On the Complexity of Nondeterministically Testable Hypergraph Parameters

Marek Karpinski* Roland Markó†

Abstract

The paper proves the equivalence of the notions of nondeterministic and deterministic parameter testing for uniform dense hypergraphs of arbitrary order. It generalizes the result previously known only for the case of simple graphs. By a similar method we establish also the equivalence between nondeterministic and deterministic hypergraph property testing, answering the open problem in the area. We introduce a new notion of a cut norm for hypergraphs of higher order, and employ regularity techniques combined with the ultralimit method.

1 Introduction

Hypergraph parameters are real-valued functions defined on the space of uniform hypergraphs of some given order invariant under relabeling the vertex set. Testing a parameter value associated to an instance in the dense model means to produce an estimation by only having access to a small portion of the data that describes it. The test data is selected by choosing a uniform random subset of the vertex set and exposing the induced substructure of the hypergraph on this subset. A certain parameter is said to be testable if for every given tolerated error the estimation is within the error range of the parameter value with high probability, and the size of the selected random subset does only depend on the size of this permitted error and not on the size of the instance, precise definitions are provided below. Similar notions apply to testing graph properties, in that situation one also uses uniform sampling in order to separate the cases where an instance has the property or is far from having it, where the distance is measured by the number of edge modifications required. For the related notions of approximation theory and limits see [1], [3], and [4]. The general reader is referred to [9], [11], and [13] for some related developments.

*Dept. of Computer Science and the Hausdroff Center for Mathematics, University of Bonn. Supported in part by DFG grants, the Hausdorff grant EXC59-1/2. Research partly supported by Microsoft Research New England. E-mail: marek@cs.uni-bonn.de

†Hausdorff Center for Mathematics, University of Bonn. Supported in part by a Hausdorff scholarship. E-mail: roland.marko@hcm.uni-bonn.de

The notion nondeterministic testability was introduced by Lovász and Vesztergombi [14] in the framework of graph property testing, and encompasses an a priori weaker characteristic than the original testability. They defined that a certain property is nondeterministically testable if there exists another property of colored (edge or node) graphs that is testable in the normal sense and serves as a certificate for the original. It was shown by the authors of [14] that for graph properties the two notions are equivalent, demonstrating that if a property is nondeterministically testable, then it is also testable. Their proof used the machinery of graph limits and for this reason it was of non-effective nature. Subsequently, an explicit construction of a tester was given by Gishboliner and Shapira [8] for nondeterministically testable graph properties containing the tester of the colored witness property as a subroutine. They used Szemerédi's Regularity Lemma combined with developments by Alon et al. [2], and provided a tower-type dependence between the sample complexity of the investigated property the sample complexity of the witness property.

In [14], additionally the study nondeterministic testing for parameters was initiated, the definition is similar to the property testing situation. A different approach by Karpinski and Markó [10] relying on weaker regularity methods led to an effective upper bound on the sample complexity that is a 3-fold iteration of the exponential function applied to the sample size required by the witness parameter.

The previous works mentioned above dealt with graphs, it was asked in [14] if the concept can be employed for hypergraphs. The notion of an r-uniform hypergraph (in short, r-graph) parameter and its testability can be defined completely analogously to the graph case, the same applies for nondeterministic testability. Naturally, first the question arises whether or not the deterministic and the nondeterministic testability are equivalent for higher order hypergraphs, and secondly, if the answer to the first question is positive, then what can be said about the relationship of the sample complexity of the parameter and that of its witness parameter. The statements that are analogous to the main results of [8], [10], and [14] do not follow immediately for uniform hypergraphs of higher order from the proof for graphs, like-wise to the generalizations of the Regularity Lemma new tools and notions are required to handle these cases. In the current paper we prove the equivalence of the two testability notions for uniform hypergraphs of higher order and settle the first question posed above. Unfortunately, we were not able to obtain an explicit upper bound for the sample complexity, this is the consequence of us applying of the limit theory for hypergraphs developed by Elek and Szegedy [6] using methods of non-standard analysis, therefore the second problem still remains open. We also show that testing nondeterministically testable properties is as hard as parameter testing with our method in the sense that the same complexity bounds apply.

The paper is organized as follows. In Section 2 we give the preliminaries required and formulate the precise definitions followed by our main result, Theorem 2.3. Section 3 contains the testability results for r-cut norms together with a brief summary of the notions and results regarding the ultralimit method that are needed for our purposes, Section 4 comprises some auxiliary results required. Section 5 describes the proof of our main result.

In Section 6 we give an application for property testing of hypergraphs, and in Section 7 we pose some questions related to possible further research.

2 Preliminaries and main result

A simple r-uniform hypergraph on n vertices is a subset G of $\binom{[n]}{r}$, the size of G is n, and the elements of $\binom{[n]}{r}$ are r-edges. Let A_G denote the symmetric $\{0,1\}$-valued r-array or symmetric subset of $[n]^r$ that represents G, we will sometimes use also only the term G to refer to a symmetric subset of $[n]^r \setminus \mathrm{diag}([n]^r)$ corresponding to the array representation. Let k be a positive integer, and let $\mathcal{G}_n^{r,k}$ denote the set of *k-colored* r-uniform hypergraphs of size n, that are partitions $\mathbf{G} = (G^\alpha)_{\alpha\in[k]}$ of $\binom{[n]}{r}$ into k classes, so in all what follows here colored r-graph means a complete r-graph where to each edge e we assign exactly one color $\mathbf{G}(e)$ from the set $[k]$. In this sense simple r-graphs are regarded as 2-colored. In the k-colored case is also possible to speak about the array representation, A_{G^α} stands for the symmetric $\{0,1\}$-valued r-array that represents the color class of α, again with slight abuse of notation we will use G^α for A_{G^α}. Additionally we have to introduce the color ι and the corresponding array A_ι that always is the indicator array of the set of diagonal elements of $[n]^r$ (those having repetitions in their coordinates, denoted by $\mathrm{diag}([n]^r)$). For any finite set C the term C-colored graph is defined analogously.

A *k-coloring* of a t-colored r-graph $\mathbf{G} = (G^\alpha)_{\alpha\in[t]}$ is a tk-colored r-graph $\hat{\mathbf{G}} = (G^{(\alpha,\beta)})_{\alpha\in[t],\beta\in[k]}$ with colors from the set $[t]\times[k]$, where each of the original color classes indexed by $\alpha \in [t]$ is retrieved by taking the union of the new classes corresponding to (α,β) over all $\beta \in [k]$, that is $G^\alpha = \cup_{\beta\in[k]} G^{(\alpha,\beta)}$. This last operation is called *k-discoloring* of a $[t]\times[k]$-colored graph, we denote it by $[\hat{\mathbf{G}}, k] = \mathbf{G}$. We will sometimes write *tk-colored* for $[t]\times[k]$-colored graphs when it is clear from the context what we mean.

Further, for a finite set S, let $\mathrm{h}(S)$ denote the set of nonempty subsets of S, and $\mathrm{h}(S,m)$ the set of nonempty subsets of S of cardinality at most m. A real 2^r-1-dimensional vector $x_{\mathrm{h}(S)}$ denotes $(x_{T_1},\dots,x_{T_{2^r-1}})$, where $T_1,\dots,T_{2^r-1}$ is a fixed ordering of the nonempty subsets of S with $T_{2^r-1} = S$, for a permutation π of the elements of S the vector $x_{\pi(\mathrm{h}(S))}$ means $(x_{\pi'(T_1)},\dots,x_{\pi'(T_{2^r-1})})$, where π' is the action on the subsets of S induced by π.

We will require some basic notation from graph limit theory, and we summarize their relevance outlined in previous works, Lovász [11] is a comprehensive reference for the area.

Let $q \geq 1$ and $\mathbf{G} \in \mathcal{G}_n^{r,k}$, then $\mathbb{G}(q,\mathbf{G})$ denotes the random r-graph on q vertices that is obtained by uniformly picking a random subset S of $[n]$ of cardinality q and taking the induced subgraph $\mathbf{G}[S]$. For any $\mathbf{F} \in \mathcal{G}_q^{r,k}$ and $\mathbf{G} \in \mathcal{G}^{r,k}$ the $\mathbf{F}$-density of $\mathbf{G}$ is defined as $t(\mathbf{F},\mathbf{G}) = \mathbb{P}(\mathbf{F} = \mathbb{G}(q,\mathbf{G}))$.

Let the r-kernel space $\mathcal{W}_0^r$ denote the space of the bounded measurable functions $W\colon [0,1]^{\mathrm{h}([r],r-1)} \to \mathbb{R}$, and the subspace $\mathcal{W}^r$ of $\mathcal{W}_0^r$ symmetric r-kernels that are invariant under coordinate permutations induced by $\pi \in S_r$, that is $W(x_{\mathrm{h}([r],r-1)}) = W(x_{\pi(\mathrm{h}([r],r-1))})$ for each $\pi \in S_r$. We will refer to this invariance in the paper both for r-kernels and for

measurable subsets of $[0,1]^{\mathrm{h}([r])}$ as *satisfying the usual symmetries*. Assume that the functions $W \in \mathcal{W}_I^r$ take their values in the interval I, for $I = [0,1]$ we call these special symmetric r-kernels *r-graphons*. In what follows, λ always denotes the usual Lebesgue measure in $\mathbb{R}^d$, where d is everywhere clear from the context.

Analogously to the graph case we define the space of *k-colored r-graphons* $\mathcal{W}^{r,k}$ whose elements are referred to as $\mathbf{W} = (W^\alpha)_{\alpha\in[k]}$ with each of the W^α's being an r-graphon. The special color ι that stands for the absence of any colors in the diagonal in some sense can be also employed in this setting, see below for the case when we represent a k-colored r-graph as a graphon. The corresponding r-graphon W^ι is $\{0,1\}$-valued. Furthermore $\sum_{\alpha\in[k]} W^\alpha(x) = 1 - W^\iota(x)$ everywhere on $[0,1]^{\mathrm{h}([r],r-1)}$. For $x \in [0,1]^{\mathrm{h}([r])}$ the expression $\mathbf{W}(x)$ denotes the color at x, we have $\mathbf{W}(x) = \alpha$ whenever $\sum_{i=1}^{\alpha-1} W^i(x_{\mathrm{h}([r],r-1)}) \le x_{[r]} \le \sum_{i=1}^{\alpha} W^i(x_{\mathrm{h}([r],r-1)})$.

Similar to the finitary case, a *k-coloring* of a $\mathbf{W} \in \mathcal{W}^{r,k}$ is a tk-colored r-graphon $\hat{\mathbf{W}} = (W^{(\alpha,\beta)})_{\alpha\in[t],\beta\in[k]}$ with colors from the set $[t]\times[k]$ so that $\sum_{\alpha\in[t],\beta\in[k]} W^{(\alpha,\beta)}(x) = W^\alpha(x)$ for each $x \in [0,1]^{\mathrm{h}([r],r-1)}$ and $\alpha \in [t]$. The k-discoloring $[\hat{\mathbf{W}}, k]$ of $\hat{\mathbf{W}}$ and the term C-colored graphon is defined analogously, and simple r-graphons are treated as 2-colored.

For $q \ge 1$ and $\mathbf{W} \in \mathcal{W}^{r,k}$ the random $[k]\cup\{\iota\}$-colored r-graph $\mathbb{G}(q,\mathbf{W})$ is generated as follows. The vertex set of $\mathbb{G}(q,\mathbf{W})$ is $[q]$, we have to pick uniformly a random point $(X_S)_{S\in\mathrm{h}([q],r-1)} \in [0,1]^{\mathrm{h}([q],r-1)}$, then conditioned on this choice we conduct independent trials to determine the color of each edge $e \in \binom{[q]}{r}$ with the distribution given by $\mathbb{P}_e(\mathbb{G}(q,\mathbf{W})(e) = \alpha) = W^\alpha(X_{\mathrm{h}(e,r-1)})$ corresponding to e. Recall that ι is a special color which we want to avoid in most cases, therefore we will highlight the conditions imposed on the above random variables so that $\mathbb{G}(q,\mathbf{W}) \in \mathcal{G}^{r,k}$.

For $\mathbf{F} \in \mathcal{G}_q^{r,k}$ the $\mathbf{F}$-density of $\mathbf{W}$ is defined as $t(\mathbf{F},\mathbf{W}) = \mathbb{P}(\mathbf{F} = \mathbb{G}(q,\mathbf{W}))$, which can be written following the above description of the random graph as

$$t(\mathbf{F},\mathbf{W}) = \int_{[0,1]^{\mathrm{h}([q],r-1)}} \prod_{e\in\binom{[q]}{r}} W^{\mathbf{F}(e)}(x_{\mathrm{h}(e,r-1)})\mathrm{d}\lambda(x).$$

The above notions were introduced in order to provide a concise representation for the limit space of r-graphs in [6] and [12], in the current work we will not draw on this development explicitly but mention their relevance here. In a nutshell, a sequence of r-graphs converges if the corresponding numerical $\mathbf{F}$-density sequences converge for all r-graphs $\mathbf{F}$. One of the main results of [12] for graphs and [6] in the general case is that for every convergent sequence of r-graphs there exists an r-graphon they converge to in the sense that the $\mathbf{F}$-densities approach the $\mathbf{F}$-density of the limiting r-graphon. This was later reproved by [15] for general r with purely combinatorial methods that are similar to concepts employed in the current paper.

We can associate to each $\mathbf{G} \in \mathcal{G}_n^{r,k}$ an element $\mathbf{W}_\mathbf{G} \in \mathcal{W}^{r,k}$ by subdividing the unit cube $[0,1]^{\mathrm{h}([r],1)}$ into n^r small cubes the natural way and defining the function $W' : [0,1]^{\mathrm{h}([r],1)} \to [k]$ that takes the value $\mathbf{G}(\{i_1,\dots,i_r\})$ on $[\frac{i_1-1}{n},\frac{i_1}{n}]\times\cdots\times[\frac{i_r-1}{n},\frac{i_r}{n}]$ for distinct $i_1,\dots,i_r$, and the value ι on the remaining diagonal cubes. Then set $(\mathbf{W}_\mathbf{G})^\alpha(x_{\mathrm{h}([r],r-1)}) =$

$\mathbb{I}(W'(p_{\mathrm{h}([r],1)}(x_{\mathrm{h}([r],r-1)})) = \alpha)$ for each $\alpha \in [k] \cup \{\iota\}$, where $p_{\mathrm{h}([r],1)}$ is the projection to the suitable coordinates. Note that

$$|t(\mathbf{F}, \mathbf{G}) - t(\mathbf{F}, \mathbf{W}_{\mathbf{G}})| \leq \frac{\binom{q}{2}}{n - \binom{q}{2}} \tag{2.1}$$

for each $\mathbf{F} \in \mathcal{G}_q^{r,k}$, hence the previous representation is compatible in the sense that $\lim_{n\to\infty} t(\mathbf{F}, \mathbf{G}_n) = \lim_{n\to\infty} t(\mathbf{F}, \mathbf{W}_{\mathbf{G}_n})$ for any sequence $\{\mathbf{G}_n\}_{n=1}^{\infty}$ with $|V(\mathbf{G}_n)|$ tending to infinity.

We proceed by providing the necessary formal definitions of the parameter testability in the dense hypergraph model.

Definition 2.1. *An r-graph parameter f is testable if for any $\varepsilon > 0$ there exists a positive integer $q_f(\varepsilon)$ such that for any simple r-graph G with at least $q_f(\varepsilon)$ nodes we have that*

$$\mathbf{P}(|f(G) - f(\mathbb{G}(q_f(\varepsilon), G)| > \varepsilon) < \varepsilon.$$

The smallest function q_f satisfying the previous inequality is called the sample complexity of f. The testability of parameters of k-colored r-graphs is defined analogously.

An a priori weaker characteristic than the one above, nondeterministic testability, is the second cornerstone of the current work, and was introduced in [14].

Definition 2.2. *An r-graph parameter f is non-deterministically testable if there exist an integer k and a testable $2k$-colored directed r-graph parameter g called witness such that for any simple graph G the value $f(G) = \max_{\mathbf{G}} g(\mathbf{G})$ where the maximum goes over the set of k-colorings of G (regarded as an element of $\mathcal{G}^{r,2}$).*

Originally in [14], the witness parameter was a function of k-colored graphs, and the maximum was taken over the set of (k, m)-colorings of the original graph in order to determine the parameter value, meaning that the present edges are colored by elements of $[m]$, absent ones by the remaining colors from $[k] \setminus [m]$. Our modification is equivalent to that setting and is motivated by notational purposes.

In the current paper we only deal with undirected structures, but similar results can be obtained when the witness parameter is defined on the space of directed r-graphs. In this case, in order to obtain G from $\mathbf{G}$ as above after the discoloring we additionally have to undirect the edges and neglect multiplicities created by the former operation.

The maximization expression in Definition 2.2 is somewhat arbitrary and could be replaced for example by minimization, this would however not affect the testability characteristic of the parameter. Our main result extends the equivalence of the two testability notions for arbitrary r, this was first proved by Lovász and Vesztergombi [14] for $r = 2$.

Theorem 2.3. *Every non-deterministically testable r-graph parameter f is testable.*

Our proof follows the proof skeleton of [10], but requires a more sophisticated approach.The reason for this is that the analogous norm for hypergraphs to the cut-norm that comes with a counting lemma has some shortcomings, for instance the sample is in some cases far away from the original in the natural distance notion induced by the norm. Therefore the corresponding regularity lemma cannot be applied directly as in [10].

The definitions of the relevant norms is given next.

Definition 2.4. *Let $r \geq 1$ and A be a real r-array of size n. Then the cut norm of A is*

$$\|A\|_{\square,r} = \frac{1}{n^r} \max_{\substack{S_i \subset [n]^{r-1} \setminus \mathrm{diag}([n]^{r-1}) \\ i \in [r]}} |A(r; S_1, \ldots, S_r)|,$$

where $A(r; S_1, \ldots, S_r) = \sum_{i_1,\ldots,i_r=1}^{n} A(i_1, \ldots, i_r) \prod_{j=1}^{r} \mathbb{I}_{S_j}(i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_r)$, and the maximum goes over sets S_i that are invariant under coordinate permutations.

If $\mathcal{P} = (P_i)_{i=1}^{t}$ is a partition of $[n]^{r-1} \setminus \mathrm{diag}([n]^{r-1})$ with symmetric classes, then the cut-$\mathcal{P}$-norm of A is

$$\|A\|_{\square,r,\mathcal{P}} = \frac{1}{n^r} \max_{S_i \subset [n]^{r-1}, i \in [r]} \sum_{j_1,\ldots,j_r=1}^{t} |A(r; S_1 \cap P_{j_1}, \ldots, S_r \cap P_{j_r})|.$$

The cut norm of an r-kernel W is

$$\|W\|_{\square,r} = \sup_{\substack{S_i \subset [0,1]^{\mathrm{h}([r],r-1)} \\ i \in [r]}} \left| \int_{\cap_{i \in [r]} p_{[r]\setminus\{i\}}^{-1}(S_i)} W(x_{\mathrm{h}([r],r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r],r-1)}) \right|,$$

where the supremum is taken over sets S_i that satisfy the usual symmetries, and p_e is the natural projection from $[0,1]^{\mathrm{h}([r],r-1)}$ onto $[0,1]^{\mathrm{h}(e)}$. Furthermore, for a symmetric partition $\mathcal{P} = (P_i)_{i=1}^{t}$ of $[0,1]^{\mathrm{h}([r-1])}$ the cut-$\mathcal{P}$-norm of an r-kernel is defined by

$$\|W\|_{\square,r,\mathcal{P}} = \sup_{\substack{S_i \subset [0,1]^{\mathrm{h}([r-1])} \\ i \in [r]}} \sum_{j_1,\ldots,j_r=1}^{t} \left| \int_{\cap_{i \in [r]} p_{[r]\setminus\{i\}}^{-1}(S_i \cap P_{j_i})} W(x_{\mathrm{h}([r],r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r],r-1)}) \right|,$$

where the supremum is taken over sets S_i that satisfy the usual symmetries.

We remark that it is also true that

$$\|W\|_{\square,r} = \sup_{f_1,\ldots,f_r \in [0,1]^{\mathrm{h}([r-1])}} \left| \int_{[0,1]^{\mathrm{h}([r],r-1)}} \prod_{i=1}^{r} f_i(x_{\mathrm{h}([r]\setminus\{i\})}) W(x_{\mathrm{h}([r],r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r],r-1)}) \right|,$$

where the supremum goes over functions f_i that satisfy the usual symmetries, and similarly for any symmetric partition $\mathcal{P} = (P_i)_{i=1}^{t}$ of $[0,1]^{\mathrm{h}([r-1])}$ we have with the same conditions for the f_i's as above that

$$\|W\|_{\square,r} = \sup_{f_1,\ldots,f_r \in [0,1]^{\mathrm{h}([r-1])}} \sum_{j_1,\ldots,j_r=1}^{t} \left| \int_{[0,1]^{\mathrm{h}([r],r-1)}} \prod_{i=1}^{r} f_i(x_{\mathrm{h}([r]\setminus\{i\})}) \mathbb{I}_{P_{j_i}}(x_{\mathrm{h}([r]\setminus\{i\})}) W(x_{\mathrm{h}([r],r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r],r-1)}) \right|.$$

In several previous works, see e.g. [1], the cut norm for r-arrays denotes a term that is significantly different from the one in Definition 2.4 and is not suitable for our present purposes. The above norms give rise to a distance between r-graphons, and analogously for r-graphs.

Definition 2.5. *For two k-colored r-graphons $\mathbf{U} = (U^\alpha)_{\alpha\in[k]}$ and $\mathbf{W} = (W^\alpha)_{\alpha\in[k]}$ their cut distance is defined as*

$$d_{\square,r}(\mathbf{U},\mathbf{W}) = \sum_{\alpha=1}^{k} \|U^\alpha - W^\alpha\|_{\square,r},$$

and their cut-$\mathcal{P}$-distance as

$$d_{\square,r,\mathcal{P}}(\mathbf{U},\mathbf{W}) = \sum_{\alpha=1}^{k} \|U^\alpha - W^\alpha\|_{\square,r,\mathcal{P}}.$$

For two k-colored r-graphs $\mathbf{G} = (G^\alpha)_{\alpha\in[k]}$ and $\mathbf{H} = (H^\alpha)_{\alpha\in[k]}$ their corresponding distances are defined as

$$d_{\square,r}(\mathbf{G},\mathbf{H}) = d_{\square,r}(\mathbf{W_G},\mathbf{W_H}),$$

and

$$d_{\square,r,\mathcal{P}}(\mathbf{G},\mathbf{H}) = d_{\square,r,\mathcal{P}}(\mathbf{W_G},\mathbf{W_H}).$$

Distances between an r-graph and an r-graphon, as well as for r-kernels, is analogously defined.

Note that the norms introduced above are in general smaller or equal than the 1-norm of integrable functions, also $d_{\square,r}(\mathbf{U},\mathbf{W}) \le d_{\square,r,\mathcal{P}}(\mathbf{U},\mathbf{W})$ hods for every pair. Their relevance will be clearer in the context of the next counting lemma, we include the standard proof only for completeness' sake.

Lemma 2.6. *Let $\mathbf{U}$ and $\mathbf{W}$ be two k-colored r-graphons with $\|U\|_\infty, \|W\|_\infty \le 1$. Then for every $\mathbf{F} \in \mathcal{G}_q^{r,k}$ it holds that*

$$|t(\mathbf{F},\mathbf{W}) - t(\mathbf{F},\mathbf{U})| \le \binom{q}{r} d_{\square,r}(\mathbf{U},\mathbf{W}).$$

Proof. Fix q and $\mathbf{F} \in \mathcal{G}_q^{r,k}$. Then

$$\begin{aligned}|t(\mathbf{F},\mathbf{W}) - t(\mathbf{F},\mathbf{U})| &= \Big| \int_{[0,1]^{\mathrm{h}([q],r-1)}} \prod_{e\in\binom{[q]}{r}} W^{\mathbf{F}(e)}(x_{\mathrm{h}(e,r-1)}) - \prod_{e\in\binom{[q]}{r}} U^{\mathbf{F}(e)}(x_{\mathrm{h}(e,r-1)}) \mathrm{d}\lambda(x)\Big| \\ &\le \sum_{e\in\binom{[q]}{r}} \Big| \int_{[0,1]^{\mathrm{h}([q],r-1)}} [W^{\mathbf{F}(e)}(x_{\mathrm{h}(e,r-1)}) - U^{\mathbf{F}(e)}(x_{\mathrm{h}(e,r-1)})]\end{aligned}$$

$$\prod_{f\in\binom{[q]}{r},f\prec e} W^{\mathbf{F}(f)}(x_{\mathrm{h}(f,r-1)}) \prod_{g\in\binom{[q]}{r},e\prec g} U^{\mathbf{F}(g)}(x_{\mathrm{h}(g,r-1)})\mathrm{d}\lambda(x)|$$

$$\le \sum_{e\in\binom{[q]}{r}} \|W^{\mathbf{F}(e)} - U^{\mathbf{F}(e)}\|_{\square,r} \le \binom{q}{r} d_{\square,r}(\mathbf{U},\mathbf{W}),$$

where $\prec$ is an arbitrary total ordering of the elements of $\binom{q}{r}$. □

Let d_{tw} denote the total variation distance between probability measures on $\mathcal{G}_n^{r,[k]*}$, where $[k]* = [k] \cup \{\iota\}$ for $k \ge 1$ (without highlighting the specific parameters in the notion d_{tw}), that is $d_{\mathrm{tw}}(\mu,\nu) = \max_{\mathcal{F}\subset\mathcal{G}_n^{r,[k]*}} |\mu(\mathcal{F}) - \nu(\mathcal{F})|$, and let the measure $\mu(q,\mathbf{G})$, respectively $\mu(q,\mathbf{W})$, denote the probability measure of the random r-graph $\mathbb{G}(q,\mathbf{G})$, respectively $\mathbb{G}(q,\mathbf{W})$, taking values in $\mathcal{G}_q^{r,[k]*}$. It is a standard observation then that

$$d_{\mathrm{tw}}(\mu(q,\mathbf{W}),\mu(q,\mathbf{U})) = \frac{1}{2}\sum_{\mathbf{F}\in\mathcal{G}_q^{r,[k]*}} |t(\mathbf{F},\mathbf{W}) - t(\mathbf{F},\mathbf{U})|, \tag{2.2}$$

and that $\mathbb{G}(q,\mathbf{W})$ and $\mathbb{G}(q,\mathbf{U})$ can be coupled in form of the random r-graphs $\mathbf{G}_1$ and $\mathbf{G}_2$, such that

$$d_{\mathrm{tw}}(\mu(q,\mathbf{W}),\mu(q,\mathbf{U}) = \frac{1}{2}\mathbf{P}(\mathbf{G}_1 \ne \mathbf{G}_2), \tag{2.3}$$

and further, for any coupling $\mathbf{G}_1'$ and $\mathbf{G}_2'$ it hods that $d_{\mathrm{tw}}(\mu(q,\mathbf{W}),\mu(q,\mathbf{U}) \le \frac{1}{2}P(\mathbf{G}_1' \ne \mathbf{G}_2')$.

For $\mathbf{G} \in \mathcal{G}_n^{r,k}$ note that

$$d_{\mathrm{tw}}(\mu(q,\mathbf{G}),\mu(q,\mathbf{W}_\mathbf{G})) \le q^2/n, \tag{2.4}$$

where the right hand side is a simple upper bound on the probability that if we uniformly choose q elements of an n-element set, then we get at least two identical objects. The inequality (2.4) follows from the fact that conditioned on the event that the independent and uniform $X_{\{i\}}$'s for $i \in [q]$ fall in different intervals $[\frac{j-1}{n},\frac{j}{n}]$ for $j \in [n]$ the distribution of $\mathbb{G}(q,\mathbf{W}_\mathbf{G})$ is the same as the distribution of $\mathbb{G}(q,\mathbf{G})$.

The next corollary is a direct consequence of Lemma 2.6.

Corollary 2.7. *If* $\mathbf{U}$ *and* $\mathbf{W}$ *are two* k*-colored* r*-graphons, then*

$$d_{\mathrm{tw}}(\mu(q,\mathbf{W}),\mu(q,\mathbf{U}) \le \frac{k^{q^r}q^r}{2r!} d_{\square,r}(\mathbf{U},\mathbf{W}),$$

and there exists a coupling in form of $\mathbf{G}_1$ *and* $\mathbf{G}_2$ *of the random* r*-graphs* $\mathbb{G}(q,\mathbf{W})$ *and* $\mathbb{G}(q,\mathbf{U})$*, such that*

$$P(\mathbf{G}_1 \ne \mathbf{G}_2) \le \frac{k^{q^r}q^r}{2r!} d_{\square,r}(\mathbf{U},\mathbf{W}).$$

A generalization of the notion of a step function in the case of graphs to the situation where we deal with r-graphs is given next. For a partition $\mathcal{P}$ the number of its classes is denoted by $t_{\mathcal{P}}$.

Definition 2.8. *We call an k-colored r-graphon $\mathbf{W}$ with $r \geq l$ an (r,l)-step function if there exist positive integers $t_l, t_{l+1}, \ldots, t_r = k$, symmetric partitions $\mathcal{P} = (P_1, \ldots, P_{t_l})$ of $[0,1]^{\mathrm{h}([l])}$, and real arrays $A_s^\alpha : [t_{s-1}]^{\mathrm{h}([s],s-1)} \to [0,1]$ with $\alpha \in [t_s]$ for $l \leq s \leq r$ such that $\sum_{\alpha \in [t_s]} A_s^\alpha(i_{\mathrm{h}([s],s-1)}) = 1$ for any choice of $i_{\mathrm{h}([s],s-1)}$ and for $s \leq r$ so that W^α for $\alpha \in [k]$ is of the following form for each $\alpha \in [k]$.*

$$W^\alpha(x_{\mathrm{h}([r])}) = \sum_{\substack{i_S=1 \\ S\subset[r], l\leq|S|}}^{t_{|S|}} A_r^\alpha(i_{\mathrm{h}([r],r-1)})$$
$$\prod_{S\in\binom{[r]}{l}} \mathbb{I}_{P_{i_S}}(x_{\mathrm{h}(S)}) \prod_{\substack{S\subset[r] \\ l+1\leq|S|<r}} \mathbb{I}\Big(\sum_{j=1}^{i_S-1} A_{|S|}^j(i_{\mathrm{h}(S,|S|-1)}) \leq x_S \leq \sum_{j=1}^{i_S} A_{|S|}^j(i_{\mathrm{h}(S,|S|-1)})\Big).$$

We refer to the partition $\mathcal{P}$ as the steps of W.

The most simple example is the $(r, r-1)$ step function that can be written as

$$W^\alpha(x_{\mathrm{h}([r])}) = \sum_{i_1,\ldots,i_r=1}^{t_{r-1}} A_r^\alpha(i_1,\ldots,i_r) \prod_{j=1}^{r} \mathbb{I}_{P_{i_j}}(x_{\mathrm{h}([r]\setminus\{j\})}).$$

3 Testability of the r-cut norm

We define a parameter of r-uniform hypergraphs that is a generalization of the ground state energies of [5] in the case of graphs. This notion encompasses several important quantities, therefore its testability is central to many applications.

Definition 3.1. *For a set $H \subset \binom{[n]}{r}$, a real r-array J of size q, and a symmetric partition $\mathcal{P} = (P^1, \ldots, P^q)$ of $\binom{[n]}{r-1}$ we define the energy*

$$\mathcal{E}_{\mathcal{P},r-1}(H,J) = \frac{1}{n^r} \sum_{i_1,\ldots,i_r=1}^{q} J(i_1,\ldots,i_r) e_H(r; P_{i_1},\ldots,P_{i_r}),$$

where $e_H(r; S_1, \ldots, S_r) = |\{(u_1,\ldots,u_r) \in [n]^r | A_H(u_1,\ldots,u_r) = 1$ and $A_{S_j}(u_1,\ldots,u_{j-1},u_{j+1},\ldots,u_r) = 1$ for all $j = 1,\ldots,r\}|$.

Let $\mathbf{H} = (H^\alpha)_{\alpha\in[k]}$ be a k-colored r-uniform hypergraph on the vertex set $[n]$ and J^α a be real $q \times \cdots \times q$ r-array with $\|J\|_\infty \leq 1$ for each $\alpha \in [k]$. Then the energy for a partition $\mathcal{P}$ as above is

$$\mathcal{E}_{\mathcal{P},r-1}(\mathbf{H},J) = \sum_{\alpha\in[k]} \mathcal{E}_{\mathcal{P},r-1}(H^\alpha, J^\alpha).$$

The maximum of the energy over all partitions $\mathcal{P}$ of $\binom{[n]}{r-1}$ is called the ground state energy (GSE) of H with respect to J, and is denoted by

$$\mathcal{E}_{r-1}(\mathbf{H}, J) = \max_{\mathcal{P}} \mathcal{E}_{\mathcal{P},r-1}(\mathbf{H}, J).$$

The GSE can also be defined for r-graphons.

Definition 3.2. *For an r-graphon W, a real r-array J of size q, and a symmetric partition $\mathcal{P} = (P^1, \dots, P^q)$ of $[0,1]^{\mathrm{h}([r-1])}$ we define the energy*

$$\mathcal{E}_{\mathcal{P},r-1}(W, J) = \sum_{i_1,\dots,i_r \in [q]} J(i_1, \dots, i_r) \int_{\bigcap_{j\in[r]} p^{-1}_{[r]\setminus\{j\}}(S_{i_j})} W(x_{\mathrm{h}([r],r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r],r-1)}).$$

Let $\mathbf{W} = (W^\alpha)_{\alpha\in[k]}$ be a k-colored r-graphon and J^α a be real $q \times \cdots \times q$ r-array with $\|J\|_\infty \le 1$ for each $\alpha \in [k]$. Then the energy for a partition $\mathcal{P}$ as above is

$$\mathcal{E}_{\mathcal{P},r-1}(\mathbf{W}, J) = \sum_{\alpha\in[k]} \mathcal{E}_{\mathcal{P},r-1}(W^\alpha, J^\alpha).$$

and the GSE of $\mathbf{W}$ with respect to J, and is denoted by

$$\mathcal{E}_{r-1}(\mathbf{W}, J) = \sup_{\mathcal{P}} \mathcal{E}_{\mathcal{P},r-1}(\mathbf{W}, J),$$

where the supremum runs over all symmetric partitions $\mathcal{P} = (P^1, \dots, P^q)$ of $[0,1]^{\mathrm{h}([r-1])}$.

Definitions of the above energies are analogous in the directed, and the weighted case, and also for r-kernels. The next lemma tells us about the distribution of the GSE when taking a random sample $\mathbb{G}(n, \mathbf{H})$ of an $\mathbf{H} \in \mathcal{G}^{r,k}$.

Lemma 3.3. *The expression $\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J)$ is highly concentrated around its mean, that is for every $\varepsilon > 0$ it holds that*

$$\mathbf{P}(|\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J) - \mathbb{E}\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J)| \ge \varepsilon \|J\|_\infty) \le 2\exp(-\frac{\varepsilon^2 n}{8r^2}).$$

Proof. We can assume that $\|J\|_\infty \le 1$. The random r-graph $\mathbb{G}(n, \mathbf{H})$ is generated by picking random nodes from $V(\mathbf{H})$ without repetition, let X_i denote the ith random element of $V(\mathbf{H})$ that has been selected. Define the martingale $Y_i = \mathbb{E}[\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J)|X_1, \dots, X_i]$ for $0 \le i \le n$. It has the property that $Y_0 = \mathbb{E}[\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J)]$ and $Y_n = \mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}), J)$, whereas the jumps $|Y_i - Y_{i-1}|$ are bounded above by $\frac{2r}{n}$ for each $i \in [n]$. The last observation is the consequence of the fact that for any partition $\mathcal{P}$ of $\binom{[n]}{r-1}$ only at most rn^{r-1} terms in the sum constituting $\mathcal{E}_{\mathcal{P},r-1}(\mathbf{H}, J)$ are affected by changing the placing of X_{i+1} in the classes of $\mathcal{P}$. Applying the Azuma-Hoeffding inequality to the martingale verifies the statement of the lemma. □

The same concentration result as above applies also to $\mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{W}), J)$.

We will show that these hypergraph parameters are testable via the ultralimit method and the machinery developed by Elek and Szegedy [6]. From the notational perspective and theoretical background this section slightly stands out from the rest of the paper. First we give a brief summary of the notions that were used in [6] in order to produce a representation for the limit space of simple r-graphs. This representation led to a new analytical proof method for several results for simple r-graphs such as the Regularity Lemma, the Removal Lemma, or the testability assertion about hereditary r-graph properties. Subsequently, technical results proved in [6] which are relevant here are mentioned, for more details and complete proofs we refer to the source paper [6].

Recall that a sequence of r-graphs $(G_n)_{n\geq 1}$ is convergent if for every simple F the numerical sequences $t(F, G_n)$ converge when n tends to infinity.

We start by introducing the basic notations for ultraproduct measure spaces. Let us fix a non-principal ultrafilter ω on $\mathbb{N}$, and let $X_1, X_2, \dots$ be a sequence of finite sets of increasing size. We define the infinite product set $\hat{X} = \prod_{i=1}^{\infty} X_i$ and the equivalence relation $\sim$ between elements of $\hat{X}$, so that $p \sim q$ if and only if $\{i| \quad p_i = q_i\} \in \omega$. Set $\mathbf{X} = \hat{X}/\sim$, this set is called the ultraproduct of the X_i's, and it will serve as the base set of the ultraproduct probability space. Further, let $\mathcal{P}$ denote the algebra of subsets of $\mathbf{X}$ of the form $A = [\{A_i\}_{i=1}^{\infty}]$, where $A_i \subset X_i$ for each i, and $[.]$ denotes the equivalence class under $\sim$ (for convenience, $p = [\{p_i\}_{i=1}^{\infty}] \in [\{A_i\}_{i=1}^{\infty}]$ exactly in the case when $\{i|p_i \in A_i\} \in \omega$).

Define a measure on the sets belonging to $\mathcal{P}$ through the ultralimit of the counting measure on the sets X_i, that is, $\mu(A) = \lim_{\omega} \frac{|A_i|}{|X_i|}$, where the ultralimit of a bounded real numerical sequence $\{x_i\}_{i=1}^{\infty}$ is denoted by $x = \lim_{\omega} x_i$, and is defined by the property that for every $\varepsilon > 0$ we have $\{i| \quad |x - x_i| < \varepsilon\} \in \omega$. One can see that the limit exists for every bounded sequence and is unique, therefore well-defined, this is a consequence of basic properties of a non-principal ultrafilter. The set of $\mathcal{N} \subset 2^{\mathbf{X}}$ of μ-null sets is the family of sets N for those there exists an infinite sequence of supersets $\{A^i\}_{i=1}^{\infty} \subset \mathcal{P}$ such that $\mu(A^i) \leq 1/i$. Finally define the σ-algebra $\mathcal{B}$ on $\mathbf{X}$ by the σ-algebra generated by $\mathcal{P}$ and $\mathcal{N}$, and set the measure $\mu(B) = \mu(A)$ for each $B \in \mathcal{B}$, where $A \triangle B \in \mathcal{N}$ and $A \in \mathcal{P}$. Again, everything is well-defined, see [6], so we arrive at the ultraproduct measure space $(\mathbf{X}, \mathcal{B}, \mu)$.

Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be two increasing sequences of finite sets with ultraproducts $\mathbf{X}$ and $\mathbf{Y}$ respectively, then it is true that the ultraproduct of the product sequence $X_1 \times Y_1, X_2 \times Y_2, \dots$ is the product $\mathbf{X} \times \mathbf{Y}$, but the σ-algebra $\mathcal{B}_{\mathbf{X}\times\mathbf{Y}}$ of the measure space can be strictly larger than the σ-algebra generated by $\mathcal{B}_{\mathbf{X}} \times \mathcal{B}_{\mathbf{Y}}$, and this is a crucial point when the aim is to construct a separable representation of the ultraproduct measure space of product sets.

Let r be some positive integer, and again $X_1, X_2, \dots$ a sequence of finite sets as above. For any $e \subset [r]$ we define the ultraproduct measure spaces $(\mathbf{X}^e, \mathcal{B}_{\mathbf{X}^e}, \mu_e)$, also let P_e denote the natural projection from $\mathbf{X}^{[r]}$ to $\mathbf{X}^e$. Furthermore let $\sigma(e)$ denote the sub-σ-algebra of $\mathcal{B}_{\mathbf{X}^{[r]}}$ given by $P_e^{-1}(\mathcal{B}_{\mathbf{X}^e})$, and $\sigma(e)^*$ be the sub-σ-algebra $\langle P_f^{-1}(\mathcal{B}_{\mathbf{X}^f}) | f \subset e, |f| < |e| \rangle$. Note that in general $\sigma(e)$ is strictly larger than $\sigma(e)^*$. We denote the measure $\mu_{\mathbf{X}^e}$ simply by μ_e and the σ-algebra $\mathcal{B}_{\mathbf{X}^e}$ by $\mathcal{B}_e$.

Definition 3.4. *Let r be a positive integer. We call a measure preserving map $\phi : \mathbf{X}^{[r]} \to [0,1]^{\mathrm{h}([r])}$ a separable realization if*

1. *for any permutation $\pi \in S_{[r]}$ of the coordinates we have for all $x \in \mathbf{X}^{[r]}$ that $\Pi(\phi(x)) = \phi(\phi(x))$, where Π is the permutation of the power set of $[r]$ induced by π, and*

2. *for any $e \subset \mathrm{h}([r])$ and any measurable $A \subset [0,1]$ we have that $\phi_e^{-1}(A) \in \sigma(e)$ and $\phi_e^{-1}(A)$ is independent of $\sigma(e)^*$.*

We are interested in the limiting behavior of sequences of k-partitions (or edge-k-colored r-graphs on the vertex sets $X_1, X_2, \dots$) of the sequence $X_1^r, X_2^r, \dots$, where convergence is defined in the following general way.

Let $G_i = (G_i^1, \dots, G_i^k)$ be a symmetric partition of X_i^r for each $i \in \mathbb{N}$, then $(G_i)_{i=1}^\infty$ converges if for every k-colored r-graph F the numerical sequences $t(F, G_i)$ converge, as in Section 2. The ultralimit method enables us to handle the cases where the convergence does not hold without going to subsequences, we describe the method next. Let us denote the size of F by m and let $F(e)$ be the color of $e \in \binom{[m]}{r}$, then $t(F, G_i)$ can be written as the measure of a subset of X_i^m. We show this by explicitly presenting the set denoted by $T(F, G_i)$, so let

$$T(F, G_i) = \bigcap_{e \in \binom{[m]}{r}} P_e^{-1}(P_{s_e}(G_i^{F(e)})), \tag{3.1}$$

where P_e is the natural projection from $X_i^{[m]}$ to X_i^e, and P_{s_e} is a bijection going from $X_i^{[r]}$ to X_i^e induced by an arbitrary but fixed bijection s_e between e and $[r]$. We define the induced subgraph density of the ultraproduct of k-colored r-graphs formally following (3.1), if $\mathbf{G} = (\mathbf{G}^1, \dots, \mathbf{G}^k)$ is a $\mathcal{B}_{[r]}$-measurable k-partition of $\mathbf{X}^{[r]}$ and F is as above then let

$$T(F, \mathbf{G}) = \bigcap_{e \in \binom{[m]}{r}} P_e^{-1}(P_{s_e}(\mathbf{G}^{F(e)})). \tag{3.2}$$

It is easy to see that $\lambda(T(F, G_i)) = t(F, G_i)$. Forming the ultraproduct of a series of sets commutes with finite intersection, therefore $\lim_\omega T(F, G_i) = T(F, \lim_\omega G_i)$ and $\lim_\omega t(F, G_i) = t(F, \lim_\omega G_i)$. Observe that all of the above notation makes perfect sense and the identities hold true for directed colored r-graphs, that is, when the adjacency arrays of the G^α's are not necessarily symmetric.

We call a measurable subset of $[0,1]^{\mathrm{h}([r])}$ an r-set graphon satisfying the usual symmetries in the coordinates induced by S_r permutations, we can turn it into a proper r-graphon in the sense of Section 2 by generating the marginal with respect to the coordinate corresponding to $[r]$. Analogously a k-colored r-set graphon is a measurable partition of $[0,1]^{\mathrm{h}([r])}$ into k classes invariant under coordinate permutations induced by permuting $[r]$. These objects can serve as representations of the ultralimits of r-graph sequences in the sense that the numerical sequences of subgraph densities converge to densities defined for r-set graphons in accordance with the notation in Section 2, we will provide the definition next.

Definition 3.5. *Let* $\mathbf{F}$ *be a* k*-colored* r*-graph on* m *vertices, and* $\mathbf{W} = (W^1, \dots, W^k)$ *be a* k*-colored* r*-set graphon. Then* $T(\mathbf{F}, \mathbf{W}) \subset [0,1]^{\mathrm{h}([m],r)}$ *denotes the set of the symmetric maps* $g : \mathrm{h}([m], r) \to [0,1]$ *that satisfy that for each* $e \in \binom{[m]}{r}$ *it holds that* $(g(f))_{f \in \mathrm{h}(e)} \in W^{\mathbf{F}(e)}$*. For the Lebesgue measure of* $T(\mathbf{F}, \mathbf{W})$ *we write* $t(\mathbf{F}, \mathbf{W})$*, this expression is referred to as the density of* $\mathbf{F}$ *in* $\mathbf{W}$*.*

The reader may easily verify that the above definition of density agrees with the content of Section 2. One of the main technical results of [6] is the following.

Theorem 3.6. *[6] Let* r *be an arbitrary positive integer and let* $\mathcal{A}$ *be a separable sub-*σ*-algebra of* $\mathcal{B}_{[r]}$*. Then there exists a separable realization* $\phi : \mathbf{X}^{[r]} \to [0,1]^{\mathrm{h}([r])}$ *such that for every* $A \in \mathcal{A}$ *there exists a measurable* $B \subset [0,1]^{\mathrm{h}([r])}$ *such that* $\mu_{[r]}(A \triangle \phi^{-1}(B)) = 0$*.*

A lifting of a separable realization $\phi : \mathbf{X}^{[r]} \to [0,1]^{\mathrm{h}([r])}$ of degree n for $n \geq r$ is a measure preserving map $\psi : \mathbf{X}^{[n]} \to [0,1]^{\mathrm{h}([n],r)}$ that satisfies $p_{\mathrm{h}([r])} \circ \psi = \phi \circ P_{[r]}$, and it is equivariant under coordinate permutations in S_n, where $p_{\mathrm{h}([r])}$ and $P_{[r]}$ are the natural projections from $[0,1]^{\mathrm{h}([n],r)}$ to $[0,1]^{\mathrm{h}([r])}$, and from $\mathbf{X}^{[n]}$ to $\mathbf{X}^{[r]}$ respectively. The next lemma is central to relate the sub-r-graph densities of ultraproducts to the corresponding densities in r-set graphons.

Lemma 3.7. *[6] For every separable realization* ϕ *and integer* $n \geq r$ *there exists a degree* n *lifting* ψ*.*

The next statement is the colored version of the homomorphism correspondence in [6] (Lemma 3.3. in that paper).

Lemma 3.8. *Let* ϕ *be a separable realization and* $\mathbf{W} = (W^1, \dots, W^k)$ *be a* k*-colored* r*-graphon, and let* $\mathbf{H} = (\mathbf{H}^1, \dots, \mathbf{H}^k)$ *be a* k*-colored ultraproduct with* $\mu_{[r]}(\mathbf{H}^\alpha \triangle \phi^{-1}(W^\alpha)) = 0$ *for each* $\alpha \in [k]$*. Let* ψ *be a degree* m *lifting of* ϕ *and* $\mathbf{F}$ *be a* k*-colored* r*-graph on* m *vertices. Then* $\mu_{[m]}(\psi^{-1}(T(\mathbf{F}, \mathbf{W})) \triangle T(\mathbf{F}, \mathbf{H})) = 0$*, and consequently* $t(\mathbf{F}, \mathbf{W}) = t(\mathbf{F}, \mathbf{H})$ *for each* $\mathbf{F}$*.*

Proof. By definition we have that

$$T(F, \mathbf{H}) = \bigcap_{e \in \binom{[m]}{r}} P_e^{-1}(P_{s_e}(\mathbf{H}^{F(e)}))$$

and

$$T(F, W) = \bigcap_{e \in \binom{[m]}{r}} p_{\mathrm{h}([r])}^{-1}(p_{s_e}(W^{F(e)})).$$

Due to the fact that ψ commutes with coordinate permutations from S_n and the conditions we imposed on the symmetric difference of $\mathbf{H}^\alpha$ and $\phi^{-1}(W^\alpha)$ the statement follows. □

We turn to describe the relationship of two r-set graphons whose $\mathbf{F}$-densities coincide for each $\mathbf{F}$. For this purpose we have to introduce two types of transformations and clarify their connection. Let us define the σ-algebras $\mathcal{A}_S$, $\mathcal{A}_S^*$, and $\mathcal{B}_S \subset \mathcal{L}_{[0,1]^{\mathrm{h}([r])}}$ for each $S \subset [r]$, the σ-algebra $\mathcal{B}_S = p_S^{-1}(\mathcal{L}_{[0,1]^{\mathrm{h}(S)}})$, $\mathcal{A}_S$ is $\langle \mathcal{B}_T | T \subset S \rangle$, and $\mathcal{A}_S^*$ is $\langle \mathcal{B}_T | T \subset S, T \neq S \rangle$, where $\mathcal{L}_{[0,1]^t}$ denotes the Lebesgue measurable subsets of the unit cube with the dimension given by the index.

Definition 3.9. *We say that the measurable map $\phi : [0,1]^{\mathrm{h}([r])} \to [0,1]^{\mathrm{h}([r])}$ is structure preserving if it is measure preserving, for any $S \subset [r]$ we have $\phi^{-1}(\mathcal{A}_S) \subset \mathcal{A}_S$, for any measurable $I \subset [0,1]$ we have $\phi^{-1}(p_S^{-1}(I))$ is independent of $\mathcal{A}_S^*$, and for any $\pi \in S_r$ we have $\Pi \circ \phi = \phi \circ \Pi$, where Π is the coordinate permutation action induced by π.*

Let $\mathcal{L}^{\mathrm{h}([r])}$ denote the measure algebra of $([0,1]^{\mathrm{h}([r])}, \mathcal{L}_{[0,1]^{\mathrm{h}([r])}}, \lambda)$.

Definition 3.10. *We call an injective homomorphism $\Phi : \mathcal{L}^{\mathrm{h}([r])} \to \mathcal{L}^{\mathrm{h}([r])}$ a structure preserving embedding if it is measure preserving, for any $S \subset [r]$ we have $\Phi(\mathcal{B}_S) \subset \mathcal{A}_S$, $\Phi(\mathcal{B}_S)$ is independent from $\mathcal{A}_S^*$, and for any $\pi \in S_r$ we have $\Pi \circ \Phi = \Phi \circ \Pi$.*

Another result from [6] sheds light on the build-up of structure preserving embeddings.

Lemma 3.11. *[6] Suppose that $\Phi : \mathcal{L}^{\mathrm{h}([r])} \to \mathcal{L}^{\mathrm{h}([r])}$ is a structure preserving embedding of a measure algebra into itself. Then there exists a structure preserving map $\phi : [0,1]^{\mathrm{h}([r])} \to [0,1]^{\mathrm{h}([r])}$ that represents Φ in the sense that for each $[U] \in L^{\mathrm{h}([r])}$ it holds that $\Phi([U]) = [\phi^{-1}(U)]$, where U is a representative of $[U]$.*

A random coordinate system τ is the ultraproduct function on $\mathbf{X}^{[r]}$ of the random symmetric functions $\tau_n : [n]^r \to [0,1]^{\mathrm{h}([n],r)}$ that are for each n given by a uniform random point Z_n in $[0,1]^{\mathrm{h}([n],r)}$ so that $(\tau_n(i_1,\ldots,i_r))_e = (Z_n)_{p_e(i_1,\ldots,i_r)}$. An important property of the random mapping τ_n is that for any r-set graphon and positive integer n it holds that $(\tau_n)^{-1}(U) = \mathbb{G}(n,U)$, when the random sample Z_n used to generate the two objects is the same.

Lemma 3.12. *[6] Let U be an r-set graphon, and let $\mathbf{H} = [\{\mathbb{G}(n,U)\}_{n=1}^{\infty}]$. Then the random coordinate system $\tau = [\{\tau_n\}_{n=1}^{\infty}]$ is a separable realization such that with probability one we have $\mu_{[r]}(\mathbf{H} \triangle \tau^{-1}(U)) = 0$.*

A direct consequence is the statement for k-colored r-set graphons.

Corollary 3.13. *Let $\mathbf{U} = (U^1,\ldots,U^k)$ be a k-colored r-set graphon, and let $\mathbf{H} = (\mathbf{H}^1,\ldots,\mathbf{H}^k)$ be a k-colored ultraproduct in $\mathbf{X}^{[r]}$, where $\mathbf{H}^\alpha = [\{\mathbb{G}(n,U^\alpha)\}_{n=1}^{\infty}]$ for each $\alpha \in [k]$. Then a random separable realization τ is such that with probability one we have $\mu_{[r]}(\mathbf{H}^\alpha \triangle \tau^{-1}(U^\alpha)) = 0$ for each $\alpha \in [k]$.*

The following result is a generalization of the uniqueness assertion of [6], and states that subgraph densities determine an r-set graphon up to structure preserving transformations.

Theorem 3.14. *Let $\mathbf{U} = (U^1,\ldots,U^k)$ and $\mathbf{V} = (V^1,\ldots,V^k)$ be two k-colored r-set graphons such that for each k-colored r-graph $\mathbf{F}$ it holds that $t(\mathbf{F},\mathbf{U}) = t(\mathbf{F},\mathbf{V})$. Then there exist two structure preserving maps ν_1 and ν_2 from $[0,1]^{\mathrm{h}([r])}$ to $[0,1]^{\mathrm{h}([r])}$ such that $\mu_{[r]}(\nu_1^{-1}(U^\alpha) \triangle \nu_2^{-1}(V^\alpha)) = 0$ for each $\alpha \in [k]$.*

Proof. The equality $t(\mathbf{F},\mathbf{U}) = t(\mathbf{F},\mathbf{V})$ for each $\mathbf{F}$ implies that $\mathbf{G}(n,\mathbf{U})$ and $\mathbf{G}(n,\mathbf{V})$ have the same distribution Y_n for each n. Let $\mathbf{H} = [\{Y_n\}_{n=1}^{\infty}]$, then Corollary 3.13 implies that there exist separable realizations ϕ_1 and ϕ_2 such that $\mu_{[r]}(\mathbf{H}^\alpha \triangle \phi_1^{-1}(U^\alpha)) = 0$ and

$\mu_{[r]}(\mathbf{H}^\alpha \triangle \phi_2^{-1}(V^\alpha)) = 0$ for each $\alpha \in [k]$, therefore also $\mu_{[r]}(\phi_1^{-1}(U^\alpha) \triangle \phi_2^{-1}(V^\alpha)) = 0$. Set $\mathcal{A} = \sigma(\phi_1^{-1}(\mathcal{L}_{[0,1]^{\mathrm{h}([r])}}), \phi_1^{-1}(\mathcal{L}_{[0,1]^{\mathrm{h}([r])}}))$ that is a separable σ-algebra on $\mathbf{X}^{[r]}$ so by Theorem 3.6 there exists a separable realization ϕ_3 such that for each measurable $A \subset [0,1]^{\mathrm{h}([r])}$ the element $\phi_i^{-1}(A)$ of $\mathcal{A}$ can be represented by a subset of $[0,1]^{\mathrm{h}([r])}$ denoted by $\psi_i(A)$. It is easy to check that the maps ψ_1 and ψ_2 defined this way are structure preserving embeddings from $\mathcal{L}^{\mathrm{h}([r])} \to \mathcal{L}^{\mathrm{h}([r])}$ satisfying $\lambda(\psi_1(U^\alpha) \triangle \psi_2(V^\alpha)) = 0$ for each $\alpha \in [k]$. We conclude that by Lemma 3.11 there are structure preserving ν_1 and ν_2 such that $\lambda(\nu_1^{-1}(U^\alpha) \triangle \nu_2^{-1}(V^\alpha)) = 0$ for each $\alpha \in [k]$. □

The next result is perhaps also meaningful beyond the framework of this paper and is the main contribution in the current section. Recall the definition of the ground state energies (GSE), Definition 3.1 and Definition 3.2.

Theorem 3.15. *For any $J = (J^1, \dots, J^k)$ with J^α being a real r-array of size q for each $\alpha \in [k]$ the parameter of k-colored r-graphs $\mathcal{E}_{r-1}(., J)$ is testable.*

Proof. We may assume that $\|J^\alpha\|_\infty \leq 1$ for every α without losing generality. We proceed by contradiction. Suppose there exist an $\varepsilon > 0$ and a sequence of k-colored r-uniform hypergraphs $\{\mathbf{H}_n\}_{n=1}^\infty$ with $V(\mathbf{H}_n) = [m_n]$ tending to infinity that are such that for each n with probability at least ε we have that $\mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon \leq \mathcal{E}_{r-1}(\mathbb{G}(n, \mathbf{H}_n), J)$. Let $\mathbf{G}_n = (G_n^1, \dots, G_n^k)$ denote the random k-colored hypergraph $\mathbb{G}(n, \mathbf{H}_n)$ for each n with $G_n^\alpha = \mathbb{G}(n, H_n^\alpha)$. The previous event can be reformulated as stating that for each n with probability at least ε there is a partition $\mathcal{P}_n = (P_n^1, \dots, P_n^q)$ of $\binom{[n]}{r-1}$ such that the expression

$$\frac{1}{n^r} \sum_{\alpha=1}^{k} \sum_{i_1,\dots,i_r=1}^{q} J^\alpha(i_1, \dots, i_r) e_{G_n^\alpha}(r; P_n^{i_1}, \dots, P_n^{i_r})$$

is larger than

$$\frac{1}{m_n^r} \sum_{\alpha=1}^{k} \sum_{i_1,\dots,i_r=1}^{q} J^\alpha(i_1, \dots, i_r) e_{H_n^\alpha}(r; R_n^{i_1}, \dots, R_n^{i_r}) + \varepsilon$$

for any partition $\mathcal{R}_n = (R_n^1, \dots, R_n^q)$ of $\binom{[m_n]}{r-1}$.

Let $\mathbf{H}$ denote the ultralimit of the hypergraph sequence $\{\mathbf{H}_n\}_{n=1}^\infty$ that is a k-partition in the measure space $(\mathbf{X}_1^{[r]}, \mathcal{B}_1, \mu_1)$, and let $\sigma_1(S)$ and $\sigma_1(S)^*$ denote the sub-σ-algebras of $\mathcal{B}_1$ corresponding to subsets S of $[r]$. Due to Theorem 3.6 there exists a separable realization $\phi_1 : \mathbf{X}_1^{[r]} \to [0,1]^{\mathrm{h}([r])}$ such that there is a k-colored r-set graphon $\mathbf{W} = (W^1, \dots, W^k)$ satisfying $\mu_1(\phi_1^{-1}(W^\alpha) \triangle \mathbf{H}^\alpha) = 0$ for each $\alpha \in [k]$. Let $\mathbf{G}(s)$ stand for the point-wise ultralimit realization of the $\{\mathbf{G}_n(s)\}_{n=1}^\infty \subset \mathbf{X}_2^{[r]}$ for all $s \in \mathbb{S}$, where $(\mathbb{S}, \mathcal{S}, \nu)$ denotes the underlying joint probability space for the random hypergraphs, and $(\mathbf{X}_2^{[r]}, \mathcal{B}_2, \mu_2)$ is the ultraproduct measure space in the case of the sample sequence, $\sigma_2(S)$ and $\sigma_2(S)^*$ are the corresponding sub-σ-algebras. Note that the ultralimits $\mathbf{G}(s)$ are not k-partitions of the same ultraproduct space as $\mathbf{H}$,

moreover, it is possible that the σ-algebra generated by $\{\mathbf{G}(s)|s \in \mathbb{S}\}$ together with μ_2 form a non-separable measure algebra that prevents us from using Theorem 3.6 directly.

Suppose that for some n we have that $\mathbb{E}\mathcal{E}_{r-1}(\mathbf{G}_n, J) < \mathcal{E}_{r-1}(\mathbf{H}_n, J)+3/4\varepsilon$. This assumption implies by Lemma 3.3 that $\mathbb{P}(\mathcal{E}_{r-1}(\mathbf{G}_n, J) \geq \mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon) \leq \mathbb{P}(\mathcal{E}_{r-1}(\mathbf{G}_n, J) \geq \mathbb{E}\mathcal{E}_{r-1}(\mathbf{G}_n, J) + \varepsilon/4) \leq 2\exp(-\frac{\varepsilon^2 n}{64r^2})$. The last bound is strictly smaller than ε when n is chosen sufficiently large, therefore it contradicts the main assumption for large n. Therefore we can argue that $\mathbb{E}\mathcal{E}_{r-1}(\mathbf{G}_n, J) \geq \mathcal{E}_{r-1}(\mathbf{H}_n, J) + 3/4\varepsilon$ for large n, throwing away a starting piece of the sequence $\{\mathbf{H}_n\}_{n=1}^{\infty}$ we may assume that it holds for all n.

A second application of Lemma 3.3 leads to a lower bound on the probability that $\mathcal{E}_{r-1}(\mathbf{G}_n, J)$ is close to $\mathcal{E}_{r-1}(\mathbf{H}_n, J)$, namely $\mathbb{P}(\mathcal{E}_{r-1}(\mathbf{G}_n, J) \leq \mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon/2) \leq 2\exp(-\frac{\varepsilon^2 n}{64r^2})$. Hence, by invoking the Borel-Cantelli Lemma, we infer that with probability one the event $\mathcal{E}_{r-1}(\mathbf{G}_n, J) \leq \mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon/2$ can occur only for finitely many n, let the M_1 denote the (random) threshold for which is true that $\mathcal{E}_{r-1}(\mathbf{G}_n, J) > \mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon/2$ for every $n \geq M_1$. It follows that $\lim_\omega \mathcal{E}_{r-1}(\mathbf{G}_n, J) > \lim_\omega \mathcal{E}_{r-1}(\mathbf{H}_n, J) + \varepsilon/2$ with probability 1.

Next we will show that with probability one $\mathbf{G}$ is equivalent to $\mathbf{H}$ in the sense that for each k-colored r-graph $\mathbf{F}$ it holds that $t(\mathbf{F}, \mathbf{G}) = t(\mathbf{F}, \mathbf{H})$. Then, since there are countably many test graphs $\mathbf{F}$, we can conclude that the equality holds simultaneously for all $\mathbf{F}$ with probability 1.

We have seen above in the paragraph after (3.2) that for every fixed k-colored r-uniform hypergraph $t(\mathbf{F}, \mathbf{H}) = \lim_\omega t(\mathbf{F}, \mathbf{H}_n)$. On the other hand the subgraph densities in random induced subgraphs are highly concentrated around their mean, that is

$$\mathbb{P}(|t(\mathbf{F}, \mathbf{G}_n) - t(\mathbf{F}, \mathbf{H}_n)| \geq \delta) \leq 2\exp(-\frac{\delta^2 n}{2|V(\mathbf{F})|^2})$$

for any $\delta > 0$, this follows with basic martingale techniques, see Theorem 11 in [6] for the almost identical statement together with a proof. The Borel-Cantelli Lemma implies then for every fixed $\mathbf{F}$ that with probability one for each $\delta > 0$ there exists a (random) $n_0(\delta)$ such that for each $n \geq n_0(\delta)$ it is true that $|t(\mathbf{F}, \mathbf{G}_n) - t(\mathbf{F}, \mathbf{H}_n)| < \delta/2$. Let us fix $\delta > 0$ and $\mathbf{F} \in \mathcal{G}^{r,k}$. Since the set $\{n| \quad |t(F, \mathbf{H}_n) - t(F, \mathbf{H})| < \delta/2\}$ belongs to ω by the definition of the ultralimit function, it holds that $\{n| \quad |t(\mathbf{F}, \mathbf{G}_n) - t(F, \mathbf{H})| < \delta\} \in \omega$ as a consequence of

$$\begin{aligned}&\{n| \quad |t(\mathbf{F}, \mathbf{G}_n) - t(\mathbf{F}, \mathbf{H})| < \delta\}\\ &\quad \supset (\{n| \quad |t(\mathbf{F}, \mathbf{G}_n) - t(\mathbf{F}, \mathbf{H}_n)| < \delta/2\} \cap \{n| \quad |t(\mathbf{F}, \mathbf{H}_n) - t(\mathbf{F}, \mathbf{H})| < \delta/2\}) \in \omega.\end{aligned}$$

Consequently, $\lim_\omega t(\mathbf{F}, \mathbf{G}_n) = t(\mathbf{F}, \mathbf{H})$ with probability one for each $\mathbf{F}$, and the limit equation holds simultaneously for each $\mathbf{F}$ also with probability one, since their number is countable.

Let us pick a realization $\{\mathbf{G}_n(s)\}_{n=1}^{\infty}$ of $\{\mathbf{G}_n\}_{n=1}^{\infty}$ such that it satisfies $\lim_\omega \mathcal{E}_{r-1}(\mathbf{G}_n(s), J) - \lim_\omega \mathcal{E}_{r-1}(\mathbf{H}_n, J) \geq \varepsilon/2$ and $\lim_\omega t(\mathbf{F}, \mathbf{G}_n(s)) = t(\mathbf{F}, \mathbf{H})$ for each $\mathbf{F}$, the preceding discussion implies that such a realization exists, in fact almost all of them are like this. Furthermore, let us consider the sequence of partitions $\mathcal{P}_n = (P_n^1, \ldots, P_n^q)$ of $\binom{[n]}{r-1}$ that realize $\mathcal{E}_{r-1}(\mathbf{G}_n(s), J)$, and define $T_n^{i,j} \subset [n]^r \setminus \mathrm{diag}([n]^r)$ through the inverse images of the projections $A_{T_n^{i,j}} = (p_j^n)^{-1}(A_{P_n^i})$ for $i \in [q]$, $j \in [r]$, and $n \in \mathbb{N}$, where p_j^n is the projection that maps an r-array

of size n onto an $(r-1)$-array by erasing the jth coordinate. Note that the $T_n^{i,j}$'s are not completely symmetric, but are invariant under coordinate permutations from $S_{[r]\setminus\{j\}}$ for the corresponding $j \in [r]$. A further property is that and T_n^{i,j_1} can be obtained from T_n^{i,j_2} swapping the coordinates corresponding to j_1 and j_2.

We additionally define the ultraproducts of these sets by $\mathbf{P}^i = [\{P_n^i\}_{n=1}^\infty] \subset \mathbf{X}_2^{[r-1]}$ and $\mathbf{T}^{i,j} = [\{T_n^{i,j}\}_{n=1}^\infty] \subset \mathbf{X}_2^{[r]}$, it is clear that $\mathbf{T}^{i,j} \in \sigma_2([r] \setminus \{j\})$ for each pair of i and j, so $\cap_{(i,j)\in I}\mathbf{T}^{i,j} \in \sigma_2([r])^*$ for any $I \subset [q] \times [r]$, and that $\mathbf{X}_2^{[r-1]} = \cup_i \mathbf{P}^i$. The same symmetry assumptions apply for the $\mathbf{T}^{i,j}$'s as for the $T_n^{i,j}$'s described above.

We also require the fact that these ultraproduct sets defined above establish a correspondence between the GSE of $\mathbf{G}(s)$ and the ultralimit of the sequence of energies $\{\mathcal{E}_{r-1}(\mathbf{G}_n(s), J)\}_{n=1}^\infty$.

This can be seen as follows: Recall that

$$\mathcal{E}_{r-1}(\mathbf{G}_n(s), J) = \frac{1}{n^r}\sum_{\alpha=1}^{k}\sum_{i_1,\dots,i_r=1}^{q} J^\alpha(i_1,\dots,i_r)|G_n^\alpha \cap (\cap_{j=1}^q T_n^{i_j,j})|$$

This formula together with the identities $[\{G_n^\alpha(s) \cap (\cap_{j=1}^q T_n^{i_j})\}_{n=1}^\infty] = \mathbf{G}^\alpha(s) \cap (\cap_{j=1}^q \mathbf{T}^{i_j,j})$, and that the ultralimit of subgraph densities equals the subgraph density of the ultraproduct imply that

$$\lim_\omega \mathcal{E}_{r-1}(\mathbf{G}_n(s), J) = \sum_{\alpha=1}^{k}\sum_{i_1,\dots,i_r=1}^{q} J^\alpha(i_1,\dots,i_r)\mu_2(\mathbf{G}^\alpha(s) \cap (\cap_{j=1}^q \mathbf{T}^{i_j,j})).$$

Now consider the separable sub-σ-algebra $\mathcal{A}$ of $\mathcal{B}_2$ generated by the collection of the sets $\mathbf{G}^1(s),\dots,\mathbf{G}^k(s),\mathbf{T}^{1,1},\dots,\mathbf{T}^{q,r}$. Then by Theorem 3.6 there exists a separable realization $\phi_2 : \mathbf{X}_2^{[r]} \to [0,1]^{\mathrm{h}([r])}$ and measurable sets $U^1,\dots,U^k,V^{1,1},\dots,V^{q,r}$ such that $\mu_2(\phi_2^{-1}(U^\alpha)\triangle\mathbf{G}^\alpha(s)) = 0$ for each $\alpha \in [k]$ and $\mu_2(\phi_2^{-1}(V^{i,j})\triangle\mathbf{T}^{i,j}) = 0$ for every $i \in [q], j \in [r]$. Additionally, we can modify the $V^{i,j}$'s on a set of measure 0 such that each of them only depends on the coordinates corresponding to the sets in $\mathrm{h}([r] \setminus \{j\})$, is invariant under coordinate permutations induced by elements of $S_{[r]}$ that fix j, and V^{i,j_1} can be obtained from V^{i,j_2} by relabeling the coordinates according to the S_r permutation swapping j_1 and j_2. Also, $(U^1,\dots,U^k)$ form a k-colored r-set graphon $\mathbf{U}$ when we make modifications on null sets. Most importantly, the separable realization ϕ_2 is measure preserving, so we have that

$$\lim_\omega \mathcal{E}_{r-1}(\mathbf{G}_n(s), J) = \sum_{\alpha=1}^{k}\sum_{i_1,\dots,i_r=1}^{q} J^\alpha(i_1,\dots,i_r)\lambda(U^\alpha \cap (\cap_{j=1}^r V^{i_j,j})). \tag{3.3}$$

On the other hand we established that $t(\mathbf{F},\mathbf{G}(s)) = t(\mathbf{F},\mathbf{H})$ for each $\mathbf{F}$,which implies $t(\mathbf{F},\mathbf{U}) = t(\mathbf{F},\mathbf{W})$, therefore the uniqueness statement of Theorem 3.14 ensures the existence of two structure preserving measurable maps $\nu_1,\nu_2 : [0,1]^{\mathrm{h}([r])} \to [0,1]^{\mathrm{h}([r])}$ such that $\lambda(\nu_1^{-1}(W^\alpha)\triangle\nu_2^{-1}(U^\alpha)) = 0$ for each $\alpha \in [k]$.

Now let us define the sets $\mathbf{S}^{i,j} = \phi_1^{-1}(\nu_2(\nu_1^{-1}(V^{i,j})))$, these satisfy exactly the same symmetry properties as the $\mathbf{T}^{i,j}$'s above, by the measure preserving nature of the maps involved we have that

$$\lim_{\omega} \mathcal{E}_{r-1}(\mathbf{G}_n(s), J) = \sum_{\alpha=1}^{k} \sum_{i_1,\dots,i_r=1}^{q} J^{\alpha}(i_1,\dots,i_r)\mu_1(\mathbf{H}^{\alpha} \cap (\cap_{j=1}^{r} \mathbf{S}^{i_j,j})). \tag{3.4}$$

The properties of structure preserving maps imply that $\mathbf{S}^{i,j} \in \sigma_1([r] \setminus \{j\})$ for each i, j, so $\cap_{(i,j)\in I} \mathbf{S}^{i,j} \in \sigma_1([r])^*$ for any $I \subset [q] \times [r]$. Also, the ultraproduct construction makes it possible to assert the existence of a sequence of partitions $\mathcal{R}_n = (R_n^1, \dots, R_n^q)$ of $\binom{[m_n]}{r-1}$ for ω-almost every n such that $\mathbf{S}^{i,j} = [\{(p_j^{m_n})^{-1}(R_n^i)\}_{n=1}^{\infty}]$. But again by the correspondence principle between ultralimits of sequences and ultraproducts in Lemma 3.8 applied to (3.3) and (3.4) we have

$$\lim_{\omega} \mathcal{E}_{\mathcal{R}_n, r-1}(\mathbf{H}_n, J) = \lim_{\omega} \mathcal{E}_{r-1}(\mathbf{G}_n(s), J),$$

which contradicts $\lim_\omega \mathcal{E}_{r-1}(\mathbf{G}_n(s), J) - \lim_\omega \mathcal{E}_{r-1}(\mathbf{H}_n, J) \geq \varepsilon/2$.

□

An immediate consequence is that the above theorem is also true for r-graphons.

Corollary 3.16. *For any $J = (J^1, \dots, J^k)$ with J^α being a real r-array of size l for each $\alpha \in [k]$ there exists for any $\varepsilon > 0$ a $q(\varepsilon)$ integer such that for any k-colored r-set graphon $\mathbf{W}$ and $q \geq q(\varepsilon)$ it holds that*

$$\mathbb{P}(|\mathcal{E}_{r-1}(\mathbf{W}, J) - \mathcal{E}_{r-1}(\mathbb{G}(q, \mathbf{W}), J)| > \varepsilon) < \varepsilon.$$

Proof. We only sketch the proof here, details are left to the reader. The main idea is to find for any fixed $\varepsilon > 0$, and for each k-colored r-set graphon $\mathbf{W}$ a $\mathbf{G} \in \mathcal{G}^{r,k}$ such that their GSE are sufficiently close, and further, the distributions of $\mathbb{G}(q_0(\varepsilon/2), \mathbf{W})$ and $\mathbb{G}(q_0(\varepsilon/2), \mathbf{G})$ are close enough in terms of ε, where q_0 is the sample complexity of $\mathcal{E}_{r-1}(., J)$, whose existence is ensured by Theorem 3.15. Fix $\varepsilon > 0$, and let $\mathbf{W}$ be a k-colored r-set graphon. By measurability for any $\Delta > 0$ there exists an integer l and a k-colored r-set graphon $\mathbf{U}$ such that each U^α is a union of cubes $\times_{S\in \mathrm{h}([r])}[\frac{z_S-1}{l}, \frac{z_S}{l}]$ with $z \in \mathbb{Z}^{\mathrm{h}([r])}$ and $\sum_{\alpha=1}^{k} \|W^\alpha - U^\alpha\|_1 \leq \Delta$. For a fixed, but sufficiently small Δ, let $\mathbf{G}$ be the k-colored r-graph on l vertices that is obtained by randomization form $\mathbf{U}$ using the independent uniform $[0,1]$-valued random variables $(X_S)_{S\in \mathrm{h}([l],r)\setminus \mathrm{h}([l],1)}$. Then by standard large deviations results it follows that the 1-norm of $U^\alpha - W_{\mathbf{G}^\alpha}$ is highly concentrated around 0. By definition, the deviation of the GSE's of two r-graphons can be upper bounded by a constant multiple of their difference in the 1-norm. By Corollary 2.7 the same is true for the total variation distance of the corresponding measures for the fixed sampling depth $q_0(\varepsilon/2)$, as the cut-norm is dominated by the 1-norm. The quantity $\sum_{\alpha=1}^{k} \|W^\alpha - W_{\mathbf{G}^\alpha}\|_1$ can be made arbitrarily small by the above discussion, which proves the result. □

We can derive a substantial property of the cut norm form the above theorem. Recall the definition of the relevant norms, Definition 2.4.

Lemma 3.17. *Let $r \geq 1$. For any $\varepsilon > 0$ and $t \geq 1$ there exists an integer $l_0(r, \varepsilon, t) \geq 1$ a such that for any symmetric r-kernel U that takes values in $[-1, 1]$, and for any integer $l \geq l_0(r, \varepsilon, t)$ it holds with probability at least $1 - \varepsilon$ that*

$$\left| \sup_{Q, t_Q \leq t} \|U\|_{\square, r, Q} - \sup_{Q, t_Q \leq t} \|W_{\mathrm{G}(l,U)}\|_{\square, r, Q} \right| \leq \varepsilon,$$

where the supremum at both places goes over symmetric partitions Q of $[0,1]^{\mathrm{h}([r-1])}$ into at most t classes.

Proof. Let us fix $\varepsilon > 0, r, t \geq 1$, and let U be arbitrary. In this lemma we deal with r-graphons instead of r-set graphons, Fubini's Theorem ensures that we can apply Theorem 3.15 correctly later on.

Showing that there exits an l_0 not depending on U such that for each $l \geq l_0$ it holds that $\sup_{Q, t_Q \leq t} \|U\|_{\square, r, Q} - \sup_{Q, t_Q \leq t} \|W_{\mathrm{G}(l,U)}\|_{\square, r, Q} \leq \varepsilon$ with failure probability at most $\varepsilon/2$ is a routine exercise, we only have to consider a tuple $(S_i)_{i \in [r]}$ of symmetric subsets of $[0,1]^{\mathrm{h}([r-1])}$ and a symmetric partition Q^0 of $[0,1]^{\mathrm{h}([r-1])}$ into at most t classes such that

$$\sup_{Q, t_Q \leq t} \|U\|_{\square, r, Q} = \sum_{j_1, \ldots, j_r = 1}^{t} \left| \int_{\bigcap_{i \in [r]} p^{-1}_{[r] \setminus \{i\}}(S_i \cap Q^0_{j_i})} U(x_{\mathrm{h}([r], r-1)}) \mathrm{d}\lambda(x_{\mathrm{h}([r], r-1)}) \right|,$$

and use Markov's inequality. The difficult part is to show that if l is large enough then for each U it holds that

$$\sup_{Q, t_Q \leq t} \|W_{\mathrm{G}(l,U)}\|_{\square, r, Q} - \sup_{Q, t_Q \leq t} \|U\|_{\square, r, Q} \leq \varepsilon$$

with probability at least $1 - \varepsilon/2$.

First we have to discretize the range of U in order to apply the above result on k-colored r-graphs, Corollary 3.16. Therefore we split the interval $[-1, 1]$ into consecutive intervals $I_1, \ldots, I_k$ of length at most $\varepsilon/4$, let $y_j = \inf I_j$ for each $j \in [k]$, and define the r-kernel $W(x) = \sum_{j=1}^{k} \mathbb{I}_{I_j}(U(x)) y_j$. Then $\|U - W\|_\infty \leq \varepsilon/4$, so therefore $\left| \|U\|_{\square, r, Q} - \|W\|_{\square, r, Q} \right| \leq \varepsilon/4$ and $\left| \|W_{\mathrm{G}(l,U)}\|_{\square, r, Q} - \|W_{\mathrm{G}(l,W)}\|_{\square, r, Q} \right| \leq \varepsilon/4$ for any Q and l. Thus, it suffices to show the existence of an l_0 not depending on U or W such that for each $l \geq l_0$ we have

$$\|W_{\mathrm{G}(l,W)}\|_{\square, r, Q} - \|W\|_{\square, r, Q} \leq \varepsilon/2$$

for each partition symmetric Q of $[0,1]^{\mathrm{h}([r-1])}$ into at most t classes simultaneously with probability at least $1 - \varepsilon/2$.

We can rewrite $\sup_{Q, t_Q \leq t} \|W\|_{\square, r, Q}$ as an optimization problem, more precisely

$$\sup_{Q, t_Q \leq t} \|W\|_{\square, r, Q} \tag{3.5}$$

$$= \sup_{Q, t_Q \leq t} \max_{A \in \mathbb{A}} \sup_{\substack{T_j \subset [0,1]^{\mathrm{h}([r-1])} \\ j \in [r]}} \sum_{i_1, \ldots, i_r = 1}^{t} A(i_1, \ldots, i_r) \int_{[0,1]^{\mathrm{h}([r], r-1)}} W(x_{\mathrm{h}([r], r-1)}) \prod_{j=1}^{r} \mathbb{I}_{T_j \cap Q_{i_j}}(x_{\mathrm{h}([r] \setminus \{j\})}) \mathrm{d}\lambda(x_{\mathrm{h}([r], r-1)}), \tag{3.6}$$

where $\mathbb{A}$ denotes the set of all r-arrays of size t with $\{-1,1\}$ entries, and the set and partitions involved are symmetric.

If we swap the order of the maximization operation on the right of the above formula (3.5), then it can be turned into a generalized energy for each $A \in \mathbb{A}$. In more detail, consider W as a k-colored r-graphon with $W^\alpha = \mathbb{I}_{W=y_\alpha}$ for each $\alpha \in [k]$, with slight abuse of notation we set $W = (W^\alpha)_{\alpha\in[k]}$. We also define the r-array B_0 of size 2^r, indexed by the power set of $[r]$ so that $B_0(i_{S_1}, \dots, i_{S_r})$ is equal to 1 if for every $j \in [r]$ we have $j \in S_j$, and is equal to 0 otherwise. Let $J_A^\alpha = y_\alpha(A \otimes B_0)$ be the tensor product of A and B_0 for each $A \in \mathbb{A}$ multiplied with the scalar y_α with $\alpha \in [k]$, then J_A^α is an r-array of size $2^r t$. It follows that

$$\max_{A\in\mathbb{A}} \mathcal{E}_{r-1}(W, J_A) = \sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \|W\|_{\square,r,\mathcal{Q}}. \tag{3.7}$$

Similarly,

$$\max_{A\in\mathbb{A}} \mathcal{E}_{r-1}(W_{\mathrm{G}(l,W)}, J_A) = \sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \|W_{\mathrm{G}(l,W)}\|_{\square,r,\mathcal{Q}},$$

hence

$$\sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \|W_{\mathrm{G}(l,W)}\|_{\square,r,\mathcal{Q}} - \sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \|W\|_{\square,r,\mathcal{Q}} \le \max_{A\in\mathbb{A}} |\mathcal{E}_{r-1}(W_{\mathrm{G}(l,W)}, J_A) - \mathcal{E}_{r-1}(W, J_A)|.$$

The function $\mathcal{E}_{r-1}(., J_A)$ is testable by Corollary 3.16, say with sample complexity $q_1(\varepsilon, r, l, k)$, so $\sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \|.\|_{\square,r,\mathcal{Q}}$ is testable with sample complexity $l_0(r, \varepsilon, t) = q_1(\varepsilon/|\mathbb{A}|, r, m, 2^r t)$.

□

In fact, we will require the version of Lemma 3.17 for k-tuples r-kernels.

Lemma 3.18. *Let $r, k \ge 1$. For any $\varepsilon > 0$ and $t \ge 1$ there exists an integer $q_{\mathrm{cut}}(r, k, \varepsilon, t) \ge 1$ a such that for any k-tuple of r-kernel $U_1, \dots, U_k$ that take values from $[-1, 1]$, and any integer $q \ge q_{\mathrm{cut}}(r, k, \varepsilon, t)$ it holds with probability at least $1 - \varepsilon$ that*

$$\left| \sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \sum_{j=1}^k \|U_j\|_{\square,r,\mathcal{Q}} - \sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \sum_{j=1}^k \|W_{\mathrm{G}(q,U_j)}\|_{\square,r,\mathcal{Q}} \right| \le \varepsilon$$

where the supremum at both places goes over symmetric partitions $\mathcal{Q}$ of $[0,1]^{\mathrm{h}([r-1])}$ into at most t classes.

Proof. We only sketch the proof as it is almost identical to that of Lemma 3.17. Let $r, k, t \ge 1$ and $\varepsilon > 0$ be fixed, and let $U_1, \dots, U_k$ and q be arbitrary. The lower bound on $\sup_{\mathcal{Q}, t_\mathcal{Q}\le t} \sum_{j=1}^k \|W_{\mathrm{G}(q,U_j)}\|_{\square,r,\mathcal{Q}}$ can be obtained by the same argument as above using Markov's inequality. For the upper bound we again discretize to obtain the r-kernels $W_1, \dots, W_k$ with common range $\{y_i : \alpha \in [m]\}$ such that $\|U_j - W_j\|_\infty \le \frac{\varepsilon}{4k}$ for each $j \in [k]$, hence $m = \frac{8k}{\varepsilon}$

will do. We associate to each W_j and m-colored r-graphon $\mathbf{W}_j$ as above and set J_A^α to $y_\alpha(A \otimes B_0)$, then

$$\max_{A_1,\dots,A_k \in \mathbb{A}} \sup_{\mathcal{Q}, t_\mathcal{Q} \le t} \sum_{j=1}^{k} \mathcal{E}_{\mathcal{Q},r-1}(\mathbf{W}_j, J_{A_j}) = \sup_{\mathcal{Q}, t_\mathcal{Q} \le t} \sum_{j=1}^{k} \|W_j\|_{\square,r,\mathcal{Q}}.$$

Similarly,

$$\max_{A_1,\dots,A_k \in \mathbb{A}} \sup_{\mathcal{Q}, t_\mathcal{Q} \le t} \sum_{j=1}^{k} \mathcal{E}_{\mathcal{Q},r-1}(\mathbf{W}_{\mathrm{G}(q,\mathbf{W}_j)}, J_{A_j}) = \sup_{\mathcal{Q}, t_\mathcal{Q} \le t} \sum_{j=1}^{k} \|W_{\mathrm{G}(q,W_j)}\|_{\square,r,\mathcal{Q}}.$$

The testability of $\sup_{\mathcal{Q}, t_\mathcal{Q} \le t} \sum_{j=1}^{k} \mathcal{E}_{\mathcal{Q},r-1}(\mathbf{W}_j, J_{A_j})$ follows from Theorem 3.15 with a slight modification of the argument for any fixed tuple $A_1, \dots, A_k \in \mathbb{A}$. As the cardinality of $\mathbb{A}$ does not depend on $\mathbf{W}_1, \dots, \mathbf{W}_k$ the statement of the lemma follows. □

4 Auxiliary lemmas

We will require the version of Szemerédi's Regularity Lemma adapted to the Hilbert space setting. Let us recall this variant.

Lemma 4.1. *[13] Let $\mathcal{K}_1, \mathcal{K}_2, \dots$ be arbitrary subsets of a Hilbert space $\mathcal{H}$. Then for every $\varepsilon > 0$ and $f \in \mathcal{H}$ there is an $m \le \frac{1}{\varepsilon^2}$ and there are $f_i \in \mathcal{K}_i$ and $\gamma_i \in \mathbb{R}$ $(1 \le i \le m)$ such that for every $g \in \mathcal{K}_{m+1}$ we have that*

$$|\langle g, f - \sum_{i=1}^{m} \gamma_i f_i \rangle| \le \varepsilon \|f\| \|g\|.$$

We start with the following intermediate version of the regularity lemma for edge k-colored r-graphons, the partition obtained here satisfies stronger conditions than those imposed by the Weak Regularity Lemma [7], and weaker than by Szemerédi's original.

Lemma 4.2. *For every $r \ge 1$, $\varepsilon > 0$, $t \ge 1$, $k \ge 1$ and k-colored r-graphon $\mathbf{W}$ there exists a symmetric partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0,1]^{\mathrm{h}([r-1])}$ into $m \le (2t)^{(rk+1)^{4/\varepsilon^2}} = t_{\mathrm{reg}}(r,k,\varepsilon,t)$ parts and a symmetric $(r, r-1)$- step function $\mathbf{V} \in \mathcal{W}^{r,k}$ with steps from $\mathcal{P}$, such that for any partition $\mathcal{Q}$ of $[0,1]^{\mathrm{h}([r-1])}$ into at most mt classes we have*

$$d_{\square,r,\mathcal{Q}}(\mathbf{W}, \mathbf{V}) \le \varepsilon.$$

Proof. Our lemma is a special case of Lemma 4.1. We set $\mathcal{H}$ to be the space of of k-tuples of real measurable functions on $[0,1]^{\mathrm{h}([r],r-1)}$ with the sum of the component-wise L^2-products as the inner product, this space contains $\mathcal{W}^{r,k}$. Set $s(1) = 1$ and $s(i+1) = s(i)(s(i)t+1)^{rk}$ for each $i \ge 1$ and let $\mathcal{K}_i$ be the set of k-tuples of indicator functions that are $(r, r-1)$ step functions with $s(i)$ number of symmetric steps and taking values from the set $\{-1, 0, 1\}$.

Note that the elements of the $\mathcal{K}_i$'s are not necessarily symmetric as functions, only their steps are required to be such. Further, observe that $s(i) \le (2t)^{(rk+1)^i}$. Now apply Lemma 4.1 with the above setup for $\varepsilon/2$ and $\mathbf{W}$ to obtain $\mathbf{U}$ that satisfies all the conditions of the lemma except for symmetry, in particular

$$\sum_{\alpha=1}^{k} \|U^\alpha - W^\alpha\|_{\square,r,\mathcal{P}} < \varepsilon.$$

Define $\mathbf{V}$ with $V^\alpha(x_{\mathrm{h}([r],r-1)}) = \frac{1}{r!}\sum_{\pi\in S_r} U^\alpha(x_{\pi(\mathrm{h}([r],r-1))})$. The symmetry of $\mathbf{W}$ and the triangle inequality implies that $\mathbf{V}$ is suitable, since

$$\|V^\alpha - W^\alpha\|_{\square,r,\mathcal{P}} \le \frac{1}{r!}\sum_{\pi\in S_r} \|(U^\alpha)^\pi - (W^\alpha)^\pi\|_{\square,r,\mathcal{P}} = \|U^\alpha - W^\alpha\|_{\square,r,\mathcal{P}}$$

for any $\mathcal{P}$ and $\alpha \in [k]$, and $U^\pi(x_{\mathrm{h}([r],r-1)}) = U(x_{\pi(\mathrm{h}([r],r-1))})$.

□

The next lemma is analogous to Lemma 3.2 from [10]. It describes under what metric conditions a k-coloring of a t-colored graphon can be transfered to another one so that the two tk-colored graphons are close in a certain sense. For the sake if completeness we sketch the proof.

Lemma 4.3. *Let $\varepsilon > 0$, $\mathbf{U}$ be a t-colored r-graphon that is an $(r, r-1)$-step function with steps $\mathcal{P} = (P_1, \dots, P_m)$ and $\mathbf{V}$ be a t-colored r-graphon with $d_{\square,r,\mathcal{P}}(U, V) \le \varepsilon$. For any $k \ge 1$ and $\hat{\mathbf{U}}$ a $[t]\times[k]$-colored r-graphon that is an $(r, r-1)$-step function with steps from $\mathcal{P}$ such that $[\hat{\mathbf{U}}, k] = \mathbf{U}$ there exists a k-coloring of $\mathbf{V}$ denoted by $\hat{\mathbf{V}}$ so that*

$$d_{\square,r,\mathcal{P}}(\hat{\mathbf{U}}, \hat{\mathbf{V}}) \le k\varepsilon.$$

Proof. Fix $\varepsilon > 0$, and let $\mathbf{U} = (U^\alpha)_{\alpha\in[t]}$, $\mathbf{V} = (V^\alpha)_{\alpha\in[t]}$ and $\hat{\mathbf{U}} = (U^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ as in the statement of the lemma. Then $\sum_{\alpha=1}^{t} U^\alpha = 1$ and $\sum_{\beta=1}^{k} U^{\alpha,\beta} = U^\alpha$ for each $\alpha \in [t]$. Let us define $\hat{\mathbf{V}} = (V^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ that is a k-coloring of $\mathbf{V}$. Set $V^{\alpha,\beta} = V^\alpha[\mathbb{I}_{U^\alpha=0}\frac{1}{k} + \mathbb{I}_{U^\alpha>0}\frac{U^{\alpha,\beta}}{U^\alpha}]$, it is easy to see that the factor on the right of the formula is a $(r, r-1)$-step function with steps $\mathcal{P} = (P_1, \dots, P_m)$. We estimate the deviation of each pair $U^{\alpha,\beta}$ and $V^{\alpha,\beta}$ from above in the r-cut norm, for this we fix the symmetric $S_1, \dots, S_r \subset [0,1]^{\mathrm{h}([r-1])}$. Then we have

$$\begin{aligned}
\left|\int_{\cap_{l\in[r]} p_l^{-1}(S_l)} U^{\alpha,\beta} - V^{\alpha,\beta}\right| &\le \sum_{\alpha_1,\dots,\alpha_r=1}^{t} \left|\int_{\cap_{l\in[r]} p_l^{-1}(S_l\cap P_{\alpha_l})} U^{\alpha,\beta} - V^{\alpha,\beta}\right| \\
&= \sum_{\alpha_1,\dots,\alpha_r=1}^{t} \left|\int_{\cap_{l\in[r]} p_l^{-1}(S_l\cap P_{\alpha_l})} (U^\alpha - V^\alpha)[\mathbb{I}_{U^\alpha=0}\frac{1}{k} + \mathbb{I}_{U^\alpha>0}\frac{U^{\alpha,\beta}}{U^\alpha}]\right| \\
&\le \|U^\alpha - V^\alpha\|_{\square,r,\mathcal{P}}.
\end{aligned}$$

Taking the maximum over all symmetric measurable r-tuples $S_1, \dots, S_r$ and summing up over all choices of α and β delivers the upper bound we were after. □

5 Proof of the main result

The central tool in the main proof is the following lemma which can also be of independent interest. Informally it states that every coloring of a sampled r-graph can be transferred onto the graphon from which the graph was sampled from, such that another sampling procedure with a much smaller sample size cannot distinguish the two colored objects.

Lemma 5.1. *For every $r \geq 1$, proximity parameter $\delta > 0$, palette sizes $t, k \geq 1$, sampling depth $q_0 \geq 1$ there exists an integer $q_{\mathrm{tw}} = q_{\mathrm{tw}}(r, \delta, q_0, t, k) \geq 1$ such that for every $q \geq q_{\mathrm{tw}}$ the following holds. Let $\mathbf{U} = (U^\alpha)_{\alpha\in[t]}$ be a t-colored r-graphon and let V^α denote $W_{\mathbb{G}(q,U^\alpha)}$ for each $\alpha \in [t]$, also let $\mathbf{V} = (V^\alpha)_{\alpha\in[t]}$, so $\mathbf{W}_{\mathbb{G}(q,\mathbf{U})} = \mathbf{V}$. Then with probability at least $1-\delta$ there exist for every k-coloring $\hat{\mathbf{V}} = (V^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ of $\mathbf{V}$ a k-coloring $\hat{\mathbf{U}} = (U^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ of $\mathbf{U} = (U^\alpha)_{\alpha\in[t]}$ such that we have that*

$$d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}), \mu(q_0, \hat{\mathbf{U}}) \leq \delta.$$

Proof. We proceed by induction with respect to r. The statement is not difficult to verify for $r = 1$. In this case the 1-graphons U^α and V^α can be regarded as indicator functions of measurable subsets B^α and A^α of $[0,1]$ (so for each $\alpha \in [k]$ we have $U^\alpha = \mathbb{I}_{B^\alpha}$ and $V^\alpha = \mathbb{I}_{A^\alpha}$) that form two partitions associated to $\mathbf{U}$ and $\mathbf{V}$ respectively. Note that $(A^\alpha)_{\alpha\in[k]}$ is obtained from $(B^\alpha)_{\alpha\in[k]}$ by the sampling process. A k-coloring corresponds to a refinement of these partitions with each original class being divided into k measurable parts, that is $A^\alpha = \cup^*_{\beta\in[k]} A^{\alpha,\beta}$ and $V^{\alpha,\beta} = \mathbb{I}_{A^{\alpha,\beta}}$. Moreover, $|t(\mathbf{F}, \hat{\mathbf{U}}) - t(\mathbf{F}, \hat{\mathbf{V}})| = |\prod_{l=1}^{q_0} \lambda(B^{\mathbf{F}(l)}) - \prod_{l=1}^{q_0} \lambda(A^{\mathbf{F}(l)})|$ for any of k-coloring $\hat{\mathbf{U}}$ of $\mathbf{U}$ and for any $[t]\times[k]$-colored $\mathbf{F}$ on q_0 vertices. We may define a suitable coloring by partitioning each of the sets B^α into parts $(B^{\alpha,\beta})_{\beta\in[k]}$ so that the classes satisfy $\lambda(B^{\alpha,\beta}) = \lambda(B^\alpha)\frac{\lambda(A^{\alpha,\beta})}{\lambda(A^\alpha)}$ when $\lambda(A^\alpha) > 0$, and $\lambda(B^{\alpha,\beta}) = \lambda(B^\alpha)\frac{1}{k}$ otherwise for each $\beta \in [k]$. Then by setting $U^{\alpha,\beta} = \mathbb{I}_{B^{\alpha,\beta}}$ and $\hat{\mathbf{U}} = (U^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ we have that

$$d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}), \mu(q_0, \hat{\mathbf{U}}) = \frac{1}{2} \sum_{\mathbf{F}:|V(\mathbf{F})|=q_0} |t(\mathbf{F}, \hat{\mathbf{U}}) - t(\mathbf{F}, \hat{\mathbf{V}})| \leq \frac{q_0^{k+1}}{2} \max_{\alpha\in[k]} |\lambda(A^\alpha) - \lambda(B^\alpha)|,$$

where the sum runs over all $[t]\times[k]$-colored 1-graphs $\mathbf{F}$ on q_0 vertices.

The probability that for a fixed $\alpha \in [t]$ the deviation $|\lambda(A^\alpha) - \lambda(B^\alpha)|$ exceeds $\frac{2\delta}{q_0^{k+1}}$ is at most $2\exp(-\frac{4\delta^2 q}{3q_0^{2k+2}})$, the union bound gives the upper bound $\exp(\ln 2 + t - \frac{4\delta^2 q}{3q_0^{2k+2}})$ for the probability that

$$d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}), \mu(q_0, \hat{\mathbf{U}}) \leq \delta$$

fails for our particular choice for the coloring $\hat{\mathbf{U}}$ of $\mathbf{U}$. We note that the failure probability can be made arbitrary small with the right choice of q, so in particular smaller than δ, therefore $q_{\mathrm{tw}}(1, \delta, q_0, t, k) = \frac{(t+\ln 2-\ln\delta)3q_0^{2k+2}}{4\delta^2}$ that satisfies the conditions of the lemma.

Now assume that we have already verified the statement of the lemma for $r-1$ and any other choice of the other parameters of q_{tw}. Let us proceed to the proof of the case for r-graphons, therefore let $\delta > 0$, $t, k, q_0 \geq 1$ be arbitrary and fixed, q to be determined below

and $\mathbf{U}$, $\mathbf{V}$, and $\hat{\mathbf{V}}$ as in the condition of the lemma. We start by explicitly constructing a k-coloring $\hat{\mathbf{U}}$ for $\mathbf{U}$, in the second part of the proof we verify that the construction is suitable.

In a nutshell, we proceed as follows. We approximate $\hat{\mathbf{V}}$ by the step function $\hat{\mathbf{Z}}$, and set $\mathbf{Z} = [\hat{\mathbf{Z}}, k]$, and also approximate $\mathbf{U}$ by $\mathbf{W}_1$. Let $\mathbf{W}_2$ be the sampled version of $\mathbf{W}_1$ generated by the same process as $\mathbf{V}$. This way $\mathbf{W}_2$ and $\mathbf{Z}$ are close, hence we can color $\mathbf{W}_2$ using the coloring $\hat{\mathbf{Z}}$ of $\mathbf{Z}$ to obtain $\hat{\mathbf{W}}_2$. The coloring $\hat{\mathbf{W}}_2$ is then transferred onto $\mathbf{W}_1$ using the induction hypothesis applied to the marginals of the step sets of $\mathbf{W}_1$ and $\mathbf{W}_2$ to get $\hat{\mathbf{W}}_1$ with $[\hat{\mathbf{W}}_1, k] = \mathbf{W}_1$. Finally we color $\mathbf{U}$ exploiting the proximity of $\mathbf{U}$ and $\mathbf{W}_1$ and the colored $\hat{\mathbf{W}}_1$.

Our construction may fail to meet the criteria of the lemma, this can be caused at two points in the above outline. For one, it may happen, that $\mathbf{W}_2$ does not approximate $\mathbf{V}$ well enough, and secondly, when we transfer $\hat{\mathbf{W}}_2$ onto $\mathbf{W}_1$ using the induction hypothesis with $r-1$, as the current lemma leaves space for probabilistic error. These two events are independent from the particular choice of $\hat{\mathbf{V}}$ and their probability can be made sufficiently small, we aim for to show this. We proceed now to the technical details.

Let $\Delta = \frac{\delta r!}{4k(kt)^{q_0} q_0^r}$. Set $t_2 = t_{\mathrm{reg}}(r, tk, \Delta, 1)$ and $t_1 = t_{\mathrm{reg}}(r, t, \Delta/2, t_2)$, and define $q_{\mathrm{tw}}(r, \delta, q_0, t, k) = \max\{q_{\mathrm{tw}}(r-1, \delta/4, q_0, t_1, t_2), q_{\mathrm{cut}}(r, t, \Delta/2, t_1 t_2)\}$. Let $q \geq q_{\mathrm{tw}}(r, \delta, q_0, t, k)$ be arbitrary.

First we apply Lemma 4.2 with proximity parameter $\Delta/2$ to the t-colored r-graphon $\mathbf{U}$, the lemma ensures the existence of a symmetric partition $\mathcal{P} = (P_1, \dots, P_{t_1})$ of $[0,1]^{\mathrm{h}([r-1])}$ with $t_{\mathcal{P}} \leq t_1$ and a t-colored symmetric step function $\mathbf{W}_1 = (W_1^1, \dots, W_1^t)$ with steps in $\mathcal{P}$ that satisfies $\sup_{\mathcal{Q}, t_{\mathcal{Q}} \leq t_{\mathcal{P}} t_2} d_{\square, r, \mathcal{Q}}(\mathbf{W}_1, \mathbf{U}) \leq \Delta/2$, where the supremum runs over all symmetric partitions $\mathcal{Q}$ of $[0,1]^{\mathrm{h}([r-1])}$ with at most $t_{\mathcal{P}} t_2$ classes. Applying structure preserving transformations to $[0,1]^{\mathrm{h}([r-1])}$ the classes of $\mathcal{P}$ can be considered as piled up, meaning that for each $y \in [0,1]^{\mathrm{h}([r-1], r-2)}$ the fibers $\{y\} \times [0,1]$ are partitioned by the intersections with the classes of $\mathcal{P}$ into intervals $[0, a_1), [a_1, a_2), \dots, [a_{t_1-1}, a_{t_1}]$ with $\{y\} \times [a_{j-1}, a_j) = (\{y\} \times [0,1]) \cap P_j$. We introduce the r-dimensional real arrays $A_1, \dots, A_t$ in order to describe the explicit form of the W_1^α's. So,

$$W_1^\alpha(x_{\mathrm{h}([r], r-1)}) = \sum_{i_1, \dots, i_r = 1}^{t_{\mathcal{P}}} A_\alpha(i_1, \dots, i_r) \prod_{l=1}^{r} \mathbb{I}_{P_{i_l}}(x_{\mathrm{h}([r] \setminus \{l\})}).$$

Let $\mathbf{W}_2 = (W_2^1, \dots, W_2^t)$ denote $\mathbb{G}(q, \mathbf{W}_1)$, so W_2^α stands for $\mathbb{G}(q, W_1^\alpha)$ for each $\alpha \in [t]$, then Lemma 3.18 implies that for $q \geq q_{\mathrm{cut}}(r, t, \Delta/2, t_1 t_2)$ it holds that

$$\sup_{\mathcal{Q}, t_{\mathcal{Q}} \leq t_{\mathcal{P}} t_2} d_{\square, r, \mathcal{Q}}(\mathbf{W}_2, \mathbf{V}) \leq \Delta,$$

with probability at least $1 - \Delta/2$, since $t_{\mathcal{P}} \leq r_1$. Also,

$$W_2^\alpha(x_{\mathrm{h}([r], r-1)}) = \sum_{i_1, \dots, i_r = 1}^{t_{\mathcal{P}}} A_\alpha(i_1, \dots, i_r) \prod_{l=1}^{r} \mathbb{I}_{P'_{i_l}}(x_{\mathrm{h}([r] \setminus \{l\})}),$$

for each $\alpha \in [t]$ and

$$P'_j = \cup_{(p_1,\dots,p_{r-1})\in I_j} [\frac{p_1-1}{q}, \frac{p_1}{q}] \times \cdots \times [\frac{p_r-1}{q}, \frac{p_r}{q}] \times [0,1] \times \cdots \times [0,1]$$

with $I_j = \{(p_1,\dots,p_{r-1}) : X_{r[\{p_1,\dots,p_{r-1}\}]} \in P_j\}$ for every $j \in [t_{\mathcal{P}}]$. Note that $\mathcal{P}' = (P'_j)_{j\in[t_{\mathcal{P}}]}$ is a symmetric partition.

We apply now Lemma 4.2 with proximity parameter Δ in order to approximate the $[t]\times[k]$-colored r-graph $\hat{\mathbf{V}} = (V^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$, the outcome is a $[t]\times[k]$-colored step function $\hat{\mathbf{Z}} = (Z^{\alpha,\beta})_{\alpha\in[t],\beta\in[k]}$ with symmetric steps in $\mathcal{R} = (R_1,\dots,R_{t_2})$ of $[0,1]^{\mathrm{h}([r-1])}$ with $t_{\mathcal{R}} \le t_2$ that satisfies

$$\sup_{Q,t_Q\le t_{\mathcal{R}}} d_{\square,r,Q}(\hat{\mathbf{V}},\hat{\mathbf{Z}}) \le \Delta.$$

We introduce the t-colored step function $\mathbf{Z} = [\hat{\mathbf{Z}}, k]$ that is the k-discoloring of $\hat{\mathbf{Z}}$ that has steps in $\mathcal{R}$ and note that

$$\sup_{Q,t_Q\le t_{\mathcal{R}}} d_{\square,r,Q}(\mathbf{V},\mathbf{Z}) \le \Delta,$$

and therefore

$$\sup_{Q,t_Q\le t_{\mathcal{R}}} d_{\square,r,Q}(\mathbf{Z},\mathbf{W}_2) \le 2\Delta. \tag{5.1}$$

Define the r-arrays $B_1,\dots,B_t$ such that for each $\alpha\in[t]$ it holds that

$$Z^{\alpha}(x_{\mathrm{h}([r],r-1)}) = \sum_{i_1,\dots,i_r=1}^{t_{\mathcal{R}}} B_\alpha(i_1,\dots,i_r) \prod_{l=1}^{r} \mathbb{I}_{R_{i_l}}(x_{\mathrm{h}([r]\setminus\{l\})}),$$

further define also the r-arrays $(B^\beta_\alpha)_{\alpha\in[t],\beta\in[k]}$ such that

$$Z^{\alpha,\beta}(x_{\mathrm{h}([r],r-1)}) = \sum_{i_1,\dots,i_r=1}^{t_{\mathcal{R}}} B^\beta_\alpha(i_1,\dots,i_r) \prod_{l=1}^{r} \mathbb{I}_{R_{i_l}}(x_{\mathrm{h}([r]\setminus\{l\})})$$

for each $\alpha\in[t],\beta\in[k]$. Clearly, $B_\alpha(i_1,\dots,i_r) = \sum_{\beta=1}^{k} B^\beta_\alpha(i_1,\dots,i_r)$.

Our aim next is to find a k-coloring of $\mathbf{W}_2$ so that the new tk-colored r-graphon obtained is close to $\hat{\mathbf{Z}}$. In order to produce the coloring we apply Lemma 4.3 relying on (5.1), hence we obtain $\hat{\mathbf{W}}_2$ with $[\hat{\mathbf{W}}_2,k] = \mathbf{W}_2$. The proximity between the two tk-colored r-graphs can be quantified by

$$\sup_{Q,t_Q\le t_{\mathcal{R}}} d_{\square,r,Q}(\hat{\mathbf{Z}},\hat{\mathbf{W}}_2) \le 2k\Delta.$$

The graphon $\hat{\mathbf{W}}_2$ is a symmetric step function with steps that form the coarsest partition that both refines $\mathcal{P}'$ and $\mathcal{R}$, we denote this symmetric partition of $[0,1]^{\mathrm{h}([r-1])}$ by $\mathcal{S}$, its number of classes satisfies $t_{\mathcal{S}} = t_{\mathcal{P}'}t_{\mathcal{R}} \le t_1t_2$.

Let us define the $t_{\mathcal{P}}$-colored $(r-1)$-graphon $\mathbf{w} = (w^1,\dots,w^{t_{\mathcal{P}}})$ that is obtained from the classes of the partition $\mathcal{P}$ by integrating along the coordinate corresponding to the

set $[r-1]$, that is $w^i(x_{\mathrm{h}([r-1],r-2)}) = \int_0^1 \mathbb{I}_{P_i}(x_{\mathrm{h}([r-1])})\mathrm{d}x_{[r-1]}$. In the same way we define the $t_{\mathcal{P}}$-colored $(r-1)$-graphon $\mathbf{u} = (u^1, \ldots, u^{t_{\mathcal{P}}})$ corresponding to the partition $\mathcal{P}'$, as well as the $[t_{\mathcal{P}}] \times [t_{\mathcal{R}}]$-colored $\hat{\mathbf{u}} = (u^{i,j})_{i\in[t_{\mathcal{P}}],j\in[t_{\mathcal{R}}]}$, where it holds that $\mathbf{u} = [\hat{\mathbf{u}}, t_{\mathcal{R}}]$ and $\hat{\mathbf{u}}$ is the $t_{\mathcal{R}}$-coloring of $\mathbf{u}$ corresponding to the partition $\mathcal{S}$. Note that $\mathbf{w}$,$\mathbf{u}$, and $\hat{\mathbf{u}}$ satisfy the usual symmetries, since their origin partitions were symmetric. As the partition $\mathcal{P}'$ was constructed via the same sampling procedure as $\mathbf{V}$ and $\mathbf{W}_2$, therefore it holds that $\mathbf{u} = \mathbb{G}(q, \mathbf{w})$ and $u^i = \mathbb{G}(q, w^i)$ for each $i \in [t_{\mathcal{P}}]$.

We can assert that due to the induction hypothesis there exists a $t_{\mathcal{R}}$-coloring $\hat{\mathbf{w}} = (w^{i,j})_{i\in[t_{\mathcal{P}}],j\in[t_{\mathcal{R}}]}$ of $\mathbf{w}$ that satisfies

$$d_{\mathrm{tw}}(\mu(q_0, \mathbf{w}), \mu(q_0, \mathbf{u}) \leq \delta/4$$

with probability at least $1 - \delta/4$ for $q \geq q_{\mathrm{tw}}(r-1, \delta/4, q_0, t_1, t_2)$.

We construct a k-coloring for $\mathbf{W}_1$ next. Recall the proof of Lemma 4.3, therefore we have that

$$W_2^{\alpha,\beta}(x_{\mathrm{h}([r],r-1)}) = \sum_{i_1,\ldots,i_r=1}^{t_{\mathcal{P}}} \sum_{j_1,\ldots,j_r=1}^{t_{\mathcal{R}}} A_\alpha(i_1, \ldots, i_r) \left[\frac{B_\alpha^\beta(j_1, \ldots, j_r)}{B_\alpha(j_1, \ldots, j_r)} \mathbb{I}_{B_\alpha>0} + \frac{1}{k}\mathbb{I}_{B_\alpha=0} \right] \prod_{m=1}^{r} \mathbb{I}_{P'_{i_m}\cap R_{j_m}}(x_{\mathrm{h}([r]\setminus\{m\})}), \tag{5.2}$$

and set $A_\alpha^\beta((i_1, j_1), \ldots, (i_r, j_r)) = A_\alpha(i_1, \ldots, i_r) \left[\frac{B_\alpha^\beta(j_1,\ldots,j_r)}{B_\alpha(j_1,\ldots,j_r)} \mathbb{I}_{B_\alpha>0} + \frac{1}{k}\mathbb{I}_{B_\alpha=0} \right]$ for all $\alpha \in [t]$, $\beta \in [k]$ and $((i_1, j_1), \ldots, (i_r, j_r)) \in ([t_{\mathcal{P}}] \times [t_{\mathcal{R}}])^r$.

We utilize the $t_{\mathcal{R}}$-coloring $\hat{\mathbf{w}}$ of the $(r-1)$-graphons $\mathbf{w}$ to construct a refined partition of $\mathcal{P}$ that resembles $\mathcal{S}$ in order to enable the construction of a k-coloring of $\mathbf{W}_1$ along the same lines as in (5.2). Let

$$P_{i,j} = \{x \in [0,1]^{\mathrm{h}([r-1])} : \\ \sum_{l=1}^{i-1} w^l(x_{\mathrm{h}([r-1],r-2)}) + \sum_{l=1}^{j-1} w^{i,l}(x_{\mathrm{h}([r-1],r-2)}) \leq x_{[r-1]} < \sum_{l=1}^{i-1} w^l(x_{\mathrm{h}([r-1],r-2)}) + \sum_{l=1}^{j} w^{i,l}(x_{\mathrm{h}([r-1],r-2)})\}$$

for each $i \in [t_{\mathcal{P}}]$ and $j \in [t_{\mathcal{R}}]$. Let $\mathcal{P}'' = (P_{i,j})_{i\in[t_{\mathcal{P}}],j\in[t_{\mathcal{R}}]}$.

Clearly, $(P_{i,j})_{j\in[t_{\mathcal{R}}]}$ is a $t_{\mathcal{R}}$-partition of the set P_i, and $w^{i,j}(x_{\mathrm{h}([r-1],r-2)}) = \int_{P_{i,j}} \mathrm{d}x_{\mathrm{h}([r-1])}$. We are able now to describe the k-coloring of the $\mathbf{W}_1$, define

$$W_1^{\alpha,\beta}(x_{\mathrm{h}([r],r-1)}) = \sum_{i_1,\ldots,i_r=1}^{t_{\mathcal{P}}} \sum_{j_1,\ldots,j_r=1}^{t_{\mathcal{R}}} A_\alpha^\beta((i_1, j_1), \ldots, (i_r, j_r)) \prod_{m=1}^{r} \mathbb{I}_{P_{i_m,j_m}}(x_{\mathrm{h}([r]\setminus\{m\})}). \tag{5.3}$$

Note that $\hat{\mathbf{W}}_1$ is a step functions with a step partition that refines $\mathcal{P}$ into $\mathcal{P}''$, but the regularity property of $\mathbf{W}_1$ allows for

$$d_{\square,r,\mathcal{P}''}(\mathbf{U}, \mathbf{W}_1) \leq \Delta/2.$$

Finally we employ Lemma 4.3 that produces a k-coloring $\hat{\mathbf{U}}$ of $\mathbf{U}$, so that $\hat{\mathbf{U}}$ is $[t] \times [k]$-colored, and

$$d_{\square,r,\mathcal{P}''}(\hat{\mathbf{U}}, \hat{\mathbf{W}}_1) \leq \frac{k\Delta}{2}.$$

It remains to show that this coloring satisfies the requirements of the current lemma for a large enough q.

In the first step of the coloring construction we employed the r-graphon version of the intermediate regularity lemma, Lemma 4.2, therefore we can assert that for each tk-colored $\mathbf{F}$ we have by means of Lemma 2.6 that

$$\sum_{\mathbf{F} \in \mathcal{G}^{r,kt}_{q_0}} |t(\mathbf{F}, \hat{\mathbf{V}}) - t(\mathbf{F}, \hat{\mathbf{Z}})| \leq \frac{(kt)^{q_0^r} q_0^r}{r!} d_{\square,r}(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) \leq \frac{\delta}{4k},$$

so we can conclude that $d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{V}}), \mu(q_0, \hat{\mathbf{Z}})) \leq \frac{\delta}{8k}$.

In the next step, as a consequence of Lemma 4.3 and Corollary 2.7, we have for $\hat{\mathbf{W}}_2$ that $d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{Z}}), \mu(q_0, \hat{\mathbf{W}}_2)) \leq \delta/4$.

We will next elaborate on the correctness of the inductive step of the construction. Let us consider the tk-colored random r-graph $\mathbb{G}(q_0, \hat{\mathbf{W}}_1)$, it is generated by the independent uniformly distributed $[0,1]$-valued random variables $\{Y_S : S \in \mathrm{h}([q_0], r)\}$. The color of each edge $e = \{e_1, \ldots, e_2\} \in \binom{[q_0]}{r}$ is decided by determining first the unique tuple (up to coordinate permutations) $((i_1, j_1), \ldots, (i_r, j_r)) \in ([t_{\mathcal{P}}] \times [t_{\mathcal{R}}])^r$ such that $(Y_S)_{S \in \mathrm{h}(e \setminus \{e_l\})} \in P_{i_l, j_l}$, and then check for which pair $\alpha \in [t]$, $\beta \in [k]$ it holds that

$$\sum_{l=1}^{\alpha-1} A_l((i_1, j_1), \ldots, (i_r, j_r)) + \sum_{l=1}^{\beta-1} A^l_\alpha((i_1, j_1), \ldots, (i_r, j_r)) < Y_e$$
$$\leq \sum_{l=1}^{\alpha-1} A_l((i_1, j_1), \ldots, (i_r, j_r)) + \sum_{l=1}^{\beta} A^l_\alpha((i_1, j_1), \ldots, (i_r, j_r)),$$

then add the color (α, β) to e with corresponding index. It is convenient to view this process as first randomly $t_{\mathcal{P}''}$-coloring an $(r-1)$-uniform template hypergraph $\mathbf{G}_1$, whose edges are the simplices of the original edges, here we add a color (i, j) to an $(r-1)$-edge e' whenever $(Y_S)_{S \in \mathrm{h}(e')} \in P_{i,j}$, and conditioned on $\mathbf{G}_1$ we subsequently make independent choices for each edge to determine their color based on the arrays A^β_α by means of the random variables $\{Y_S : S \in \binom{[q_0]}{r}\}$ at the top level.

Let us turn to the tk-colored $\mathbb{G}(q_0, \hat{\mathbf{W}}_2)$, the above description of the random process generating this object remains conceptually valid also for this random graph, the r-arrays A^β_α are identical to the case above, only the partition $\mathcal{P}''$ has to be altered to $\mathcal{S}$. Similarly as above, we introduce the random $(r-1)$-uniform $t_{\mathcal{P}''}$-colored hypergraph $\mathbf{G}_2$ that is generated as above adapted to $\mathbb{G}(q_0, \hat{\mathbf{W}}_2)$. That means that the $(r-1)$-edges are colored by indices of the classes that form the partition $\mathcal{S}$ through the process that generates $\mathbb{G}(q_0, \hat{\mathbf{W}}_2)$, see above. The key observation here is that conditioned on $\mathbf{G}_1 = \mathbf{G}_2$, one can

couple $\mathbb{G}(q_0, \hat{\mathbf{W}}_1)$ and $\mathbb{G}(q_0, \hat{\mathbf{W}}_2)$ so that the two random graphs coincide with conditional probability 1. Recall that a coupling is only another name for a joint probability space for the two random objects with the marginal distributions following $\mu(q_0, \mathbf{W}_1)$ and $\mu(q_0, \mathbf{W}_2)$ respectively. As the conditional distributions for the choices of colors for the r-edges are identical provided that $\mathbf{G}_1 = \mathbf{G}_2$ the coupling is trivial. In order to construct a good unconditional coupling we require another coupling, now of $\mathbf{G}_1$ and $\mathbf{G}_2$, so that $\mathbb{P}(\mathbf{G}_1 \neq \mathbf{G}_2)$ is negligible small for our purposes, and whose existence is exactly what the induction hypothesis ensures, when q is large enough.

As $q \geq q_{\mathrm{tw}}(r-1, \delta/4, q_0, t_1, t_2)$, the induction hypothesis enables us to use that there exist for any $\hat{\mathbf{u}}$ a $\hat{\mathbf{w}}$ so that $d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{u}}), \mu(q_0, \hat{\mathbf{w}})) \leq \delta/4$ holds with probability at least $1 - \delta/4$ for each $\hat{\mathbf{u}}$ simultaneously, which in turn implies that there is a coupling of the t_1t_2-colored random $(r-1)$-graphs $\mathbf{G}_1$ and $\mathbf{G}_2$ so that $\mathbb{P}(\mathbf{G}_1 \neq \mathbf{G}_2) \leq \delta/2$.

It follows that there exists a coupling of $\mathbb{G}(q_0, \hat{\mathbf{W}}_1)$ and $\mathbb{G}(q_0, \hat{\mathbf{W}}_2)$ such that $\mathbb{P}(\mathbb{G}(q_0, \hat{\mathbf{W}}_1) \neq \mathbb{G}(q_0, \hat{\mathbf{W}}_2)) \leq \delta/2$ due to the discussion above, which in turn implies

$$d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}_1), \mu(q_0, \hat{\mathbf{W}}_2)) \leq \delta/4.$$

Since $\hat{\mathbf{W}}_1$ has at most $t_{\mathcal{P}}t_2$ steps, another application of Lemma 4.3 provides the bound

$$d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}), \mu(q_0, \hat{\mathbf{U}})) \leq \delta/16,$$

as $d_{\square,r}(\hat{\mathbf{U}}, \hat{\mathbf{W}}_1) \leq \frac{k\Delta}{2}$.

Evoking the triangle inequality and summing up the upper bounds on the respective deviations we conclude that

$$\begin{aligned} d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{V}}), \mu(q_0, \hat{\mathbf{U}})) &\leq d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{V}}), \mu(q_0, \hat{\mathbf{Z}})) + d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{Z}}), \mu(q_0, \hat{\mathbf{W}}_2)) \\ &+ d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}_2), \mu(q_0, \hat{\mathbf{W}}_1)) + d_{\mathrm{tw}}(\mu(q_0, \hat{\mathbf{W}}_1), \mu(q_0, \hat{\mathbf{U}})) \leq \left(\frac{1}{8k} + \frac{1}{4} + \frac{1}{4} + \frac{1}{16}\right)\delta < \delta, \end{aligned}$$

the overall error probability is at most $\delta/2 + \Delta/2$, which is at most δ.

□

With Lemma 5.1 at hand we can overcome the difficulties caused by properties of the r-cut norm for $r \geq 3$ in contrast to the case $r = 2$, we turn to prove the main result of the paper.

Proof of Theorem 2.3. We regard simple hypergraphs as 2-colored r-graphs, in the following the term simple should be understood this way at each appearance. Let the $2k$-colored witness parameter of the nondeterministically testable r-graph parameter f be denoted by g, whose sample complexity is at most $q_g(\varepsilon)$ for each proximity parameter $\varepsilon > 0$. Set $d(r, \varepsilon, q_0, k, t) = \frac{[q_g(\varepsilon)^r \ln(tk) - \ln(\varepsilon)][2(tk)^{q_0^r} q_0^2]}{\varepsilon^2}$.Let $\varepsilon > 0$ be fixed and define $q_f(\varepsilon) = \max\{q_{\mathrm{tw}}(r, \varepsilon/4, q_g(\varepsilon/4), k, 3); \frac{4}{\varepsilon} q_g^2(\varepsilon/2); d(r, \varepsilon/4, q_g(\epsilon/4), k, 2)\}$. We will show that for every $q \geq q_f(\varepsilon)$ the condition

$$\mathbf{P}(|f(G) - f(\mathbb{G}(q_f(\varepsilon), G)| > \varepsilon) < \varepsilon.$$

is satisfied for each G. Let $q \geq q_f(\varepsilon)$ arbitrary but fixed and G be a fixed simple graph on n vertices.

First we show that $f(\mathbb{G}(q,G)) \geq f(G) - \varepsilon/4$ with probability at least $1 - \varepsilon/4$. For this let us select a k-coloring $\mathbf{G}$ of G such that $f(G) = g(\mathbf{G})$, then the random k-colored graph $\mathbf{F} = \mathbb{G}(q,\mathbf{G})$ is a k-coloring of $\mathbb{G}(q,G)$, therefore $f(\mathbb{G}(q,G)) \geq g(\mathbf{F})$, but since $q \geq q_q(\varepsilon/4)$ we know from the testability of g that $g(\mathbf{F}) \geq g(\mathbf{G}) - \varepsilon/4$ with probability at least $1-\varepsilon/4$, which verifies our claim.

The more difficult part is to show that $f(\mathbb{G}(q,G)) \leq f(G) + \varepsilon$ with failure probability at most $\varepsilon/2$. Let us denote the random r-graph $\mathbb{G}(q,G)$ by F. We claim that with probability at least $1 - \varepsilon/2$ there exists for any k-coloring $\mathbf{F}$ of F there exists a k-coloring $\mathbf{G}$ of G such that $|g(\mathbf{F}) - g(\mathbf{G})| \leq \varepsilon$, this suffices to verify the statement of the theorem.

Our proof exploits that the difference of the g values between two colored r-graphs $\mathbf{F}$ and $\mathbf{G}$ can be upper bounded by

$$|g(\mathbf{F}) - g(\mathbf{G})| \leq |g(\mathbf{F}) - g(\mathbb{G}(q_g(\varepsilon/4),\mathbf{F}))| + |g(\mathbf{G}) - g(\mathbb{G}(q_g(\varepsilon/4),\mathbf{G}))| \leq \varepsilon/2,$$

whenever there exists a coupling of the two random $2k$-colored r-graphs $\mathbb{G}(q_g(\varepsilon/4),\mathbf{G})$ and $\mathbb{G}(q_g(\varepsilon/4),\mathbf{F})$ appearing in the above formula such that they are equal with probability larger than $\varepsilon/2$. Set $q_0 = q_g(\varepsilon/4)$. We will show that with high probability fore every $\mathbf{F}$ there exists a $\mathbf{G}$ that satisfies the previous conditions.

Recall that coupling is a probability space together with the random r-graphs $\mathbf{G}_1$ and $\mathbf{G}_2$ defined on it such that $\mathbf{G}_1$ has the same marginal distribution as $\mathbb{G}(q_0,\mathbf{G})$ and $\mathbf{G}_2$ has the same as $\mathbb{G}(q_0,\mathbf{F})$, their joint distribution is constructed in a way that serves our current purposes by maximizing the probability that they coincide. When the target spaces are finite as in our case then a coupling that satisfies this condition can be easily constructed whenever $d_{\mathrm{tw}}(\mu(q_0,\mathbf{G}),\mu(q_0,\mathbf{F})) \leq 1 - \varepsilon/2$, see (2.3).

By Lemma 5.1 for 3-colored r-graphs (there are 3 types of entries in the graphon representation of simple r-graphs, edges, non-edges, and diagonal elements) it follows that with probability at least $1 - \varepsilon/4$ for each $\mathbf{F}$ there exists a $3k$-colored $\mathbf{U}$ with $[\mathbf{U},k] = W_G$ such that $d_{\mathrm{tw}}(\mu(q_0,\mathbf{U}),\mu(q_0,\mathbf{W_F})) \leq \varepsilon/4$. Let us condition on this event and let $\mathbf{F}$ be fixed. From (5.1) we know that $d_{\mathrm{tw}}(\mu(q_0,\mathbf{G}),\mu(q_0,\mathbf{W_G})) \leq q_0^2/n \leq \varepsilon/4$ and $d_{\mathrm{tw}}(\mu(q_0,\mathbf{F}),\mu(q_0,\mathbf{W_F})) \leq q_0^2/q \leq \varepsilon/4$. Further, it follows from our condition above that there exists a $3k$-colored $\mathbf{U}$ that induces a fractional coloring of G, and $d_{\mathrm{tw}}(\mu(q_0,\mathbf{U}),\mu(q_0,\mathbf{W_F})) \leq \varepsilon/4$. It remains to produce a $3k$-coloring of W_G from any fixed $3k$-colored $\mathbf{U}$ (k of the colors of $\mathbf{U}$ correspond exclusively to diagonal cubes, so can be neglected). We do this by randomization, let $(X_{\{i\}})_{i\in[n]}$ be independent uniform random variables distributed on $[\frac{i-1}{n},\frac{i}{n}]$, and let $(X_S)_{S\in\mathrm{h}([n],r)\setminus\mathrm{h}([n],1)}$ be independent uniform random variables on $[0,1]$. We can define $\mathbf{W_G}$ to take the color $\mathbf{U}(X_{\mathrm{h}(e)})$ on the set $[\frac{e_1-1}{n},\frac{e_1}{n}]\times\cdots\times[\frac{e_r-1}{n},\frac{e_r}{n}]\times[0,1]\times\cdots\times[0,1]$ for $e=\{e_1,\dots,e_r\}$. For any fixed $\mathbf{H}\in\mathcal{G}^{r,2k}_{q_0}$ basic martingale methods deliver

$$\mathbb{P}(|t(\mathbf{H},\mathbf{W_G}) - t(\mathbf{H},\mathbf{U})| \geq \delta) \leq 2\exp(-\frac{\delta^2 n}{2q_0^2})$$

for any $\delta > 0$, therefore when setting $\delta = \frac{\varepsilon}{4(2k)^{q_0^r}}$ we get that

$$d_{\mathrm{tw}}(\mu(q_0, \mathbf{U}), \mu(q_0, \mathbf{W_G})) = \frac{1}{2} \sum_{\mathbf{H} \in \mathcal{G}_{q_0}^{r,2k}} |t(\mathbf{H}, \mathbf{W_G}) - t(\mathbf{H}, \mathbf{U})| \leq \varepsilon/4$$

with probability at least $1 - \varepsilon/4$, since $n \geq q \geq d(r, \varepsilon/4, q_0, k, 2) = \frac{[q_0^r \ln(2k) - \ln(\varepsilon/4)][32(2k)^{q_0^r} q_0^2]}{\varepsilon^2}$. Summing up terms gives

$$\begin{aligned} d_{\mathrm{tw}}(\mu(q_0,\mathbf{F}), \mu(q_0, \mathbf{G})) \leq d_{\mathrm{tw}}(\mu(q_0, \mathbf{F}), \mu(q_0, \mathbf{W_F})) + d_{\mathrm{tw}}(\mu(q_0, \mathbf{W_F}), \mu(q_0, \mathbf{U})) \\ + d_{\mathrm{tw}}(\mu(q_0, \mathbf{U}), \mu(q_0, \mathbf{W_G})) + d_{\mathrm{tw}}(\mu(q_0, \mathbf{W_G}), \mu(q_0, \mathbf{G})) \leq \varepsilon, \end{aligned}$$

with failure probability at most $\varepsilon/2$, this concludes our proof.

□

6 Nondeterministically testable hypergraph properties

The concept of nondeterministic testing was originally introduced for testing properties by Lovász and Vesztergombi [14], and remarkable progress has been made in that context, see [8] and [14], the estimation of parameters, which is our main issue in this paper, is in close relationship to that concept. For related developments in combinatorial property testing using regularity methods we refer to [2].

We present the definition of testability of properties in the usual and in the nondeterministic sense and construct a tester from the tester of the witness property with the aid of Lemma 5.1 that achieves the same sample complexity as in the parameter testing case. This result connects our contribution to previous efforts more directly and answers the question posed in [14] asking if the equivalence of the two testability notions persists for uniform hypergraphs of higher order similar to the case of graphs.

Definition 6.1. *An r-graph property $\mathcal{P}$ is testable, if there exists another r-graph property $\hat{\mathcal{P}}$ called the sample property, such that*

(a) $\mathbb{P}(\mathbb{G}(q, G)) \in \hat{\mathcal{P}}) \geq \frac{2}{3}$ *for every $G \in \mathcal{P}$ and $q \geq 1$, and*

(b) for every $\varepsilon > 0$ there is an integer $q_{\mathcal{P}}(\varepsilon) \geq 1$ such that for every $q \geq q_{\mathcal{P}}(\varepsilon)$ and every G that is ε-far from $\mathcal{P}$ we have that $\mathbb{P}(\mathbb{G}(q, G)) \in \hat{\mathcal{P}}) \leq \frac{1}{3}$.

Testability for colored r-graphs is defined analogously.

We remark that ε-far here means that one has to modify at least $\varepsilon |V(G)|^2$ edges in order to obtain an element of $\mathcal{P}$. Note that $\frac{2}{3}$ and $\frac{1}{3}$ in the definition can be replaced by arbitrary constants $1 > a > b > 0$, this change may alter the corresponding certificate $\hat{\mathcal{P}}$ and the function $q_{\mathcal{P}}$, but not the characteristic of $\mathcal{P}$ being testable or not. Let $\mathcal{P}_n$ denote the elements of $\mathcal{P}$ of size n.

Next we formulate the definition of nondeterministic testability.

Definition 6.2. *An r-graph property $\mathcal{P}$ is nondeterministically testable, if there exists an integer $k \geq 1$ and a $2k$-colored r-graph property $\mathcal{Q}$ called the witness property that is testable in the sense of Definition 6.1 satisfying $[\mathcal{Q}, k] = \{[\mathbf{G}, k] | \mathbf{G} \in \mathcal{Q}\} = \mathcal{P}$ (see Definition 2.2 above for the discoloring operation).*

We formulate next the main theorem in this section.

Theorem 6.3. *Every nondeterministically testable r-graph property is testable.*

Proof. Let $\mathcal{P}$ be a nondeterministically testable property with witness property $\mathcal{Q}$ of $2k$-colored r-graphs. We employ the combinatorial language with counting subgraph densities when referring to $\mathcal{Q}$ and its testability, and the probabilistic language of picking random subgraphs in a uniform way when handling $\mathcal{P}$ in order to facilitate readability.

Let $\hat{\mathcal{Q}}$ be the corresponding sample property that certifies the testability of $\mathcal{Q}$ and $q_{\mathcal{Q}}$ be the sample complexity corresponding to the thresholds $1/5$ and $4/5$, that is

(i) if $\mathbf{G} \in \mathcal{Q}$, then for every and $q \geq 1$ we have $t(\hat{\mathcal{Q}}_q, \mathbf{G}) \geq 4/5$, and

(ii) for every $\varepsilon > 0$ if $\mathbf{G}$ is ε-far from $\mathcal{Q}$, then for every $q \geq q_{\mathcal{Q}}(\varepsilon)$ we have that $t(\hat{\mathcal{Q}}_q, \mathbf{G}) \leq 1/5$.

Our task is to construct a property $\hat{\mathcal{P}}$ together with a function $q_{\mathcal{P}}$ such that they fulfill the conditions of Definition 6.1. We are free to specify the error thresholds by the remark after Definition 6.1, we set them to $2/5$ and $3/5$.

Let n be a positive integer and let $\varepsilon_n > 0$ be the infimum of all positive reals δ that satisfy $n \geq \max\{q_{\mathrm{tw}}(r, 1/10, q_Q(\delta), 3, k); 100q_{\mathcal{Q}}^2(\delta); d(r, 1/10, q_{\mathcal{Q}}(\delta), k, 2)\}$ from Lemma 5.1. Define for each n the set

$$\hat{\mathcal{P}}_n = \{H \in \mathcal{G}_n^r | \text{there exists a } k\text{-coloring } \mathbf{H} \text{ of } H \text{ such that } t(\hat{\mathcal{Q}}_{q_{\mathcal{Q}}(\varepsilon_n)}, \mathbf{H}) \geq 3/5\},$$

and let $\hat{\mathcal{P}} = \cup_{n=1}^{\infty} \hat{\mathcal{P}}_n$. We set $q_{\mathcal{P}}(\varepsilon) = \max\{q_{\mathrm{tw}}(r, 1/10, q_Q(\varepsilon), 2, k); 100q_{\mathcal{Q}}^2(\varepsilon); d(r, 1/10, q_{\mathcal{Q}}(\delta), k, 2)\}$. We are left to check if the two conditions for testability of $\mathcal{P}$ hold with $\hat{\mathcal{P}}$ and $q_{\mathcal{P}}$ described as above. Assume for the rest of the proof that $n \geq q_{\mathcal{P}}(\varepsilon_n)$ for each n for simplicity, the general case follows along the same lines with some technical difficulties.

First let $G \in \mathcal{P}$, we have to show that for every $q \geq 1$ integer we have that $\mathbb{G}(q, G) \in \hat{\mathcal{P}}_q$ with probability at least $3/5$.

The condition $G \in \mathcal{P}$ implies that there exists a a k-coloring $\mathbf{G}$ of G such that $\mathbf{G} \in \mathcal{Q}$. From the testability of $\mathcal{Q}$ it follows that $t(\hat{\mathcal{Q}}_l, \mathbb{G}(l, \mathbf{G})) \geq 4/5$ for any $l \geq 1$. Let $q \geq 1$ be arbitrary, and let F denote $\mathbb{G}(q, G)$, furthermore let $\mathbf{F} = \mathbb{G}(q, \mathbf{G})$ generated by the same random process as F, so $\mathbf{F}$ is a k-coloring of F. We know by a standard sampling argument that

$$\mathbb{P}(|t(\hat{\mathcal{Q}}_{q_{\mathcal{Q}}(\varepsilon_q)}, \mathbf{G}) - t(\hat{\mathcal{Q}}_{q_{\mathcal{Q}}(\varepsilon_q)}, \mathbf{F})| \geq 1/5) \leq 2\exp\left(-\frac{q}{50q_{\mathcal{Q}}^2(\varepsilon_q)}\right), \tag{6.1}$$

and the right hand side of (6.1) is less than 2/5 by the choice of ε_q, since by definition $q \geq 100q_Q^2(\varepsilon_q)$. It follows that $t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_q)}, \mathbf{F}) \geq 3/5$ with probability at least 3/5, so by the definition of $\hat{P}$ we have that $F \in \hat{\mathcal{P}}_q$ with probability at least 3/5, which is what we wanted to show.

To verify the second condition we proceed by contradiction. Suppose that G is ε-far from $\mathcal{P}$, but at the same time there exists an $l \geq q_{\mathcal{P}}(\varepsilon)$ such that $F \in \hat{\mathcal{P}}_l$ with probability larger than 2/5, where $F = \mathbb{G}(l, G)$.

In this case, the latter condition implies that with probability larger than 2/5 there exists a k-coloring $\mathbf{F}$ of F such that $t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_l)}, \mathbf{F}) \geq 3/5$. By Lemma 5.1 and the proof of Theorem 2.3 there exists a k-coloring $\mathbf{G}$ of G such that $d_{\mathrm{tw}}(\mu(q_Q(\varepsilon_l), \mathbf{F}), \mu(q_Q(\varepsilon_l), \mathbf{G})) \leq 22/100$ with probability at least 4/5, in particular

$$|t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_l)}, \mathbf{F}) - t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_l)}, \mathbf{G})| \leq \frac{22}{100},$$

which implies that with probability at least 1/5 there exist a $\mathbf{G}$ such that $t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_l)}, \mathbf{G}) > \frac{3}{10}$. We can drop the probabilistic assertion and can say that there exists a k-coloring $\mathbf{G}$ of G such that $t(\hat{\mathbf{Q}}_{q_Q(\varepsilon_l)}, \mathbf{G}) > \frac{3}{10}$, because G and the density expression are deterministic.

On the other hand, the fact that G is ε-far from $\mathcal{P}$ implies that any k-coloring $\mathbf{G}$ of G is ε-far from $\mathcal{Q}$, which means that $t(\hat{\mathbf{Q}}_q, \mathbf{G}) \leq 1/5$ for any k-coloring $\mathbf{G}$ of G and $q \geq q_Q(\varepsilon)$. But we know that $q_Q(\varepsilon_l) \geq q_Q(\varepsilon)$, since $\varepsilon_l \leq \varepsilon$ which delivers the contradiction. The last inequality is the consequence of our definitions, ε_l is the infimum of the $\delta > 0$ that satisfy $l \geq q_{\mathcal{P}}(\delta)$, and on the other hand, $l \geq q_{\mathcal{P}}(\varepsilon)$.

□

7 Further research

It would be very interesting to shed light on the explicit sample complexity bounds for the witness parameter in Theorem 2.3. The only ingredient of our proof which is non-effective is the part which deals with the ultralimit method in the proof of Theorem 3.15, to our knowledge an effective proof regarding this result is only known for $r = 2$.

Acknowledgement

We thank Gábor Elek for a number of interesting discussions.

References

[1] Noga Alon, W. Fernandez de la Vega, Ravi Kannan, and Marek Karpinski. Random sampling and approximation of MAX-CSP problems. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 232–239, 2002. Also appeared in J. Comput. System Sci., 67(2):212–243,2003.

[2] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. *SIAM J. Comput.*, 39(1):143–167, 2009.

[3] Sanjeev Arora, David R. Karger, and Marek Karpinski. Polynomial time approximation schemes for dense instances of *NP*-hard problems. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, pages 284–293, 1995. Also appeared in J. Comput. System Sci., 58(1):193–210, 1999.

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

[7] Alan M. Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[8] Lior Gishboliner and Asaf Shapira. Deterministic vs non-deterministic graph property testing. *Israel J. Math.*, 204(1):397–416, 2014.

[9] Marek Karpinski and Roland Markó. Limits of CSP problems and efficient parameter testing, 2014. preprint, arXiv:1406.3514.

[10] Marek Karpinski and Roland Markó. Complexity of nondeterministic graph parameter testing, 2014. preprint, arXiv:1408.3590.

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

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

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

[14] László Lovász and Katalin Vesztergombi. Non-deterministic graph property testing. *Combin. Probab. Comput.*, 22(5):749–762, 2013.

[15] Yufei Zhao. Hypergraph limits: a regularity approach, 2013. preprint, arXiv:1302.1634.