}_1(X,Y), \widetilde{U}_2(X,Y), \dots)$ is a martingale. Since every $\widetilde{U}_s$ is bounded, the Martingale Convergence Theorem[1] implies that the martingale converges with probability 1, and hence there is some $\widetilde{W}\colon [0,1]^2 \to [0,1]$ which is the pointwise almost everywhere limit of $\widetilde{U}_s$'s. One then checks that $\widetilde{W}$ is the desired limit.

In summary, the above proof consists of two main steps:

(1) For each error tolerance ε, apply a weak regularity lemma to get a finite-dimensional step function approximation of each graphon. Take a subsequence so that the step functions converge.
(2) Take a decreasing sequence of ε tending to zero, we obtain refining chains of regularity partitions, and the corresponding subsequential limits $\widetilde{U}_s$ form a martingale. The existence of the final limit graphon follows by the Martingale Convergence Theorem.

3. Limits of 3-uniform hypergraphs

In this section we sketch the idea for 3-uniform hypergraph limits. To keep things simple, consider a sequence $H_1, H_2, \dots$ of 3-uniform hypergraphs (as opposed to hypergraphons).

We begin with an initial attempt that does not quite work. For a 3-variable function $W\colon [0,1]^3 \to \mathbb{R}$, we might extend the cut norm (5) as follows (assume everything is measurable from now on):

$$\text{(bad cut norm)} \qquad \|W\|_\square = \sup_{R,S,T \subseteq [0,1]} \left| \int_{R \times S \times T} W(x,y,z)\, dx dy dz \right|. \tag{4}$$

For each hypergraph H, one can easily extend the weak regularity lemma, Lemma 2.2, to obtain a partition $\mathcal{Q}$ of the vertex set of H into at most $2^{3/\varepsilon^2}$ parts so that $\left\|W^H - W^H_{\mathcal{Q}}\right\|_\square \le \varepsilon$ (regard W^H as a 3-variable function for now, and $W^H_{\mathcal{Q}}$ is derived from W by averaging over each cells

[1]The Martingale Convergence Theorem (see [18, Thm. 11.5]) says that every L^1-bounded martingale converges almost surely. Our martingales are actually bounded uniformly within $[0,1]$.

induced by $\mathcal{Q}$). Theorem 2.3 also extends with virtually no change in the proof. That is, allowing permutations of vertices, some subsequence of H_n converges with respect to the vertex-cut norm (4) to a 3-variable symmetric function $\widetilde{W}\colon [0,1]^3 \to [0,1]$.

Unfortunately, the vertex-cut norm (4) is not strong enough to guarantee a counting lemma. We want to say that if H_1 and H_2 are close with respect to some cut norm, then $t(F, H_1)$ and $t(F, H_2)$ are close. If we carry through the proof of Lemma 2.1, we find that $|t(F, H_1) - t(F, H_2)| \le e(F)\|W^{H_1} - W^{H_2}\|_\square$ holds when F is a *linear hypergraph*, i.e., where every two edges of F intersect in at most one vertex. However, when F is not linear, say $F = K_4^{(3)}$, then this claim is completely false, as $t(F, H_1)$ and $t(F, H_2)$ can be separated even when $\|W^{H_1} - W^{H_2}\|_\square$ is small. A counterexample for 3-uniform hypergraphs can be built by taking triangles of the random graph $\mathbb{G}(n, p)$, and then keeping each triangle as a 3-uniform edge with some probability q. With parameters $(p, q) = (1/2, 1)$ and $(1, 1/8)$, we obtain 3-uniform hypergraphs that are close with respect to the vertex-cut norm, and yet they have very different $K_4^{(3)}$ densities.

Now let us scrap the vertex-cut norm (4). The proof of the counting lemma, Lemma 2.1, extends with respect to the following modified cut norm (again we use a 3-variable W for now):

$$\text{(better cut norm)} \qquad \|W\|_{\square^2} = \sup_{\substack{u,v,w\colon [0,1]^2\to[0,1]\\ \text{symmetric}}} \left| \int_{[0,1]^3} W(x,y,z)u(x,y)v(x,z)w(y,z)\, dxdydz \right|. \qquad (5)$$

For this cut norm, the counting lemma $|t(F, H_1) - t(F, H_2)| \le e(F)\|W^{H_1} - W^{H_2}\|_{\square^2}$ holds. However, like trying to fit a large rug in a small room, we quickly run into another issue: this norm is too strong and we do not have the compactness result corresponding to Theorem 2.3. Indeed, taking the sequence H_n of triangles of $\mathbb{G}(n, 1/2)$ from §1.2, the two hypergraphs H_n and H_m are typically not close with respect to $\|\cdot\|_{\square^2}$, although they are close in homomorphism densities.

Even though we do not have compactness with respect to $\|\cdot\|_{\square^2}$, we can still hope for a slightly weaker topology that gives convergence of homomorphism densities. We can extend the weak regularity lemma, Lemma 2.2, to $\|\cdot\|_{\square^2}$, where now instead of partitioning the vertex set $V = V(H)$, we partition the edges of the underlying complete graph $K_V = \binom{V}{2}$, i.e., the collection of unordered pairs of V. So now $\mathcal{Q}$ is a partition $K_V = G_1 \cup \cdots \cup G_m$ of the edges of K_V into m graphs. The partition $\mathcal{Q}$ of K_V induces a partition $\mathcal{Q}^*$ on triples of vertices:

$$(x, y, z) \sim_{\mathcal{Q}^*} (x', y', z') \Leftrightarrow (x, y) \sim_{\mathcal{Q}} (x', y'),\ (x, z) \sim_{\mathcal{Q}} (x', z'),\ \text{and } (y, z) \sim_{\mathcal{Q}} (y', z').$$

Being somewhat sloppy with notation for the time being, we can form $W^H_{\mathcal{Q}}$ by averaging W^H inside each cell of $\mathcal{Q}^*$. Then the weak regularity lemma guarantees us a partition $\mathcal{Q}$ of K_V into at most $2^{3/\varepsilon^2}$ parts so that $\|W^H - W^H_{\mathcal{Q}}\|_{\square^2} \le \varepsilon$, and $|t(F, W^H) - t(F, W^H_{\mathcal{Q}})| \le e(F)\varepsilon$ by the counting lemma.

For each hypergraph in the sequence $H_1, H_2, \ldots$, apply the weak regularity lemma (for a uniform ε) to obtain a partition $\mathcal{Q}_n$ of the complete graph on $V(H_n)$ into m graphs: $K_{V(H_n)} = G_{n,1} \cup \cdots \cup G_{n,m}$, where m depends on ε but not on n.

By applying Theorem 2.3 on the graph sequence $(G_{n,1})_{n\ge1}$, we can find a graphon $\widetilde{Y}_1\colon [0,1]^2 \to [0,1]$ so that $G_{n,1}$ converges to $\widetilde{Y}_1$ as $n \to \infty$ along some subsequence . By further restricting to subsequences, we can find a $\widetilde{Y}_j$ for each $1 \le j \le m$ so that $G_{n,j}$ converges to $\widetilde{Y}_j$ as $n \to \infty$ along a subsequence.

For each n, $\{G_{n,1}, \ldots, G_{n,m}\}$ is a partition of $K_{V(H_n)}$, so the same holds for the resulting limit[2], in the sense that $\widetilde{Y}_1 + \cdots + \widetilde{Y}_m = 1$ almost everywhere as functions $[0,1]^2 \to [0,1]$. Next we build a partition $\widetilde{Q}$ of the cube $[0,1]^3 = [0,1]^{r[2]}$ (coordinates indexed by x_1, x_2, x_{12}) by stacking together

[2]Provided that the limits of the various graph sequences are taken in a compatible way. This is a source of technical/notational annoyance later on, and it is the reason for introducing branching partitions in §6.

subsets whose heights are given by $\widetilde{Y}_j$. More precisely, $\widetilde{Q} = \{\widetilde{Q}_1, \dots, \widetilde{Q}_m\}$ where

$$\widetilde{Q}_j = \{(x_1, x_2, x_{12}) \in [0,1]^3 : (\widetilde{Y}_1 + \cdots + \widetilde{Y}_{j-1})(x_1, x_2) \le x_{12} < (\widetilde{Y}_1 + \cdots + \widetilde{Y}_j)(x_1, x_2)\}.$$

This is the first place where the "extra" coordinates such as x_{12} arise even though we started with hypergraphs not requiring these extra coordinates. They arise because the limit graphon $\widetilde{Y}_1$ of a sequence of graphs $G_{n,1}$ is not always a $\{0,1\}$-valued function.

The partition $\widetilde{\mathcal{Q}}$ of $[0,1]^{r[2]}$ induces a partition $\widetilde{\mathcal{Q}}^*$ of $[0,1]^6 = [0,1]^{r_<[2]}$:

$$(x_1, x_2, x_3, x_{12}, x_{13}, x_{23}) \sim_{\widetilde{\mathcal{Q}}^*} (x_1', x_2', x_3', x_{12}', x_{13}', x_{23}') \Leftrightarrow (x_i, x_j, x_{ij}) \sim_{\widetilde{\mathcal{Q}}} (x_i', x_j', x_{ij}') \ \forall 1 \le i < j \le 3.$$

The partition $\widetilde{\mathcal{Q}}^*$ should not be viewed as a regularization partition for any H_n (indeed, the extra coordinates do not even appear in H_n). Instead, the partitions $\mathcal{Q}_n$ themselves become increasing close to $\widetilde{\mathcal{Q}}$. There is a correspondence of cells of $\mathcal{Q}_n$ with those of $\widetilde{\mathcal{Q}}$, and this induces a correspondence between cells of $\mathcal{Q}_n^*$ with those of $\widetilde{\mathcal{Q}}^*$.

Now we construct the first limiting hypergraphon $\widetilde{U}_1$ as a step function $[0,1]^6 \to [0,1]$ that is constant on each part of $\widetilde{\mathcal{Q}}^*$. On each part of $\widetilde{\mathcal{Q}}^*$, we assign to $\widetilde{U}_1$ the limiting value of the average of W_n on the corresponding cell of $\mathcal{Q}_n^*$, limit taken as $n \to \infty$ along a further restricted subsequence. We have constructed $\widetilde{U}_1$, which plays a similar role as $\widetilde{U}_1$ near the end of §2.

However, unlike §2, $\widetilde{U}_1$ is not close in $\|\cdot\|_{\square^2}$ to H_n for large n. It is a limit in the following sense: we first ε-regularized H_n, and then took the graph limit of the partitions, created a new partition of $[0,1]^6$ using these lower order limits, and then constructed a step-function U_1 using this limiting partition and the limiting values on the steps. We knew from the earlier counting lemma (referred to later on as *Counting Lemma I*) that

$$\left|t(F, H) - t(F, W^H_{\mathcal{Q}_n})\right| \le e(F)\varepsilon. \tag{6}$$

By what we will call *Counting Lemma II*, we have (here $n \to \infty$ along a subsequence)

$$\lim_{n \to \infty} t(F, W^H_{\mathcal{Q}_n}) = t(F, \widetilde{U}_1). \tag{7}$$

Here is some intuition why (7) holds. Both $W^H_{\mathcal{Q}_n}$ and $\widetilde{U}_1$ are step functions. We can split them up into weighted sums of indicator functions, on which the claim reduces to checking homomorphism densities for the graphons corresponding to parts of the partitions $\mathcal{Q}_n$ and $\widetilde{\mathcal{Q}}$. We know that the graphs which are the parts of $\mathcal{Q}_n$ converge to the graphons from which $\widetilde{\mathcal{Q}}$ is built. So the graph homomorphism densities converge.

This shows that $\widetilde{U}_1$ is a $O(e(F)\varepsilon)$-approximation to a subsequence of H_n in terms of F-densities. Now, take a smaller $\varepsilon' < \varepsilon$, and build another $\widetilde{U}_2$, where the new partitions $\mathcal{Q}_n$ are refinements of the previous ones. Continuing this process, we obtain a sequence $\widetilde{U}_1, \widetilde{U}_2, \dots$ which is a martingale as before. The Martingale Convergence Theorem gives a pointwise almost everywhere limit $\widetilde{W}$ of $\widetilde{U}_s$, $s \to \infty$, and $\widetilde{W}$ is the desired limit.

In proving 3-uniform hypergraph limits, we used the existence of graph limits. In general, we prove the existence of k-uniform hypergraph limits by induction on k. There are a few further technical difficulties. For example, we need to make sure that the limit of a sequence of partitions remains a partition, so the limit needs to be taken in a compatible way. Since we are working with multiple partitions, we will need to deal with homomorphisms from F to a vector of hypergraphons, where the edges of F individually land in different hypergraphons. The details are addressed in the rest of this paper.

4. Notation

One (not so trivial) source of difficulty in working with hypergraphs is the complexity of notation. This section collects some of the notation and conventions used in the rest of this paper. Some notations were already introduced in §1.

We shall omit the word "measurable" as everything we consider is assumed to be measurable.

4.1. Hypergraphs.

A k-uniform hypergraph F is some finite collection of k-element subsets of some ground set, which we denote by $V(F)$. So when we talk about an element of F, we mean an edge of F, and $|F|$ means the number of edges of F.

4.2. Subsets, partitions, and hypergraphons.

Definition 4.1 (Symmetric sets and partitions)**.** A *symmetric (measurable) subset* of $[0,1]^{r[k]}$ is one which is invariant under the action of all permutations of $[k]$. A *symmetric (measurable) partition* of $[0,1]^{r[k]}$ is a partition of $[0,1]^{r[k]}$ into a finite collection of symmetric subsets.

A symmetric subset $P \subseteq [0,1]^{r[k]}$ is associated to a k-hypergraphon $W^P\colon [0,1]^{r_<[k]} \to [0,1]$ by integrating out the top coordinate:

$$W^P(\mathbf{x}_{r_<[k]}) := \int_0^1 1_P(\mathbf{x}_{r[k]})\, dx_{[k]}. \tag{8}$$

For example, for $k = 3$, we have $P \subseteq [0,1]^3$, with coordinates indexed by $r[2] = \{1, 2, 12\}$, and

$$W^P(x_1, x_2) = \int_0^1 1_P(x_1, x_2, x_{12})\, dx_{12}.$$

This operation collapses the final coordinate in P. It will be helpful to think of P and W^P as representing the same object. For example, when $k = 2$ this means we do not care how P is placed along the x_{12} coordinate, as we only care about how much P intersects line segments of the form $\{x_1\} \times \{x_2\} \times [0,1]$. And conversely, for given a $W\colon [0,1]^2 \to [0,1]$, there are many $P \subseteq [0,1]^2$ satisfying $W^P = W$, e.g., any set of the form $P = \{(x,y,z) : a(x,y) \le z \le b(x,y)\}$ where $b(x,y) - a(x,y) = W(x,y)$.

4.3. Homomorphism densities.

For any tuple of k-uniform hypergraphons $\mathbf{W} = (W_1, \dots, W_m)$, any k-uniform hypergraph F, and any map $\alpha\colon F \to [m]$, define the homomorphism density

$$t_\alpha(F, \mathbf{W}) := \int_{[0,1]^{r(V(F),k-1)}} \prod_{e \in F} W_{\alpha(e)}(\mathbf{x}_{r_<(e)})\, d\mathbf{x}.$$

Example 4.2. If $k = 2$, $F = K_3 = \{12, 13, 23\}$, $\alpha = (12 \mapsto 1, 13 \mapsto 2, 23 \mapsto 3)$, then

$$t_\alpha(F, \mathbf{W}) = \int_{[0,1]^3} W_1(x_1,x_2) W_2(x_1,x_3) W_3(x_2,x_3)\, dx_1 dx_2 dx_3$$

For any symmetric partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0,1]^{r[k]}$, define

$$\mathbf{W}^{\mathcal{P}} := (W^{P_1}, \dots, W^{P_m}) \qquad \text{and} \qquad t_\alpha(F, \mathcal{P}) := t_\alpha(F, \mathbf{W}^{\mathcal{P}}). \tag{9}$$

4.4. Quotient and stepping operators.

Let $W\colon [0,1]^{r_<[k]} \to [0,1]$ be a k-uniform hypergraphon and $\mathcal{Q}$ a symmetric partition of $[0,1]^{r[k-1]}$ into q parts $Q_1, Q_2, \dots, Q_q \subseteq [0,1]^{r[k-1]}$. The *quotient* $W/\mathcal{Q}$ is a $2q^k$-tuple of numbers in $[0,1]$ defined by assigning to each k-tuple $f = (f_1, \dots, f_k) \in [q]^k$ a pair (v_f, w_f), referred to as (volume, average), as follows:

- Volume: v_f equals the integral

$$v_f := \int_{\mathbf{x} \in [0,1]^{r_<[k]}} 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) 1_{Q_{f_2}}(\mathbf{x}_{r([k]\setminus\{2\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})})\, d\mathbf{x}. \tag{10}$$

- Average: If $v_f = 0$, then we set $w_f = 0$. Otherwise, w_f is defined to be

$$w_f := \frac{1}{v_f} \int_{\mathbf{x} \in [0,1]^{r<[k]}} W(\mathbf{x}_{r<[k]}) 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) 1_{Q_{f_2}}(\mathbf{x}_{r([k]\setminus\{2\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})})\, d\mathbf{x}. \tag{11}$$

Intuitively, the partition $\mathcal{Q}$ induces a partition $\mathcal{Q}^*$ of $[0,1]^{r[k]}$ into parts enumerated by $f \in [q]^k$. Each cell of $\mathcal{Q}^*$ has a volume v_f and an average value w_f of W on the cell.

If we have another k-uniform hypergraphon W', and a symmetric partition $\mathcal{Q}'$ of $[0,1]^{r[k-1]}$ into q parts ($\mathcal{Q}$ and $\mathcal{Q}'$ have the same number of parts) with volumes and weights (v'_f, w'_f), we define

$$d_1(W/\mathcal{Q}, W'/\mathcal{Q}') := \sum_{f \in [q]^k} (|v_f - v'_f| + |v_f w_f - v'_f w'_f|). \tag{12}$$

For any symmetric subset $P \subseteq [0,1]^{r[k]}$, we write

$$P/\mathcal{Q} := W^P/\mathcal{Q}.$$

A *$\mathcal{Q}$-step function* $U\colon [0,1]^{r<[k]} \to \mathbb{R}$ is a function of the form

$$U(\mathbf{x}) = \sum_{f=(f_1,\dots,f_k)\in[q]^k} u_f 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) 1_{Q_{f_2}}(\mathbf{x}_{r([k]\setminus\{2\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})}) \tag{13}$$

for some real values u_f. Since $\mathcal{Q}$ is a partition, the indicator functions in (13) all have disjoint support, which together partition the domain $[0,1]^{r<[k]}$. Usually U is a symmetric function, which is equivalent to having an additional symmetry constraint on u_f, namely that $u_f = u_{f'}$ whenever f' is obtained from f' by a permutation of the coordinates.

The *$\mathcal{Q}$-stepping operator*, denoted by a subscript $\mathcal{Q}$, turns a k-uniform hypergraphon W into a symmetric $\mathcal{Q}$-step function $W_\mathcal{Q}$ by averaging over each induced cell of $\mathcal{Q}^*$. More precisely, we define $W_\mathcal{Q}\colon [0,1]^{r<[k]} \to [0,1]$ to be (using v_f and w_f from $W/\mathcal{Q}$ defined earlier)

$$W_\mathcal{Q}(\mathbf{x}) := \sum_{f=(f_1,\dots,f_k)\in[q]^k} w_f 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) 1_{Q_{f_2}}(\mathbf{x}_{r([k]\setminus\{2\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})})$$

We can also apply the stepping operator to a tuple of hypergraphons. If $\mathbf{W} = (W_1, \dots, W_m)$, then

$$\mathbf{W}_\mathcal{Q} := ((W_1)_\mathcal{Q}, \dots, (W_m)_\mathcal{Q}).$$

In particular, if $\mathcal{P} = \{P_1, \dots, P_m\}$ is a partition of $[0,1]^{r[k]}$, then we write

$$\mathbf{W}^\mathcal{P}_\mathcal{Q} := ((W^{P_1})_\mathcal{Q}, \dots, (W^{P_m})_\mathcal{Q}) = (W^{P_1}_\mathcal{Q}, \dots, W^{P_m}_\mathcal{Q})$$

4.5. **Cut norm.**

Definition 4.3. For any symmetric function $W\colon [0,1]^{r<[k]} \to \mathbb{R}$, define

$$\|W\|_{\square^{k-1}} := \sup_{\substack{u_1,\dots,u_k\colon [0,1]^{r[k-1]}\to[0,1] \\ \text{symmetric}}} \left| \int_{[0,1]^{r<[k]}} W(\mathbf{x}_{r<[k]}) \prod_{i=1}^k u_i(\mathbf{x}_{r([k]\setminus\{i\})})\, d\mathbf{x} \right|. \tag{14}$$

Note that by linearity of the expression inside the absolute value in (14), it suffices to consider functions u_i's which are indicator functions 1_{B_i} of symmetric subsets $B_i \subseteq [0,1]^{r[k-1]}$. The usual cut norm corresponds to the case $k = 2$. The following example shows $k = 3$.

Example 4.4. For any symmetric function $W\colon [0,1]^{r<[3]} \to \mathbb{R}$, $\|W\|_{\square^2}$ equals to

$$\sup_{u_1,u_2,u_3} \left| \int_{[0,1]^6} W(x_1,x_2,x_3,x_{12},x_{13},x_{23}) u_1(x_2,x_3,x_{23}) u_2(x_1,x_3,x_{13}) u_3(x_1,x_2,x_{12})\, dx_1 dx_2 dx_3 dx_{12} dx_{13} dx_{23} \right|$$

where u_1, u_2, u_3 vary over all symmetric functions $[0,1]^{r[2]} \to [0,1]$.

5. Regularity and counting lemmas

Definition 5.1. Let W be a k-uniform hypergraphon and $\mathcal{Q}$ a symmetric partition of $[0,1]^{r[k-1]}$. We say that $(W, \mathcal{Q})$ is *weakly ε-regular* if $\|W - W_{\mathcal{Q}}\|_{\square^{k-1}} \le \varepsilon$.

For a symmetric subset $P \subseteq [0,1]^{r[k]}$, we say that $(P, \mathcal{Q})$ is weakly ε-regular if $(W^P, \mathcal{Q})$ is.

Lemma 5.2 (Weak regularity lemma). *Let $k \ge 2$ and $\varepsilon > 0$. Let $\mathbf{W} = (W_1, \dots, W_m)$ be a tuple of k-uniform hypergraphons. Let $\mathcal{Q}$ be a symmetric partition of $[0,1]^{r[k-1]}$. Then there exists a partition $\mathcal{Q}'$ refining $\mathcal{Q}$ so that every part of $\mathcal{Q}$ is refined into exactly $\lceil 2^{km/\varepsilon^2} \rceil$ parts (allowing empty parts) so that $(W_i, \mathcal{Q}')$ is weakly ε-regular for every $1 \le i \le m$.*

Proof. We build the partition incrementally, starting with $\mathcal{Q}$. At a given stage, suppose the partition is $\mathcal{R}$. If $(W_i, \mathcal{R})$ is weakly ε-regular for every i then we stop. Otherwise there is some i with $\|W_i - (W_i)_{\mathcal{R}}\|_{\square^{k-1}} > \varepsilon$, so there exists symmetric subsets $B_1, \dots, B_k \subseteq [0,1]^{r([k-1])}$ such that

$$\left| \int_{[0,1]^{r_<[k]}} (W_i - (W_i)_{\mathcal{R}})(\mathbf{x}_{r_<([k])}) \prod_{i=1}^{k} 1_{B_i}(\mathbf{x}_{r([k]\setminus\{i\})}) \, d\mathbf{x} \right| > \varepsilon. \tag{15}$$

Let $B\colon [0,1]^{r_<[k]} \to [0,1]$ be the function (not necessarily symmetric)

$$B(\mathbf{x}) := \prod_{i=1}^{k} 1_{B_i}(\mathbf{x}_{r([k]\setminus\{i\})}) \, d\mathbf{x}.$$

For two functions $U, U'\colon [0,1]^{r_<[k]} \to [0,1]$, define the inner product

$$\langle U, U' \rangle = \int_{[0,1]^{r_<[k]}} U(\mathbf{x}) U'(\mathbf{x}) \, d\mathbf{x}.$$

We will use the following easy fact: if U' is a $\mathcal{Q}$-step function, then $\langle U, U' \rangle = \langle U_{\mathcal{Q}}, U' \rangle$.

Now let $\mathcal{R}'$ be the the minimal partition refining $\mathcal{R}$ and $B_1, \dots, B_k$. Since $((W_i)_{\mathcal{R}'})_{\mathcal{R}} = (W_i)_{\mathcal{R}}$, applying the fact above, we obtain

$$\langle (W_i)_{\mathcal{R}'}, (W_i)_{\mathcal{R}} \rangle = \langle (W_i)_{\mathcal{R}}, (W_i)_{\mathcal{R}} \rangle \tag{16}$$

Since B is an $\mathcal{R}'$-step function, we have $\langle (W_i)_{\mathcal{R}'}, B \rangle = \langle W_i, B \rangle$. So by (15)

$$|\langle (W_i)_{\mathcal{R}'} - (W_i)_{\mathcal{R}}, B \rangle| = |\langle W_i - (W_i)_{\mathcal{R}}, B \rangle| > \varepsilon. \tag{17}$$

Since $\|U\|_2^2 = \langle U, U \rangle$ for any U, we obtain by (16), the Cauchy-Schwarz inequality, and (17)

$$\|(W_i)_{\mathcal{R}'}\|_2^2 - \|(W_i)_{\mathcal{R}}\|_2^2 = \|(W_i)_{\mathcal{R}'} - (W_i)_{\mathcal{R}}\|_2^2 \ge |\langle (W_i)_{\mathcal{R}'} - (W_i)_{\mathcal{R}}, B \rangle|^2 > \varepsilon^2. \tag{18}$$

Furthermore, for every $1 \le j \le m$, $\|(W_j)_{\mathcal{R}'}\|_2^2 \ge \|(W_j)_{\mathcal{R}}\|_2^2$ by convexity since $((W_j)_{\mathcal{R}'})_{\mathcal{R}} = (W_j)_{\mathcal{R}}$.

The quantity $\|(W_1)_{\mathcal{R}}\|_2^2 + \cdots + \|(W_m)_{\mathcal{R}}\|_2^2$ is at most m, and each iteration above increases the sum by at least ε^2. So there can be at most m/ε^2 iterations. At the end we obtain a partition $\mathcal{Q}'$ so that $(W_i, \mathcal{Q}')$ is weakly ε-regular for every $1 \le i \le m$. Each time we introduced at most k new sets to refine the partition, so $\mathcal{R}'$ refines each part of $\mathcal{R}$ into at most 2^k subparts. After at most m/ε^2 iterations, each part of the original partition $\mathcal{Q}$ is refined into at most $2^{km/\varepsilon^2}$ parts. We can throw in some empty parts so that each part of $\mathcal{Q}$ is refined into exactly $\lceil 2^{km/\varepsilon^2} \rceil$ parts. ■

Lemma 5.3 (Counting lemma I). *Let $\mathbf{U} = (U_1, \dots, U_m)$ and $\mathbf{W} = (W_1, \dots, W_m)$ be two m-tuple of k-uniform hypergraphons and $\mathcal{Q}$ a symmetric partition of $[0,1]^{r([k-1])}$. Suppose that $\|W_i - U_i\|_{\square^{k-1}} \le \varepsilon$ for each i. Then for any k-uniform hypergraph F and any map $\alpha\colon F \to [m]$, we have*

$$|t_\alpha(F, \mathbf{U}) - t_\alpha(F, \mathbf{W})| \le |F| \, \varepsilon.$$

Proof. Let $V = V(F)$ and $F = \{e_1, \dots, e_{|F|}\}$. Write as a telescoping sum

$$\begin{aligned} & t_\alpha(F,\mathbf{U}) - t_\alpha(F,\mathbf{W}) \\ &= \int_{[0,1]^{r(V,k-1)}} \left(\prod_{i=1}^{|F|} U_{\alpha(e_i)}(\mathbf{x}_{r_<(e_i)}) - \prod_{i=1}^{|F|} W_{\alpha(e_i)}(\mathbf{x}_{r_<(e_i)}) \right) d\mathbf{x}_{r(V,k-1)} \\ &= \sum_{j=1}^{|F|} \int_{[0,1]^{r(V,k-1)}} \left(\prod_{i=1}^{j-1} U_{\alpha(e_i)}(\mathbf{x}_{r_<(e_i)}) \right) (U_{\alpha(e_j)} - W_{\alpha(e_j)})(\mathbf{x}_{r_<(e_j)}) \left(\prod_{i=j+1}^{|F|} W_{\alpha(e_i)}(\mathbf{x}_{r_<(e_i)}) \right) d\mathbf{x}. \end{aligned}$$

The j-th term in the final sum is bounded by $\|U_{\alpha(e_j)} - W_{\alpha(e_j)}\|_{\square^{k-1}} \le \varepsilon$. Indeed, if we fix all variables other than $\mathbf{x}_{r_<(e_j)}$, then all the factors except for $(U_{\alpha(e_j)} - W_{\alpha(e_j)})(\mathbf{x}_{r_<(e_j)})$ have the form $u(\mathbf{x}_{r(f)})$ for some $f \subsetneq e_j$, where f is the intersection of e_j with another edge $e_{j'}$. So the the integral can be bounded by the $(k-1)$-cut norm, as claimed. ■

Lemma 5.4 (Counting lemma II). *Let $\mathbf{U} = (U_1, \dots, U_m)$ and $\mathbf{W} = (W_1, \dots, W_m)$ be two m-tuples of k-uniform hypergraphons. Let $\mathcal{Q} = \{Q_1, \dots, Q_q\}$ and $\mathcal{R} = \{R_1, \dots, R_q\}$ be symmetric partitions of $[0,1]^{r[k-1]}$. Suppose that $d_1(U_i/\mathcal{Q}, W_i/\mathcal{R}) \le \delta$ for each i. Then for any k-uniform hypergraph F and any map $\alpha\colon F \to [m]$,*

$$|t_\alpha(F,\mathbf{U}_\mathcal{Q}) - t_\alpha(F,\mathbf{W}_\mathcal{R})| \le |F|\,\delta + \sum_{\beta\colon \partial F \to [q]} |t_\beta(\partial F, \mathcal{Q}) - t_\beta(\partial F, \mathcal{R})|\,,$$

where the sum is taken over all maps $\beta\colon \partial F \to [q]$, and ∂F is the $(k-1)$-uniform hypergraph on $V(F)$ consisting of $(k-1)$-element subsets of $V(F)$ that are contained in some edge of F.

Proof. We can replace each U_i by $(U_i)_\mathcal{Q}$ as this does not change $U_i/\mathcal{Q}$ or $t_\alpha(F,\mathbf{U}_\mathcal{Q})$. So we may assume that every U_i is a symmetric $\mathcal{Q}$-step function, i.e., $\mathbf{U}_\mathcal{R} = \mathbf{U}$. Similarly, assume that every W_i is a symmetric $\mathcal{R}$-step function.

For each $f \in [q]^k$, let $(v_{i,f}, w_{i,f})$ denote the volume and average corresponding to f in $U_i/\mathcal{Q}$, and let $(v'_{i,f}, w'_{i,f})$ denote the same for $W_i/\mathcal{R}$.

For each $1 \le i \le m$, construct a symmetric $\mathcal{Q}$-step function U'_i from U_i by changing its value on the step corresponding to f from $w_{i,f}$ to $w'_{i,f}$. So $U'_i/\mathcal{Q}$ has $(v_{i,f}, w'_{i,f})$ as its volumes and averages. In other words,

$$U_i(\mathbf{x}_{r_<[k]}) = \sum_{f=(f_1,\dots,f_q)\in[q]^k} w_{i,f} 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})}); \tag{19}$$

$$U'_i(\mathbf{x}_{r_<[k]}) = \sum_{f=(f_1,\dots,f_q)\in[q]^k} w'_{i,f} 1_{Q_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})}); \tag{20}$$

$$W_i(\mathbf{x}_{r_<[k]}) = \sum_{f=(f_1,\dots,f_q)\in[q]^k} w'_{i,f} 1_{R_{f_1}}(\mathbf{x}_{r([k]\setminus\{1\})}) \cdots 1_{R_{f_k}}(\mathbf{x}_{r([k]\setminus\{k\})}). \tag{21}$$

Write $\mathbf{U}' = (U'_1, \dots, U'_m)$. We have

$$\begin{aligned} \left\|U_i - U'_i\right\|_1 &= \sum_{f\in[q]^k} v_{i,f}|w_{i,f} - w'_{i,f}| \le \sum_{f\in[q]^k} (|v_{i,f}w_{i,f} - v'_{i,f}w'_{i,f}| + w'_{i,f}|v'_{i,f} - v_{i,f}|) \\ &\le \sum_{f\in[q]^k} (|v_{i,f}w_{i,f} - v'_{i,f}w'_{i,f}| + |v'_{i,f} - v_{i,f}|) = d_1(U_i/\mathcal{Q}, W_i/\mathcal{R}) \le \delta. \end{aligned}$$

So $\|U_i - U'_i\|_1 \le \delta$ for each i. It follows that

$$\left|t_\alpha(F,\mathbf{U}) - t_\alpha(F,\mathbf{U}')\right| \le |F|\,\delta. \tag{22}$$

(This follows from Counting Lemma I, but it's in fact even easier.) From (20) we have

$$t_\alpha(F,\mathbf{U}') = \int_{[0,1]^{r(V(F),k-1)}} \prod_{e=\{j_1,\dots,j_k\}\in F} \left(\sum_{f=(f_1,\dots,f_q)\in[q]^k} w'_{\alpha(e),f} 1_{Q_{f_1}}(\mathbf{x}_{r(e\setminus\{j_1\})}) \cdots 1_{Q_{f_k}}(\mathbf{x}_{r(e\setminus\{j_k\})}) \right) d\mathbf{x}$$

$$= \sum_{\beta\colon \partial F\to[q]} \left(\prod_{e\in F} w'_{\alpha(e),\beta(\partial e)} \right) t(\partial F, \mathcal{Q}). \tag{23}$$

Here $\beta(\partial e) = (\beta(e\setminus\{j_1\}),\dots,\beta(e\setminus\{j_k\})) \in [q]^k$ when $e = \{j_1,\dots,j_k\}$. The last equality above needs some pondering. Essentially we expand the product of sums in the previous line and note that since $\mathcal{Q}$ is a partition, the nonzero terms in the expansion correspond to assigning an f to every e in a compatible way: if two edges $e = \{j_1,\dots,j_k\}$ and $e' = e\cup\{j'_l\}\setminus\{j_l\}$ intersect in exactly $k-1$ vertices, and f is assigned to e, and f' is assigned to e', then $f_l = f'_l$. These assignments are in bijection with $\beta\colon \partial F\to[q]$, where β corresponds to the assignment assigning e to $\beta(\partial e)$.

Similar to (23) we have

$$t_\alpha(F,\mathbf{W}) = \sum_{\beta\colon \partial F\to[q]} \left(\prod_{e\in F} w'_{\alpha(e),\beta(\partial e)} \right) t(\partial F, \mathcal{R}). \tag{24}$$

Combing (23) and (24) using the triangle inequality and noting that $0 \le w'_{i,f} \le 1$, we have

$$\left|t_\alpha(F,\mathbf{U}') - t_\alpha(F,\mathbf{W})\right| \le \sum_{\beta\colon \partial F\to[q]} \left|t_\beta(\partial F,\mathcal{Q}) - t_\beta(\partial F,\mathcal{R})\right|. \tag{25}$$

The lemma follows from combining (22) and (25) using the triangle inequality. ■

6. Branching partitions

Now we are almost ready to build the limiting object. We will proceed by induction on k (for k-uniform hypergraphons). The situation is very simple when $k=1$, since in this case a hypergraphon is simply a number between 0 and 1. To build the limiting hypergraphon in general, we will need to repeatedly apply the weak regularity lemma to obtain a refining chain of partitions. Since we need to apply induction on k, we need to have a stronger induction hypothesis that involves a sequence of not just single hypergraphons, but refining chains of partitions. This motivates the following definition of a branching partition, which is a special case a *filtration*, in the language of probability. See Figure 1.

Definition 6.1. A *degree* $p = (p_1,p_2,\dots) \in \mathbb{N}^{\mathbb{N}}$ *(symmetric) branching partition* $\mathscr{P}$ of $[0,1]^{r[k]}$ is a collection of symmetric subsets P_i of $[0,1]^{r[k]}$, collected into *levels*, where each level $\mathcal{P}_l$ is a symmetric partition of $[0,1]^{r[k]}$:

- Level 0: $\mathcal{P}_0 = \{[0,1]^{r[k]}\}$
- Level 1: $\mathcal{P}_1 = \{P_1,P_2,\dots,P_{p_1}\}$ is a symmetric partition of $[0,1]^{r[k]}$.
- Level l ($l\ge 2$): $\mathcal{P}_l$ is a refinement of $\mathcal{P}_{l-1}$, where each part of $\mathcal{P}_{l-1}$ gets further refined into exactly p_l parts.

An *index* at level l is a tuple $i = (i_1,i_2,\dots,i_l) \in [p_1]\times[p_2]\times\cdots\times[p_l]$, which points to the symmetric subset $P_i = P_{i_1,\dots,i_l} \in \mathcal{P}_l$ at level l, where P_i is the i_l-th part in the refinement of the part $P_{i_1,\dots,i_{l-1}}$ at level $l-1$, whenever $l\ge 2$ (all partitions are ordered).

Font convention. $\mathscr{P}$ is a branching partition, $\mathcal{P}$ is a partition, and P is a subset of $[0,1]^{r[k]}$.

Example 6.2. A symmetric subset $P \subseteq [0,1]^{r[k]}$ or a k-uniform hypergraphon W (related by (8)) can be thought of as a degree $(2,1,1,1,\dots)$ branching partition: level 1 is P and P^c (the complement of P in $[0,1]^{r[k]}$) and all subsequent levels are trivial refinements.

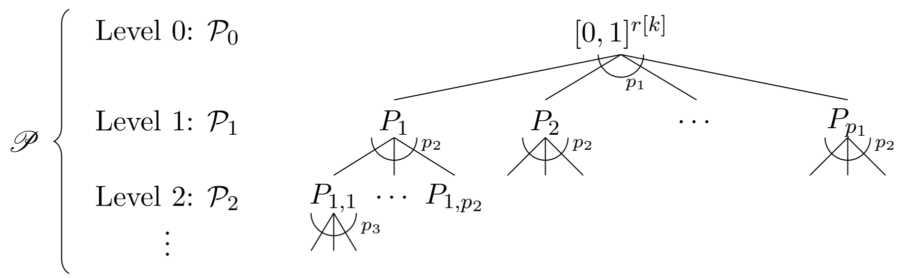

Figure 1. A branching partition

We can generalize the notion of regularity from Definition 5.1 to branching partitions as follows.

Definition 6.3. Let $\mathscr{P}$ be a branching partition of $[0,1]^{r[k]}$ and $\mathscr{Q}$ a branching partition of $[0,1]^{r[k-1]}$. We say that $(\mathscr{P},\mathscr{Q})$ is *weakly* $(\varepsilon_1,\varepsilon_2,\dots)$*-regular* if for every $s \geq 1$, whenever $P \subseteq [0,1]^{r[k]}$ is a member of $\mathscr{P}$ of level at most s, and $\mathcal{Q}_s$ is the level s partition of $[0,1]^{r[k-1]}$ in $\mathscr{Q}$, the pair $(P,\mathcal{Q}_s)$ is weakly ε_s-regular.

Lemma 6.4 (Weak regularity lemma for branching partitions)**.** *For every $k \geq 2$, $p = (p_1,p_2,\dots) \in \mathbb{N}^{\mathbb{N}}$ and $\varepsilon = (\varepsilon_1,\varepsilon_2,\dots) \in \mathbb{R}^{\mathbb{N}}_{>0}$, we can find a $q = (q_1,q_2,\dots) \in \mathbb{N}^{\mathbb{N}}$ so that the following holds: for every degree p branching partition $\mathscr{P}$ of $[0,1]^{r[k]}$, there exists a degree q branching partition $\mathscr{Q}$ of $[0,1]^{r[k-1]}$ so that $(\mathscr{P},\mathscr{Q})$ is weakly ε-regular.*

Proof. Take $q_s = \lceil 2^{k(p_1+p_1p_2+\cdots+p_1p_2\cdots p_s)/\varepsilon_s^2}\rceil$. We build $\mathscr{Q}$ successively by level. To obtain the level s partition in $\mathscr{Q}$, applying Lemma 5.2 with $\varepsilon = \varepsilon_s$, $\mathbf{W}$ the collection of hypergraphons corresponding to all members of $\mathscr{P}$ of level at most s, and $\mathcal{Q}$ the level $s-1$ partition in $\mathscr{Q}$. ∎

Now we introduce two notions of convergence for branching partitions. The first notion, called *left-convergence*, is based on convergence of homomorphism densities. The second notion, called *partitionable convergence*, is based on convergence of regularity partitions. We will show, using our counting lemmas, that partitionable convergence implies left-convergence.

Notation. Given degree $p = (p_1,p_2,\dots)$ branching partitions $\mathscr{P}_1,\mathscr{P}_2,\dots$ and $\widetilde{\mathscr{P}}$ of $[0,1]^{r[k]}$ and degree $q = (q_1,q_2,\dots)$ branching partitions $\mathscr{Q}_1,\mathscr{Q}_2,\dots$ and $\widetilde{\mathscr{Q}}$ of $[0,1]^{r[k-1]}$, we use the following notation to refer to the partitions and parts in these branching partitions.

- For each $l \geq 1$, $\mathcal{P}_{n,l}$ is the level l partition in $\mathscr{P}_n$, and $\widetilde{\mathcal{P}}_l$ is the level l partition in $\widetilde{\mathscr{P}}$.
- For each $s \geq 1$, $\mathcal{Q}_{n,s}$ is the level s partition in $\mathscr{Q}_n$, and $\widetilde{\mathcal{Q}}_s$ is the level s partition in $\widetilde{\mathscr{Q}}$.
- For each index $i = (i_1,i_2,\dots,i_l) \in [p_1]\times\cdots\times[p_l]$, $P_{n,i}$ is the index i element of $\mathscr{P}_n$ and $\widetilde{P}_i$ is the index i element of $\widetilde{\mathscr{P}}$.

Definition 6.5 (Left-convergence: $\mathscr{P}_n \to \widetilde{\mathscr{P}}$)**.** We say that a sequence $\mathscr{P}_1,\mathscr{P}_2,\dots$ of degree p branching partitions of $[0,1]^{r[k]}$ *left-converges* to another degree p branching partition $\widetilde{\mathscr{P}}$ of $[0,1]^{r[k]}$, written $\mathscr{P}_n \to \widetilde{\mathscr{P}}$, if

$$\lim_{n\to\infty} t_\alpha(F,\mathcal{P}_{n,l}) = t_\alpha(F,\widetilde{\mathcal{P}}_l) \qquad \text{for all } F,l,\alpha \tag{26}$$

where F ranges over all k-uniform hypergraphs, l ranges over all positive integers, and α ranges over all maps $F \to [p_1\cdots p_l]$. Recall from (9) that $t_\alpha(F,\mathcal{P}) := t_\alpha(F,\mathbf{W}^{\mathcal{P}})$ for a partition $\mathcal{P}$.

Definition 6.6 (Partitionable convergence: $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$)**.** We say that a sequence $\mathscr{P}_1,\mathscr{P}_2,\dots$ of degree $p = (p_1,p_2,\dots)$ branching partitions of $[0,1]^{r[k]}$ *partitionably converges* to another degree p branching partition $\widetilde{\mathscr{P}}$ of $[0,1]^{r[k]}$, written $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$, if the following is satisfied (the definition is inductive on k).

When $k = 1$, for every index $i = (i_1, \dots, i_l) \in [p_1] \times \cdots \times [p_l]$, we have $\lim_{n\to\infty} \lambda(P_{n,i}) = \lambda(\widetilde{P}_i)$, where λ is the Lebesgue measure on $[0, 1]$.

When $k \geq 2$, there exists some $q \in \mathbb{N}^{\mathbb{N}}$ and degree q branching partitions $\mathscr{Q}_1, \mathscr{Q}_2, \dots$ and $\widetilde{\mathscr{Q}}$ of $[0, 1]^{r[k-1]}$ satisfying:

(a) $(\mathscr{P}_n, \mathscr{Q}_n)$ is weakly $(1, 1/2, 1/3, \dots)$-regular for every n;
(b) $\mathscr{Q}_n \dashrightarrow \widetilde{\mathscr{Q}}$ as $n \to \infty$ (defined inductively);
(c) For every $s \geq 1$ and every index $i \in [p_1]\times\cdots\times[p_l]$, one has $\lim_{n\to\infty} d_1(P_{n,i}/\mathcal{Q}_{n,s}, \widetilde{P}_i/\widetilde{\mathcal{Q}}_s) = 0$;
(d) For every member $\widetilde{P} \subseteq [0, 1]^{r[k]}$ of $\widetilde{\mathscr{P}}$, one has $(W^{\widetilde{P}})_{\widetilde{\mathcal{Q}}_s} \to W^{\widetilde{P}}$ pointwise almost everywhere as $s \to \infty$.

Lemma 6.7 (Partitionable convergence implies left-convergence)**.** *If $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$ then $\mathscr{P}_n \to \widetilde{\mathscr{P}}$.*

Proof. We use induction on k. When $k = 1$, the claim is trivial. Now assume $k \geq 2$.

We need to show that (26) holds. Fix F, l, α. Let $m = p_1 \cdots p_s$. Let $\mathscr{Q}_n$ and $\widetilde{\mathscr{Q}}$ be as in Definition 6.6, and let $q = (q_1, q_2, \dots)$ be the degree of $\widetilde{\mathscr{Q}}$.

Let $\varepsilon > 0$. By Definition 6.6(d), $\mathbf{W}^{\widetilde{\mathcal{P}}_l}_{\widetilde{\mathcal{Q}}_s}$ converges pointwise almost everywhere in each coordinate to $\mathbf{W}^{\widetilde{\mathcal{P}}_l}$ as $s \to \infty$, so $\lim_{s\to\infty} t_\alpha(F, \mathbf{W}^{\widetilde{\mathcal{P}}_l}_{\widetilde{\mathcal{Q}}_s}) = t_\alpha(F, \mathbf{W}^{\widetilde{\mathcal{P}}_l})$. We can find an $s \geq \max\{l, |F| /\varepsilon\}$ so that $|t_\alpha(F, \mathbf{W}^{\widetilde{\mathcal{P}}_l}_{\widetilde{\mathcal{Q}}_s}) - t_\alpha(F, \mathcal{P}_l)| \leq \varepsilon$. Fix this value of s.

By Definition 6.6(b) we have $\mathscr{Q}_n \dashrightarrow \widetilde{\mathscr{Q}}$, so $\mathscr{Q}_n \to \mathscr{Q}$ by the induction hypothesis. Thus

$$\lim_{n\to\infty} t_\beta(\partial F, \mathcal{Q}_{n,s}) = t_\beta(\partial F, \widetilde{\mathcal{Q}}_s) \tag{27}$$

for all $\beta \colon \partial F \to [q_1 q_2 \cdots q_s]$. See Lemma 5.4 for the definition of ∂F. We have

$$\begin{aligned} &|t_\alpha(F, \mathcal{P}_{n,l}) - t_\alpha(F, \widetilde{\mathcal{P}}_l)| \\ &\leq |t_\alpha(F, \mathcal{P}_{n,l}) - t_\alpha(F, \mathbf{W}^{\mathcal{P}_{n,l}}_{\mathcal{Q}_{n,s}})| + |t_\alpha(F, \mathbf{W}^{\mathcal{P}_{n,l}}_{\mathcal{Q}_{n,s}}) - t_\alpha(F, \mathbf{W}^{\widetilde{\mathcal{P}}_l}_{\widetilde{\mathcal{Q}}_s})| + |t_\alpha(F, \mathbf{W}^{\widetilde{\mathcal{P}}_l}_{\widetilde{\mathcal{Q}}_s}) - t_\alpha(F, \widetilde{\mathcal{P}}_l)| \end{aligned} \tag{28}$$

As $n \to \infty$, the first term on the right hand side of (28) has a limsup of at most $|F| /s \leq \varepsilon$ by Counting Lemma I (Lemma 5.3) since $(P, \mathcal{Q}_{n,s})$ is $1/s$-regular for every $P \in \mathcal{P}_{n,l}$ by Definition 6.6(a). The second term on the RHS of (28) goes to zero by Counting Lemma II (Lemma 5.4), Definition 6.6(c), and (27). The third term on the RHS of (28) is at most ε using our choice of s. It follows that $\limsup_{n\to\infty} |t_\alpha(F, \mathcal{P}_{n,l}) - t_\alpha(F, \widetilde{\mathcal{P}}_l)| \leq 2\varepsilon$. Since ε can be made arbitarily small, we obtain $\lim_{n\to\infty} t_\alpha(F, \mathcal{P}_{n,l}) = t_\alpha(F, \widetilde{\mathcal{P}}_l)$ as desired. ■

Proposition 6.8. *Let $p \in \mathbb{N}^{\mathbb{N}}$. Let $\mathscr{P}_1, \mathscr{P}_2 \cdots$ be a sequence of degree p branching partitions of $[0, 1]^{r[k]}$. Then there exists another degree p branching partition $\widetilde{\mathscr{P}}$ of $[0, 1]^{r[k]}$ so that $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$ as $n \to \infty$ along some infinite subsequence.*

Proof. We use induction on k. The claim is easy when $k = 1$, since we can pick a subsequence so that for each index i, the measure $\lambda(P_{n,i})$ converges to some value a_i as $n \to \infty$, and we can take the limit $\widetilde{\mathscr{P}}$ to be a branching partition where $\widetilde{P}_i$ is an interval with length a_i.

Now assume $k \geq 2$. By Lemma 6.4, there exists a $q \in \mathbb{N}^{\mathbb{N}}$ so that for every n we can find a degree q branching partition $\mathscr{Q}_n$ of $[0, 1]^{r[k-1]}$ so that $(\mathscr{P}_n, \mathscr{Q}_n)$ is weakly $(1, 1/2, 1/3, \dots)$-regular, thereby satisfying (a) in Definition 6.6. Applying the induction hypothesis, we can restrict to a subsequence so that $\mathscr{Q}_n \dashrightarrow \widetilde{\mathscr{Q}}$ for some branching partition $\widetilde{\mathscr{Q}}$ of $[0, 1]^{r[k-1]}$ (here and onwards in this proof we abuse notation by only considering convergence as $n \to \infty$ along some subsequence. We will be repeatedly taking subsequences, and the conclusion will follow by a standard diagonalization argument). So (b) is satisfied.

By further restricting to a subsequence, we may assume that for each $s \geq 1$ and each index i, the quotient $P_{n,i}/\mathcal{Q}_{n,s}$ converges coordinate-wise as $n \to \infty$. Let $W_{n,i} := W^{P_{n,i}}$ be the hypergraphon associated to $P_{n,i}$. Let $\widetilde{W}_{i,s}\colon [0,1]^{r<[k]} \to [0,1]$ be a symmetric $\widetilde{\mathcal{Q}}_s$-step function, with values assigned so that $d_1(W_{n,i}/\mathcal{Q}_{n,s}, \widetilde{W}_{i,s}/\widetilde{\mathcal{Q}}_s) \to 0$ as $n \to \infty$. This is possible since we previously assumed that $P_{n,i}/\mathcal{Q}_{n,s}$ converges coordinatewise as $n \to \infty$, so that are now simply putting in the limiting values of the "average" coordinates into a template for a symmetric $\widetilde{\mathcal{Q}}_s$-step function in order to construct $\widetilde{W}_{i,s}$. To see that the "volume" coordinates (10) of $\mathcal{Q}_{n,s}$ converge to those of $\widetilde{\mathcal{Q}}_s$, note that this amount to the claim that $\lim_{n\to\infty} t_\beta(K_k^{(k-1)}, \mathcal{Q}_{n,s}) = t_\beta(K_k^{(k-1)}, \widetilde{\mathcal{Q}}_s)$ for every $\beta\colon K_k^{(k-1)} \to [q]$, where $K_k^{(k-1)}$ is the $(k-1)$-uniform simplex, i.e., the collection of all $(k-1)$-element subsets of $[k]$. The convergence of these homomorphism densities follows from $\mathscr{Q}_n \to \widetilde{\mathscr{Q}}$ which in turn follows from $\mathscr{Q}_n \dashrightarrow \widetilde{\mathscr{Q}}$ and Lemma 6.7.

Claim 1. $(\widetilde{W}_{i,s+1})_{\widetilde{\mathcal{Q}}_s} = \widetilde{W}_{i,s}$.

Proof of Claim 1. We have

$$\lim_{n\to\infty} d_1(W_{n,i}/\mathcal{Q}_{n,s}, \widetilde{W}_{i,s}/\widetilde{\mathcal{Q}}_s) = 0 \tag{29}$$

and

$$\lim_{n\to\infty} d_1(W_{n,i}/\mathcal{Q}_{n,s+1}, \widetilde{W}_{i,s+1}/\widetilde{\mathcal{Q}}_{s+1}) = 0 \tag{30}$$

Since $\widetilde{\mathcal{Q}}_{s+1}$ is a refinement of $\widetilde{\mathcal{Q}}_s$, by merging together parts in $W_{n,i}/\mathcal{Q}_{s+1}$ and $\widetilde{W}_{i,s+1}/\widetilde{\mathcal{Q}}_{s+1}$, we deduce from (30)

$$\lim_{n\to\infty} d_1(W_{n,i}/\mathcal{Q}_{n,s}, \widetilde{W}_{i,s+1}/\widetilde{\mathcal{Q}}_s) = 0 \tag{31}$$

From (29) and (31) we obtain $\widetilde{W}_{i,s+1}/\widetilde{\mathcal{Q}}_s = \widetilde{W}_{i,s}/\widetilde{\mathcal{Q}}_s$, which implies $(\widetilde{W}_{i,s+1})_{\widetilde{\mathcal{Q}}_s} = \widetilde{W}_{i,s}$ since both sides are $\widetilde{\mathcal{Q}}_s$-step functions □

It follows that $\widetilde{W}_{i,1}, \widetilde{W}_{i,2}, \widetilde{W}_{i,3}, \dots$ is a martingale with respect to the filtration[3] induced by $\widetilde{\mathcal{Q}}_1, \widetilde{\mathcal{Q}}_2, \dots$. By the Martingale Convergence Theorem, there exists some $\widetilde{W}_i$, so that $\widetilde{W}_{i,s} \to \widetilde{W}_i$ pointwise almost everywhere as $s \to \infty$. Furthermore $(\widetilde{W}_i)_{\widetilde{\mathcal{Q}}_s} = \widetilde{W}_{i,s}$.

Claim 2. Let $l \geq 1$, and let $i = (i_1, \dots, i_{l-1}) \in [p_1] \times \cdots \times [p_{l-1}]$ an index at level $l-1$, which points to a part in $\mathcal{P}_1$ that splits into indices $\{j_1, \dots, j_{p_l}\} = i \times [p_l]$ at level l. Then

$$\widetilde{W}_{j_1} + \cdots + \widetilde{W}_{j_{p_l}} = \widetilde{W}_i \quad \text{almost everywhere.}$$

Proof of Claim 2. Since $P_{n,j_1}, \dots, P_{n,j_{p_s}}$ is a partition of $P_{n,i}$, we have

$$W_{n,j_1} + \cdots + W_{n,j_{p_s}} = W_{n,i}$$

Taking the $\mathcal{Q}_{n,s}$ quotient of both sides and then take the limit as $n \to \infty$, we find the following equality for these $\widetilde{\mathcal{Q}}_s$-step functions.

$$\widetilde{W}_{j_1,s} + \cdots + \widetilde{W}_{j_s,s} = \widetilde{W}_{i,s}$$

Taking $s \to \infty$ and using the pointwise almost everywhere convergence of $\widetilde{W}_{j,s} \to \widetilde{W}_j$ as $s \to \infty$ for every index j, we obtain Claim 2. □

Claim 2 tells us that we can find a branching partition $\widetilde{\mathscr{P}}$ of $[0,1]^{r[k]}$ so that the part $\widetilde{P}_i$ satisfies $W^{\widetilde{P}_i} = \widetilde{W}_i$. Visually we can build the level s of $\widetilde{\mathscr{P}}$ by stacking together subsets of $[0,1]^{r[k]}$ that correspond to $\widetilde{W}_j$, ranged over all indices j at level s. Then $P_{n,i}/\mathcal{Q}_{n,s} = W_{n,i}/\mathcal{Q}_{n,s} \to_{d_1} \widetilde{W}_{i,s}/\widetilde{\mathcal{Q}}_s =$

[3] To be more precise, let $[0,1]^{r[k]}$ be the probability space equipped with the uniform Lebesgue measure. For each $s \geq 1$ let $\mathcal{B}_s$ be the minimal σ-algebra on $[0,1]^{r[k]}$ generated by functions of the form $1_Q(\mathbf{x}_{r([k]\setminus\{j\})})$ ranged over $Q \in \mathcal{Q}_s$ and $j \in [k]$. Then $\widetilde{W}_{i,s}$ is a $\mathcal{B}_s$-measurable random variable, and Claim 1 implies that $\widetilde{W}_{i,1}, \widetilde{W}_{i,2}, \cdots$ is a martingale adapted to the filtration $\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \cdots$

$\widetilde{W}_i/\widetilde{\mathscr{Q}}_s = \widetilde{P}_i/\widetilde{\mathscr{Q}}_s$, so (c) is satisfied. Also (d) is satisfied since $W^{\widetilde{P}_i}_{\widetilde{\mathscr{Q}}_s} = \widetilde{W}_{i,s} \to \widetilde{W}_i = W^{\widetilde{P}_i}$ pointwise almost everywhere as $s \to \infty$ (from our application of the Martingale Convergence Theorem). ■

Proof of Theorem 1.6. Let $\mathscr{P}_n$ be the degree $(2, 1, 1, 1, \dots)$ branching partition built from W_n as in Example 6.2. Proposition 6.8 implies that there exists a branching partition $\widetilde{\mathscr{P}}$ so that $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$ along a subsequence, and hence $\mathscr{P}_n \to \widetilde{\mathscr{P}}$ along a subsequence by Lemma 6.7. Let $\widetilde{P}$ be the index (1) element of $\widetilde{\mathscr{P}}$. The associated hypergraphon $\widetilde{W} = W^{\widetilde{P}}$ is the desired limit of W_n. By applying (26) with $l = 1$ and $\alpha \equiv 1$, we see that $t(F, W_n) \to t(F, \widetilde{W}_n)$ along the subsequence. ■

We conclude the paper with a conjecture that partitionable convergence is equivalent to left-convergence, thereby proposing a converse to Lemma 6.7.

Conjecture 6.9. *$\mathscr{P}_n \to \widetilde{\mathscr{P}}$ if and only if $\mathscr{P}_n \dashrightarrow \widetilde{\mathscr{P}}$.*

Acknowledgments

The author would like to thank Jacob Fox, László Lovász, Jennifer Chayes and Christian Borgs for helpful conversations, and also the anonymous referees for helpful comments and for pointing out the connections between graph limits and exchangeable random arrays.

References

[1] D. J. Aldous. Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.*, 11(4):581–598, 1981.

[2] D. J. Aldous. Exchangeability and continuum limits of discrete random structures. In *Proceedings of the International Congress of Mathematicians. Volume I*, pages 141–153, New Delhi, 2010. Hindustan Book Agency.

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

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

[8] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[9] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)*, 166(3):897–946, 2007.

[10] D. N. Hoover. Relations on probability spaces and arrays of random variables, 1979. Preprint, Institute for Advanced Study, Princeton, NJ.

[11] O. Kallenberg. Symmetries on random arrays and set-indexed processes. *J. Theoret. Probab.*, 5(4):727–765, 1992.

[12] L. Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.

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

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

[15] B. Nagle, V. Rödl, and M. Schacht. The counting lemma for regular k-uniform hypergraphs. *Random Structures Algorithms*, 28(2):113–179, 2006.

[16] T. Tao. A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A*, 113(7):1257–1280, 2006.

[17] T. Tao. A correspondence principle between (hyper)graph theory and probability theory, and the (hyper)graph removal lemma. *J. Anal. Math.*, 103:1–45, 2007.

[18] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Department of Mathematics, MIT, Cambridge, MA 02139-4307.
E-mail address: yufeiz@math.mit.edu