Complexity of Nondeterministic Graph Parameter Testing

Marek Karpinski* Roland Markó†

Abstract

We study the sample complexity of nondeterministically testable graph parameters and improve existing bounds by several orders of magnitude. The technique used would be also of independent interest. We also discuss some generalization and the special case of nondeterministic testing with polynomial sample size.

1 Introduction

We call a non-negative function on the set of labeled simple graphs a *graph parameter* if it is invariant under graph isomorphism. We define parameters of edge-k-colored directed graphs, that will be considered in this paper as loop-free, and graphons analogously. From now on colored means edge-colored if not noted otherwise, furthermore, each directed edge carries exactly one color. The central characteristic of parameters investigated in the current paper is the possibility of value estimation via uniform sampling. For a graph G (directed and k-colored possibly) the expression $\mathbb{G}(k, G)$ denotes the random induced subgraph of G with the vertex set chosen uniformly among all subsets of $V(G)$ that have cardinality k.

Definition 1.1. *The graph parameter f is testable if for any $\epsilon > 0$ there exists a positive integer $q_f(\epsilon)$ such that for any simple graph G with at least $q_f(\epsilon)$ nodes*

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

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

An a priori weaker notion than testability is the second cornerstone of the current work, it was introduced in [14].

*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

Definition 1.2. *The graph parameter f is non-deterministically testable if there exist integers $m \geq k$ and a testable edge-k-colored directed graph parameter g called witness such that for any simple graph G the value $f(G) = \max_{\mathbf{G}'=G} g(\mathbf{G})$ where the maximum goes over the set $\mathcal{K}$ of (k,m)-colorings of G. The edge-k-colored directed graph $\mathbf{G}$ is a (k,m)-coloring of G, if we erase all edges of $\mathbf{G}$ colored with an element of $[m+1,\dots,k]$ and forget about the orientation and coloring of the remaining edges, then we end up with G (G is the shadow of $\mathbf{G}$).*

The problem regarding the relationship of the set of parameters that are testable and those who are non-deterministically testable was first studied in the framework of graph limits by Lovász and Vesztergombi [14] in the spirit of the general "P vs. NP" question, that is a central problem in theoretical computer science. In the dense setting with this particular notion of nondeterministicity they were able to prove that any non-deterministically testable graph property is also testable, which implies the analogous statement for parameters.

Theorem 1.3. *[14] Every non-deterministically testable graph parameter f is testable.*

However, no explicit relationship between the sample size required by f and the two factors, the number of colors k and the sample complexity of the witness g was provided. The reason for the absence of this is that the authors exploited various consequences of the next remarkable fact.

Fact. If $(W_n)_{n\geq 1}$ and $\|W_n\|_\square \to 0$ when n tends to infinity, the for any measurable function $Z\colon [0,1]^2 \to [-1,1]$ it is true that $\|W_n Z\|_\square \to 0$, where the product is taken pointwise.

The norm $\|.\|_\square$ above will be precisely defined in the paper, for now it is enough to know that it is weaker than the L^1-norm, and combined with an optimal labeling procedure it is possible to define with its aid a distance whose unit ball is compact. Although the above statement is true for all Z, the convergence is not uniform and its rate depends heavily on the structure of Z.

The relationship of the magnitude of the sample complexity of a testable property $\mathcal{P}$ and its witness $\mathcal{Q}$ was analyzed by Gishboliner and Shapira [9] relying on Szemerédi's regularity lemma and its connections to graph property testing unveiled by Alon, Fischer, Newman, and Shapira [1]. The height of the exponential tower in the estimate of [9] was not bounded and growing in function of $\frac{1}{\epsilon}$, the main result can be rephrased as follows:.

Theorem 1.4. *[9] Every non-deterministically testable graph parameter f is testable. If the sample complexity of the witness parameter g for each $\epsilon > 0$ is $q_g(\epsilon)$, then the sample complexity of f for each $\epsilon > 0$ is at most $\mathrm{tf}((q_g(\epsilon/2)))$, where $\mathrm{tf}(t)$ is the value of the t-fold iteration of the function 2^x at 2.*

In the current note, motivated by the fact that most of the dense graph limit theory does only rely on the Weak Regularity Lemma as a central tool, see [4], [5], we improve on the result of [9] using a weaker kind of regularity approach which eliminates the tower-type dependence on the error parameter ϵ.

Theorem 1.5. *Let f be a testable graph parameter with non-deterministic witness parameter g of k-colored graphs, and let the corresponding sample complexities be q_f and q_g. Then there for any $\epsilon > 0$ we have $q_f(\epsilon) \leq \exp^{(3)}(\mathrm{poly}(q_g(\epsilon/2)))$.*

1.1 Outline of the paper

This paper is organized as follows. In Section 2 we introduce the basic notation related to dense graph limit theory that is necessary to conduct the proof of the main result in Theorem 1.5, and we will also state and prove the main ingredient of the proof, the intermediate regularity lemma, that might be of interest on its own right. Section 3 continues with the proof of Theorem 1.5, while in Section 4 we treat some generalizations and special cases of the non-deterministic testing notion applied in the current paper, and also directions of further research are discussed.

2 Graph limits and regularity lemmas

First we provide the definition of graph convergence via subgraph densities. For the simple graphs F and G let $\hom(F,G)$ denote the number of maps $\phi \colon V(F) \to V(G)$ that preserves adjacency, that is, it is a graph homomorphism. Furthermore, let $t(F,G) = \frac{\hom(F,G)}{|V(G)|^{|V(F)|}}$ denote the subgraph density. The densities $t(\mathbf{F},\mathbf{G})$ in the case of k-colored digraphs is defined similarly. The variant $t_{\mathrm{inj}}(.,.)$ stands for the relative cardinality of injective graph homomorphisms.

Definition 2.1. *Let $(G_n)_{n\geq 1}$ be a sequence of simple graphs graphs. It is said to be convergent if for every simple F the numerical sequences $(t(F,G_n))_{n\geq 1}$ converge to some limit. Convergence is defined in the case of sequences of k-colored directed graphs similarly.*

Now we describe the space of possible limit objects of simple graphs. Let I be an interval and $\mathcal{W}_I$ be the set of all measurable functions $W \colon [0,1] \times [0,1] \to I$ that are symmetric in the sense that $W(x,y) = W(y,x)$ for all $x, y \in [0,1]$. When $I = [0,1]$, then we call $\mathcal{W}_I$ the space of graphons. The space of k-colored directed graphons can be described in a similar, though more complicated way. Let $\mathcal{W}^{(k)}$ be the set of k^2-tuples $\mathbf{W} = (W^{(i,j)})_{i,j\in[k]}$ such that for each $i,j \in [k]$ the function $W^{(i,j)} \colon [0,1] \times [0,1] \to [0,1]$ is measurable, they have a symmetry in the sense that $W^{(i,j)}(x,y) = W^{(j,i)}(y,x)$ for each $x,y \in [0,1]$, and also $\sum_{i,j\in[k]} W_{i,j}(x,y) = 1$ for each $x,y \in [0,1]$. Note that for each set $[m]$ with $m \leq k$ we have that $W'(x,y) = \sum_{i,j\in[m]} W^{(i,j)}(x,y)$ are graphons when $\mathbf{W}$ is k-colored directed graphon, furthermore, each graphon W can be regarded as a 2-colored directed graphon by setting $W^{(1,2)} = W^{(2,1)} = 0$ and $W^{(1,1)} = W$ everywhere.

We introduce the notion of the *canonical equiv-partition* of $[0,1]$ into t sets for a partition $\mathcal{P}$ whose classes are the intervals $P_i = [\frac{i-1}{t}, \frac{i}{t})$ for each $i \in [t]$. We can associate to each simple graph G on n vertices a graphon W_G that is a step function with the steps forming the canonical eguiv-partition into n sets and taking value 1 on $P_i \times P_j$ whenever $ij \in E(G)$ and 0 otherwise. Similarly, for a k-colored directed $\mathbf{G}$ we can define $\mathbf{W_G}$ as the step function with the same steps as above by setting $W_{\mathbf{G}}^{(\alpha,\beta)}$ to 1 on $P_i \times P_j$ when (i,j) is colored by α and (j,i) by β in $\mathbf{G}$, and to 0 otherwise.

For a positive integer n we call a partition of $[0,1]$ and n-partition if it is refined by the canonical partition into n sets, and we call a function on $[0,1]^2$ an n-step function if its steps form an n-partition.

Next we define the sampling process for these objects, this will be right away explained for the general case.

Definition 2.2. *Let G be a simple graph and S be a random subset of $V(G)$ chosen among all subsets of cardinality q uniformly. Then $\mathbb{G}(q,G)$ denotes the random induced subgraph of G on S. For a k-colored directed graph $\mathbf{G}$ the random subgraph $\mathbb{G}(q,\mathbf{G})$ is defined identically.*

Let W be a graphon and $q \geq 1$, furthermore, $(X_i)_{i\in[q]}$ and $(Y_{ij})_{ij\in\binom{[q]}{2}}$ mutually pair-wise independent uniform $[0,1]$ random variables. Then the random graph $\mathbb{G}(q,W)$ has vertex set $[q]$ and an edge runs between ij if $Y_{ij} \geq W(X_i,X_j)$. The random k-colored directed graph $\mathbb{G}(q,\mathbf{W})$ has also vertex set $[q]$, conditioned on $(X_i)_{i\in[q]}$, the colors for the edges are chosen independently for all pairs $ij \in \binom{[q]}{2}$ of vertices, (i,j) has color α and (j,i) has color β with probability $W^{(\alpha,\beta)}(X_i,X_j)$.

Note that in $\mathbb{G}(q,\mathbf{W})$ the colors of (i,j) and (j,i) are not even conditionally independent.

The next theorem, first proven in [11], states that the graphons truly represent the limit space of graphs. For the proof of the general case, see [6], [12] or [10].

The density of a simple graph F with vertex set $[q]$ in a graphon W is defined as

$$t(F,W) = \int\limits_{[0,1]^q} \prod_{ij\in E(F)} W(x_i,x_j)\mathrm{d}x,$$

and the density of a colored digraph $\mathbf{F}$ with the same vertex set as above in $\mathbf{W}$ is given as

$$t(\mathbf{F},\mathbf{W}) = \int\limits_{[0,1]^q} \prod_{\substack{ij:E(\mathbf{F})(i,j)=\alpha \\ E(\mathbf{F})(j,i)=\beta}} W^{(\alpha,\beta)}(x_i,x_j)\mathrm{d}x.$$

Theorem 2.3. *[11], [12] If $(G_n)_{n\geq 1}$ is a convergent sequence of simple graphs, then there exists a graphon W such that for every simple graph F we have $t(F,G_n) \to t(F,W)$, when n tends to infinity. Similarly, If $(\mathbf{G}_n)_{n\geq 1}$ is a convergent sequence of k-colored directed graphs, then there exists a k-colored digraphon $\mathbf{W}$ such that for every $\mathbf{F}$ it holds that $t(\mathbf{F},\mathbf{G}_n) \to t(\mathbf{F},\mathbf{W})$.*

It remains to add the norms and distances that will be relevant and posses valuable properties with regard to graph limits.

Definition 2.4. *The cut norm of an $n \times n$ matrix A is*

$$\|A\|_\square = \frac{1}{n^2} \max_{S,T\subset[n]} |A(S,T)|.$$

The cut distance of two labeled simple graphs F and G on the same vertex set $[n]$ is

$$d_\square(F,G) = \|A_F - A_G\|_\square,$$

where A_F and A_G stand for the respective adjacency matrices. The cut norm of a graphon W is

$$\|W\|_\square = \max_{S,T\subset[0,1]} \left| \int\limits_{S\times T} W(x)\mathrm{d}x \right|,$$

where maximum is taken over all pairs of measurable sets S and T. We speak of the n-cut norm when the maximum is only taken over such sets that can be given as the union of some classes belonging to the canonical partition into n sets, it is denoted by $\|W\|_\square^{\langle n\rangle}$. The cut norm of a k-colored directed graphon is

$$\|\mathbf{W}\|_\square = \sum_{i,j=1}^{k} \|W^{(i,j)}\|_\square.$$

The cut distance of two graphons W and U is

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

where the infimum runs over all measure-preserving permutations of $[0,1]$, and the graphon W^ϕ is defined as $W^\phi(x,y) = W(\phi(x),\phi(y))$. Similarly for k-colored directed graphons $\mathbf{W}$ and $\mathbf{U}$ we have

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

with the difference being component-wise. The cut distance for arbitrary unlabeled graphs F and G is

$$\delta_\square(F,G) = \delta_\square(W_F, W_G),$$

the definitions for the colored directed version is identical. Another variant is for the case when $V(F) = [m]$ and $V(G) = [n]$ such that m is a divisor of n. Then

$$\delta_\square^{\langle n\rangle}(F,G) = \min_\phi d_\square(F[n/m], G^\phi),$$

where $F[t]$ is the t-fold blow up of F and minimum goes over all node relabelings of G.

Observe that for two graphs F and G on the common node set $[n]$ the distance $d_\square(F,G) = \|W_F - W_G\|_\square = \|W_F - W_G\|_\square^{\langle n\rangle}$.

The connection to graph limits is given in the next theorem from [4].

Theorem 2.5. *[4] The graph sequence $(G_n)_{n\geq 1}$, and the k-colored directed graph sequence $(\mathbf{G}_n)_{n\geq 1}$ respectively, is convergent if and only if it is Cauchy in the $\delta_\square$ metric.*

We list some variants of the Weak Regularity Lemma for graphons, and going to derive the intermediate version using the general lemma, a regularity lemma in Hilbert spaces, that immediately follows, that is our key tool in the proof of Theorem 1.5.

Lemma 2.6. *[13] Let $\mathcal{K}_1, \mathcal{K}_2, \dots$ be arbitrary subsets of a Hilbert space $\mathcal{H}$. Then for every $\epsilon > 0$ and $f \in \mathcal{H}$ there is an $m \leq \frac{1}{\epsilon^2}$ and there are $f_i \in \mathcal{K}_i$ and $\gamma_i \in \mathbb{R}$ $(1 \leq i \leq m)$ such that for every $g \in \mathcal{K}_{m+1}$ we have that*

$$|\langle g, f - \sum_{i=1}^{m} \gamma_i f_i \rangle| \leq \epsilon \|f\| \|g\|. \tag{2.1}$$

Actually, we require a version that also contains a lower bound on m and the condition that the f_i's are linearly independent, these were not present in the original formulation, the inclusion does not alter the proof of the original. In the case that for some f with $\|f\| \leq 1$ Lemma 2.6 outputs an m below our desired bound, then we pick an arbitrary $f_{m+1} \in \mathcal{K}_{m+1}$ and a γ_{m+1} such that $\|f - \sum_{i=1}^{m+1} \gamma_i f_i\| \leq 1$, and apply the lemma once again for $f - \sum_{i=1}^{m+1} \gamma_i f_i$. Iterate this procedure until the desired lower bound is reached eventually. We phrase this result as a corollary.

Corollary 2.7. *Let $\mathcal{K}_1, \mathcal{K}_2, \dots$ be arbitrary subsets of a Hilbert space $\mathcal{H}$. Then for every $\epsilon > 0$ and $f \in \mathcal{H}$ with $\|f\| \leq 1$ there is an $\log \frac{1}{\epsilon} \leq m \leq \log \frac{1}{\epsilon} \frac{1}{\epsilon^2}$ and there are linearly independent $f_i \in \mathcal{K}_i$ and $\gamma_i \in \mathbb{R}$ $(1 \leq i \leq m)$ such that for every $g \in \mathcal{K}_{m+1}$ we have that*

$$|\langle g, f - \sum_{i=1}^{m} \gamma_i f_i \rangle| \leq \epsilon \|f\| \|g\|. \tag{2.2}$$

One can easily deduce Frieze and Kannan's version from the above one, that found various application in the design of efficient algorithms.

Lemma 2.8 (Weak regularity lemma)**.** *[8], [13] For every $\epsilon > 0$ and $W \in \mathcal{W}_I$ there exists a partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0,1]$ into $m \leq 2^{\frac{8}{\epsilon^2}}$ parts, such that*

$$\|W - W_{\mathcal{P}}\|_{\square} \leq \epsilon, \tag{2.3}$$

where we get $W_{\mathcal{P}}$ from W by averaging on every rectangle given by products from $\mathcal{P}$.

In the same we get the version for k-colored graphons.

Lemma 2.9 (Weak regularity lemma for k-colored directed graphons)**.** *For every $\epsilon > 0$ and k-colored digraphon there exists a partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0,1]$ into $m \leq 2^{k^2 \frac{8}{\epsilon^2}}$ equal parts, such that*

$$d_{\square}(\mathbf{W}, \mathbf{W}_{\mathcal{P}}) = \sum_{i,j=1}^{k} \|W^{(i,j)} - (W^{(i,j)})_{\mathcal{P}}\|_{\square} \leq \epsilon. \tag{2.4}$$

The following norm shares some useful properties with the cut-norm, most prominently it admits a regularity lemma with comparable number of steps to the weak one, although it does not admit a definition of a related distance by calculating the norm of the difference of two optimally overlayed objects. This comes from the general assumption that the partition $\mathcal{P}$ always will belong to one of the graphons whose deviation we wish to estimate and any "re-labeling" of $[0,1]$ should act on them simultaneously, therefore symmetry fails. Its advantages will get clearer in the next section.

Definition 2.10. *Let W be a graphon and $\mathcal{P} = (P_1, \dots, P_t)$ a partition of $[0,1]$. Then the cut-$\mathcal{P}$-norm of W is*

$$\|W\|_{\square\mathcal{P}} = \max_{S_i, T_i \subset P_i} \sum_{i,j=1}^{t} \left| \int_{S_i \times T_j} W(x,y)\mathrm{d}x\mathrm{d}y \right|. \tag{2.5}$$

For two graphons U and W let $d_{W,\mathcal{P}}(U)$ denote the cut-$\mathcal{P}$-entropy of U with respect to W that is defined by

$$d_{W,\mathcal{P}}(U) = \inf_{\phi} \|U^{\phi} - W\|_{\square\mathcal{P}}, \tag{2.6}$$

where the infimum runs over all measure preserving maps from $[0,1]$ to $[0,1]$.

For $n \geq 1$, a partition $\mathcal{P}$ of $[n]$ and a directed weighted graph H the cut-$\mathcal{P}$-norm of H is defined as

$$\|H\|_{\square\mathcal{P}} = \|W_H\|_{\square\mathcal{P}'}, \tag{2.7}$$

where $\mathcal{P}'$ is the partition of $[0,1]$ induced by $\mathcal{P}$ and the map $j \mapsto [\frac{j-1}{n}, \frac{j}{n})$.

The definition for the k-colored version is analogous.

Definition 2.11. *Let $\mathbf{W} = (W^{(1,1)}, \dots, W^{(k,k)})$ be a k-colored digraphon and $\mathcal{P} = (P_1, \dots, P_t)$ a partition of $[0,1]$. Then the cut-$\mathcal{P}$-norm of $\mathbf{W}$ is*

$$\|\mathbf{W}\|_{\square\mathcal{P}} = \sum_{i,j=1}^{k} \|W^{(i,j)}\|_{\square\mathcal{P}}. \tag{2.8}$$

For two k-colored digraphons $\mathbf{U}$ and $\mathbf{W}$ let $d_{\mathbf{W},\mathcal{P}}(\mathbf{U})$ denote the cut-$\mathcal{P}$-entropy of $\mathbf{U}$ with respect to $\mathbf{W}$ that is defined by

$$d_{\mathbf{W},\mathcal{P}}(\mathbf{U}) = \inf_{\phi} \|\mathbf{U}^{\phi} - \mathbf{W}\|_{\square\mathcal{P}} = \inf_{\phi} \sum_{i,j=1}^{k} \|(U^{(i,j)})^{\phi} - W^{(i,j)}\|_{\square\mathcal{P}}, \tag{2.9}$$

where the infimum runs over all measure preserving maps from $[0,1]$ to $[0,1]$.

It is not hard to check that the cut-$\mathcal{P}$-norm is truly a norm on the space where we identify two graphons when they differ only on a set of measure 0. From the definition it follows directly that for any U, W graphons and any partition $\mathcal{P}$ we have $\|W\|_\Box \leq \|W\|_{\Box\mathcal{P}} \leq \|W\|_1$ and $\delta_\Box(U, W) \leq d_{W,\mathcal{P}}(U) \leq \delta_1(U, W)$, the same is true for the k-colored version.

Remark 2.12. We present a different representation of the cut-$\mathcal{P}$-norm of W and $\mathbf{W}$ respectively that will allow us to rely on result concerning the cut-norm more directly. Let $\{a_{\alpha,\beta}\}_{\alpha,\beta=1}^t \in \{-1, +1\}^{t\times t}$ and let $W^a(x, y) = a_{\alpha,\beta}W(x, y)$, respectively $\mathbf{W}^a$ given by $(W^{(i,j)})^a(x, y) = a_{\alpha,\beta}W^{(i,j)}(x, y)$ for $x \in P_\alpha$ and $y \in P_\beta$. Then $\|W\|_{\Box\mathcal{P}} = \max_a \|W^a\|_\Box$ and $\|\mathbf{W}\|_{\Box\mathcal{P}} = \max_a \|\mathbf{W}^a\|_\Box$.

This newly introduced norm admits a uniform approximation assertion essential to the proof of Theorem 1.5 in the following sense.

Lemma 2.13 (Intermediate regularity lemma for edge k-colored graphs)**.** *Let $n \geq 1$ fixed, and let each partition in the statement be such that it is refined by the canonical partition of $[0, 1]$ into n parts, and each function be such that it is constant on the product sets of classes of the canonical partition.*

For every $\epsilon > 0$, $k \geq 1$ and k-colored directed graphon $\mathbf{W}$ there exists a partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0, 1]$ into $m \leq 2^{(2k^2+1)^{16/\epsilon^2}}$ parts and a step function $\mathbf{V}$ with steps from $\mathcal{P}$, such that for any partition $\mathcal{Q}$ of $[0, 1]$ into at most m classes we have

$$\|\mathbf{W} - \mathbf{V}\|_{\Box\mathcal{Q}} \leq \epsilon/2. \tag{2.10}$$

Furthermore it holds that

$$\|\mathbf{W} - \mathbf{W}_\mathcal{P}\|_{\Box\mathcal{P}} \leq \epsilon. \tag{2.11}$$

If we require the parts to have almost equal measure, then the upper bound on the number of classes is modified into $t_k(\epsilon) = 2^{(4k^2+2)^{64/\epsilon^2}}$.

Proof. We will use the result of Lemma 2.6 with a suitable choice of the space $\mathcal{H}$ and the sets $\mathcal{K}_i$. Let $\mathcal{H}$ be $\mathcal{W}^{(k)}$ with the sum of the component-wise L^2-products as the inner product and $\mathcal{K}_i$ be the set of k^2-tuples of indicator functions that have the following form. Set $s(1) = 1$ and $s(i + 1) = s(i)(s(i) + 1)^{2k^2}$ for each $i \geq 1$. Let $(S_i^{(j,l)})_{i\in[m],j,l\in[k]}$ and $(T_i^{(j,l)})_{i\in[m],j,l\in[k]}$ be such that for each $j, l \in [k]$ the tuples $(S_1^{(j,l)}, \dots, S_{s(i)}^{(j,l)})$ and $(T_1^{(j,l)}, \dots, T_{s(i)}^{(j,l)})$ consist of pairwise disjoint measurable subsets of $[0, 1]$ and let $C^{(j,l)} \subset [s(i)]^2$, and define $\mathcal{K}_i$ as the set that consists of k^2-tuples of the signed indicator functions whose components can be expressed in the form $\pm[\sum_{(\alpha,\beta)\in C^{(j,l)}} \mathbb{I}_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}}]$ for some choice of the previous sets.

Let us fix $\epsilon > 0$. Applying Lemma 2.6 with $\epsilon/4$ ensures the existence of an integer m satisfying $\log \frac{4}{\epsilon} \leq m \leq \frac{16}{\epsilon^2}$ and $\mathbf{W}_i \in \mathcal{K}_i$, $\gamma_i \in \mathbb{R}$ such that for any $\mathbf{U} = (U^{(j,l)})_{j,l\in[k]} \in \mathcal{K}_{m+1}$ of the form $\mathbf{U}^{(j,l)} = \sum_{(\alpha,\beta)\in C^{(j,l)}} \mathbb{I}_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}}$ we have

$$\sum_{j,l\in[k]} \left| \sum_{(\alpha,\beta)\in C^{(j,l)}} \int_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}} W^{(j,l)}(x, y) - \sum_{i=1}^m \gamma_i W_i^{(j,l)}(x, y)\mathrm{d}x\mathrm{d}y \right| \leq \epsilon/4. \tag{2.12}$$

Let us denote the sum $\sum_{i=1}^{m}\gamma_i W_i^{(j,l)}$ by $V^{(j,l)}$. From the definition of the sets $\mathcal{K}_i$ it follows that each $V^{(j,l)}$ is a step function with at most $t=\prod_{i=1}^{m}(s(i)+1)^{2k^2}$ common steps, let us denote them by $P_1,\dots,P_t$. It is easy to verify that $t=s(m+1)$, so in particular we have for any $S_\alpha^{(j,l)},T_\alpha^{(j,l)}\subset P_\alpha$, with $C^{(j,l)\prime}=\{(\alpha,\beta):\int_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}}W^{(j,l)}-V^{(j,l)}\geq 0\}\subset[s(m+1)]^2$ and $C^{(j,l)\prime\prime}=[s(m+1)]^2\setminus C^{(j,l)\prime}$ that

$$\sum_{j,l\in[k]}\sum_{(\alpha,\beta)\in C^{(j,l)\prime}}\left|\int\limits_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}}W^{(j,l)}-V^{(j,l)}\right|\leq\epsilon/4,\tag{2.13}$$

and

$$\sum_{j,l\in[k]}\sum_{(\alpha,\beta)\in C^{(j,l)\prime\prime}}\left|\int\limits_{S_\alpha^{(j,l)}\times T_\beta^{(j,l)}}W^{(j,l)}-V^{(j,l)}\right|\leq\epsilon/4.\tag{2.14}$$

Therefore

$$\|\mathbf{W}-\mathbf{V}\|_{\square\mathcal{P}}=\sum_{j,l\in[k]}\max_{S_i^{j,l\in[k]},T_i^{j,l\in[k]}\subset P_i}\sum_{\alpha,\beta=1}^{m}\left|\int\limits_{S_\alpha^{j,l\in[k]}\times T_\beta^{j,l\in[k]}}W^{(j,l)}-V^{(j,l)}\right|\leq\epsilon/2.\tag{2.15}$$

Claim 2.14. The cut-$\mathcal{P}$-norm is contractive with respect to averaging. That is for any $W\in\mathcal{W}^{(k)}$ any $\mathcal{Q}$ that is a refinement of $\mathcal{P}$ we have $\|W_{\mathcal{Q}}\|_{\square\mathcal{P}}\leq\|W\|_{\square\mathcal{P}}$.

Applying Claim 2.14 it follows that

$$\begin{aligned}\|\mathbf{W}-\mathbf{W}_{\mathcal{P}}\|_{\square\mathcal{P}}&\leq\|\mathbf{W}-\mathbf{V}\|_{\square\mathcal{P}}+\|\mathbf{V}-\mathbf{W}_{\mathcal{P}}\|_{\square\mathcal{P}}&(2.16)\\&=\|\mathbf{W}-\mathbf{V}\|_{\square\mathcal{P}}+\|(\mathbf{V}-\mathbf{W})_{\mathcal{P}}\|_{\square\mathcal{P}}\leq 2\|\mathbf{W}-\mathbf{V}\|_{\square\mathcal{P}}\leq\epsilon.&(2.17)\end{aligned}$$

We are left to construct an upper bound $t=s(m+1)$. Therefore define $r(1)=1$ and $r(i+1)=2^{2k^2}r^{2k^2+1}(i)$ for $i\geq 1$, and let $l(i)=\log_2 r(i)$. It is clear now that $r(i)\geq s(i)$ for any i, and $l(1)=0$ and $l(i+1)=(2k^2+1)l(i)+2k^2$ for $i\geq 1$. Simple analysis shows that $l(i)=(2k^2+1)^{i-1}-1$, which eventually leads to $2^{(2k^2+1)^{\log\frac{4}{\epsilon}}}\leq t\leq 2^{3^m}\leq 2^{(2k^2+1)^{16/\epsilon^2}}$.

In order to verify the statement regarding the equiv-partition case consider the modification of the above construction by setting $s'(1)=1$ and $s'(i+1)=[s'(i)(s'(i)+1)^{2k^2}][s'(i)(s'(i)+1)^{2k^2}+1]$, and applying Lemma 2.6 with $\epsilon/4$. Identical analysis as before delivers that $s'(i)\leq 2^{(4k^2+2)^{i-1}-1}$. Using the above notation, V is a step function with $t=s'(i)(s'(i)+1)^{2k^2}$ steps denoted by $\mathcal{P}$. Let the refinement $\mathcal{P}'$ of $\mathcal{P}$ be such that each P_i is subdivided in an arbitrary way into the sets $P_{i,1},\dots,P_{i,h_i}$ of size $1/t^2$ and a remainder set $P_{i,0}$. Subsequently consider $\mathcal{P}''$ that we obtain from $\mathcal{P}'$ by replacing the remainder sets by an arbitrary $1/t^2$ subdivision of their union to eventually obtain an equiv-partition with $s'(i+1)$ classes.

Claim 2.15. Let $\mathcal{P}$ and $\mathcal{Q}$ two partitions of $[0,1]$ with the same number of classes t (this will be not necessary, if this fails, then add empty sets to one of the partitions). Also, let $W \in \mathcal{W}_{[-1,1]}$ be such that it is 0 on the set $[\cup_{i=1}^t (P_i \cap Q_i)] \times [\cup_{i=1}^t (P_i \cap Q_i)]$. Then

$$||W_{\mathcal{P}}\|_{\square\mathcal{P}} - \|W_{\mathcal{Q}}\|_{\square\mathcal{Q}}| \leq 4 \sum_{i=1}^t \lambda(P_i \triangle Q_i).$$

We conclude with

$$\begin{aligned} \|W - W_{\mathcal{P}''}\|_{\square\mathcal{P}''} &\leq \|W - W_{\mathcal{P}'}\|_{\square\mathcal{P}'} + 4/t \\ &\leq 2\|W - V\|_{\square\mathcal{P}'} + 4/t \\ &\leq \epsilon/4 + 4/t \leq \epsilon, \end{aligned}$$

where the first inequality holds due to Claim 2.15, the second due Claim 2.14, and the third stands as a consequence of how the sets $\mathcal{K}_i$ were specified. □

As seen in the proof, the upper bound on the number of classes in the statement of the lemma is not the sharpest we can prove, we stay with the simpler bound for the sake of readability. In the simple graph case the above reads as follows.

Corollary 2.16 (Intermediate regularity lemma)**.** *For every $\epsilon > 0$ and $W \in \mathcal{W}_{[0,1]}$ there exists a partition $\mathcal{P} = (P_1, \dots, P_m)$ of $[0,1]$ into $m \leq 2^{3^{16/\epsilon^2}}$ parts, such that*

$$\|W - W_{\mathcal{P}}\|_{\square\mathcal{P}} \leq \epsilon. \tag{2.18}$$

With the additional condition that the partition classes should have the same measure the above is true with $m \leq 2^{6^{64/\epsilon^2}}$.

The following result regarding the distance of a simple graph and its induced subgraph on a uniformly chosen vertex set is crucial for our purposes. Originally it was established to verify Theorem 2.5, the equivalence of the substructure and the metric convergence.

Lemma 2.17. *[4] Let $\epsilon > 0$ and let U be a graphon with $\|U\|_\infty \leq 1$. Then for $q \geq 2^{100/\epsilon^2}$ we have*

$$\mathbb{P}(\delta_\square(U, \mathbb{H}(q,U) \geq \epsilon)) \leq \exp(-4^{100/\epsilon^2} \frac{\epsilon^2}{50}). \tag{2.19}$$

3 Proof of Theorem 1.5

We will use the continuity of a testable graph parameter with respect to the cut norm and the connection of this property to the sample complexity of the parameter. We require two results, the first one quantifies the above continuity.

Lemma 3.1. *Let g be a testable k-colored digraph parameter with sample complexity at most q_g. Then for any $\epsilon > 0$ and graphs $\mathbf{G}, \mathbf{H}$ with $\delta_\square(\mathbf{G}, \mathbf{H}) \leq 2^{-2q^2(\epsilon/2)\log k}$ we have $|g(\mathbf{G}) - g(\mathbf{H})| \leq \epsilon$, whenever $q^2/\min\{|V(\mathbf{G})|, |V(\mathbf{H})|\}^{q-1} < \epsilon$.*

Proof. Let $\mathbf{G}$, $\mathbf{H}$, and $\epsilon > 0$ as in the statement, and let $q = q(\epsilon/2)$. Then we have

$$\begin{aligned}|g(\mathbf{G}) - g(\mathbf{H})| \le{}& |g(\mathbf{G}) - g(\mathbb{G}(q, \mathbf{G}))| + |g(\mathbb{G}(q, \mathbf{W_G})) - g(\mathbb{G}(q, \mathbf{G}))| \\ &+ |g(\mathbb{G}(q, \mathbf{W_G})) - g(\mathbb{G}(q, \mathbf{W_H}))| + |g(\mathbb{G}(q, \mathbf{H})) - g(\mathbb{G}(q, \mathbf{W_H}))| \\ &+ |g(\mathbf{H}) - g(\mathbb{G}(q, \mathbf{H}))|. \end{aligned} \tag{3.1}$$

The first and the last term on the right of 3.1 can be upper bounded by $\epsilon/4$ with failure probability $\epsilon/2$, by the assumptions of the lemma. To deal with the second and the fourth term we require the fact that $\mathbb{G}(q, \mathbf{G})$ and $\mathbb{G}(q, \mathbf{W_G})$ has the same distribution conditioned on the X_i's for the second random object lie in different classes of the canonical equiv-partition of $[0, 1]$ into $|V(\mathbf{G})|$ classes. The failure probability of the latter event can be upper bounded by $q^2/2|V(\mathbf{G})|^{q-1}$.

In order to handle the third term we wish to upper bound the probability that the two random graphs are different in some appropriate coupling, since clearly in the event of equality the third term of (3.1) is 0. More precisely, we will show that $\mathbb{G}(q, \mathbf{W_G})$ and $\mathbb{G}(q, \mathbf{W_H})$ can be coupled in such a way that $\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) \neq \mathbb{G}(q, \mathbf{W_H})) < 1 - \epsilon$. We utilize that for a fixed k-colored digraph $\mathbf{F}$ on q vertices we can upper bound the deviation of the subgraph densities of $\mathbf{F}$ in $\mathbf{G}$ and $\mathbf{H}$ by the cut norm of their difference. In particular,

$$|\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) = \mathbf{F}) - \mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) = \mathbf{F})| \le \binom{q}{2} \delta_\square(\mathbf{W_G}, \mathbf{W_H}).$$

Therefore in our case

$$\sum_{\mathbf{F}} |\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) = \mathbf{F}) - \mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) = \mathbf{F})| \le k^{2\binom{q}{2}} \binom{q}{2} 2^{-2q^2 \log k} \le \epsilon, \tag{3.2}$$

where the sum goes over all labeled k-colored digraphs $\mathbf{F}$ on q vertices.

Now we can couple $\mathbb{G}(q, \mathbf{W_G})$ and $\mathbb{G}(q, \mathbf{W_H}))$ via the underlying independent uniform $[0, 1]$ random variables $\{X_i\}_{1 \le i \le q}$, and $\{Y_{i,j}\}_{1 \le i < j \le q}$, paying attention that the overlay satisfies $\mathbb{P}[\mathbb{G}(q, \mathbf{W_G})(ij) \neq \mathbb{G}(q, \mathbf{W_H})(ij) | X_i, X_j] = \sum_{\alpha,\beta=1}^k |W_{\mathbf{G}}^{(\alpha,\beta)}(X_i, X_j) - W_{\mathbf{H}}^{(\alpha,\beta)}(X_i, X_j)|$ for all $ij \in \binom{[q]}{2}$, so that in the end $\mathbb{P}[\mathbb{G}(q, \mathbf{W_G}) \neq \mathbb{G}(q, \mathbf{W_H})] \le \epsilon$. This implies that with positive probability (in fact, with at least $1 - 3\epsilon$) the sum of the five terms on the right hand side of (3.1) does not exceed ϵ, so the statement of the lemma follows. $\square$

We will also require the following statement which can be regarded as the quantitative counterpart of Lemma 3.2 from [14].

Lemma 3.2. *Let $\epsilon > 0$, U be a step function with steps $\mathcal{P} = (P_1, \dots, P_t)$ and V be a graphon with $\|U - V\|_{\square\mathcal{P}} \le \epsilon$, and also let $k \ge 1$. For any $\mathbf{U} = (U^{(1,1)}, \dots, U^{(k,k)})$ k-colored digraphon step function with steps from $\mathcal{P}$ that is an m-witness of U there exists a k-colored m-witness of V denoted by $\mathbf{V} = (V^{(1,1)}, \dots, V^{(k,k)})$ so that $\|\mathbf{U} - \mathbf{V}\|_\square = \sum_{\alpha,\beta=1}^k \|U^{(\alpha,\beta)} - V^{(\alpha,\beta)}\|_\square \le k^2\epsilon$.*

If $V = W_G$ for a simple graph G on n nodes and $\mathcal{P}$ is an n-partition of $[0, 1]$ then there is a (k, m)-coloring $\mathbf{G}$ of G that satisfies the above conditions and $\|\mathbf{U} - \mathbf{W_G}\|_\square \le 2k^2\epsilon$ whenever $n \ge 16/\epsilon^2$.

Proof. Let U, V and $\mathbf{U}$ as in the statement of the lemma. Then $\sum_{\alpha,\beta=1}^{k} U^{(\alpha,\beta)} = 1$, let M be the subset of $[k]^2$ such that its elements have at least one component at most m, and $\sum_{(\alpha,\beta)\in M} U^{(\alpha,\beta)} = U$. Now for $(\alpha,\beta) \in M$ set $V^{(\alpha,\beta)} = \frac{VU^{(\alpha,\beta)}}{U}$ on the set where $U > 0$ and $V^{(\alpha,\beta)} = \frac{V}{k^2-(k-m)^2}$ where $U = 0$, furthermore for $(\alpha,\beta) \notin M$ set $V^{(\alpha,\beta)} = \frac{(1-V)U^{(\alpha,\beta)}}{1-U}$ on the set where $1 > U$ and $V^{(\alpha,\beta)} = \frac{1-V}{(k-m)^2}$ where $U = 1$. We will show that the k-colored digraphon $\mathbf{V}$ defined this way satisfies the conditions, in particular for each $(\alpha,\beta) \in [k]^2$ we have $\|U^{(\alpha,\beta)} - V^{(\alpha,\beta)}\|_\square \le \epsilon$. We will only explicitly preform the calculation for $(\alpha,\beta) \in M$, the other case is analogous. We fix $S, T \subset [0,1]$.

$$
\begin{aligned}
\left| \int_{S\times T} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} \right| &= \left| \int_{S\times T, U>0} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} + \int_{S\times T, U=0} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} \right| \\
&\le \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j), U>0} \frac{U^{(\alpha,\beta)}}{U}(U-V) + \int_{(S\cap P_i)\times(T\cap P_j), U=0} \frac{1}{k^2-(k-m)^2}(U-V) \right| \\
&= \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j)} (U-V) \left[\mathbb{I}_{U>0} \frac{U^{(\alpha,\beta)}}{U} + \mathbb{I}_{U=0} \frac{1}{k^2-(k-m)^2} \right] \right| \\
&\le \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j)} (U-V) \right| \\
&= \|U - V\|_{\square\mathcal{P}} \le \epsilon.
\end{aligned}
$$

The second inequality is a consequence of $\left[\mathbb{I}_{U>0} \frac{U^{(\alpha,\beta)}}{U} + \mathbb{I}_{U=0} \frac{1}{k^2-(k-m)^2} \right]$ being a constant between 0 and 1 on each of the rectangles $P_i \times P_j$.

We prove now the second statement of the lemma concerning graphs with $V = W_G$ and a $\mathcal{P}$ that is an n-partition. The first part delivers the existence of $\mathbf{V}$ that is a (k,m)-coloring of W_G, which can be regarded as a fractional coloring of G, as $\mathbf{V}$ is constant on the sets associated with nodes of G. For $|V(G)| = n$ we get for each $ij \in \binom{[n]}{2}$ a probability distribution on $[k]^2$ with $\mathbb{P}[Z_{ij} = (\alpha,\beta)] = n^2 \int_{[\frac{i-1}{n},\frac{i}{n}]\times[\frac{j-1}{n},\frac{j}{n}]} V^{(\alpha,\beta)}(x,y)\mathrm{d}x\mathrm{d}y$. For each pair ij we make an independent random choice according to this measure, and color (i,j) by the first, and (j,i) by the second component of Z_{ij} to get a proper (k,m)-coloring $\mathbf{G}$ of G. We are left with the analysis of the deviation in the statement of the lemma, we will show that this is small with hight probability with respect to the randomization, which in turn implies existence. Now we have

$$
\|\mathbf{U} - \mathbf{W_G}\|_\square \le \|\mathbf{U} - \mathbf{V}\|_\square + \|\mathbf{V} - \mathbf{W_G}\|_\square
$$

$$\leq k^2\epsilon + \sum_{\alpha,\beta=1}^{k} \|V^{(\alpha,\beta)} - W_{\mathbf{G}}^{(\alpha,\beta)}\|_\square$$

For each $(\alpha, \beta) \in [k]^2$ we have that $\mathbb{P}\left(\|V^{(\alpha,\beta)} - W_{\mathbf{G}}^{(\alpha,\beta)}\|_\square \geq 4/\sqrt{n}\right) \leq 2^{-n}$, this result is exactly Lemma 4.3 in [4]. This implies for $n \geq 16/\epsilon^2$ the existence of a suitable coloring, which in turn finishes the proof of the lemma. □

Remark 3.3. Actually we can preform the same proof to verify the existence of a k-coloring $\mathbf{V}$ such that $d_{\square\mathcal{P}}(\mathbf{U}, \mathbf{V}) \leq k^2\epsilon$. On the other hand, we can not weaken the condition on the closeness of U and V, a small cut-norm of $U - V$ does not imply the existence of a suitable coloring $\mathbf{V}$, that is if the number of classes t is exponential in $1/\|U - V\|_\square$.

We proceed to the proof of the main statement of the paper. Before we do that we require yet another specific lemma.

Let $\mathcal{M}$ denote the set of U n-step functions that have steps $\mathcal{P}$ with $|\mathcal{P}| \leq t_k(\Delta/2)$ classes, and values between 0 and 1. In order to verify Theorem 1.5 we will condition on the event that is formulated in the following lemma.

Lemma 3.4. *Let G be a simple graph on n vertices and $\Delta > 0$. Then for $q \geq t_k^{14}(\Delta/2)$ we have*

$$|d_{U,\mathcal{P}}(G) - d_{U,\mathcal{P}}(\mathbb{G}(q, G))| \leq \Delta/2, \tag{3.3}$$

for each $U \in \mathcal{M}$ simultaneously, with probability at least $1 - \exp(-c\sqrt{q})$ for some constant $c > 0$ depending only on k.

Proof. Let q be such that it satisfies the conditions of the lemma and for technical convenience, n should be such that it is a multiple of q, and let us introduce the quantity $t = t_k(\Delta/2)$ and $F = \mathbb{G}(q, G)$. The proof consists of two main steps.

First we will show that for any n-partition $\mathcal{Q}$ of $[0, 1]$ into at most t classes there exists an n-permutation ϕ of $[0, 1]$ such that $\|W_G - W_F^\phi\|_{\square\mathcal{Q}} < \Delta/2$ with high probability simultaneously for each such $\mathcal{Q}$. Applying Lemma 2.16 with the error parameter $\Delta/2$ and lower bound t on the number of steps for the approximating step function for W_G we can assert that there exists an n-step function V with its steps forming the partition $\mathcal{P}$ into $t_\mathcal{P}$ steps with $t \leq t_\mathcal{P} \leq 2t$ such that for every $\mathcal{Q}$ partition into $t_\mathcal{Q}$ classes $t_\mathcal{Q} \leq t_\mathcal{P}$ it holds that

$$\|W_G - V\|_{\square\mathcal{Q}} \leq \Delta/2.$$

This property is equivalent to stating that

$$\max_{\mathcal{Q}} \max_{A\in\mathcal{A}} \max_{S,T\subset[0,1]} \sum_{i,j=1}^{t_\mathcal{P}} A_{i,j} \int_{S\times T} (W_G - V)(x, y)\mathbb{I}_{Q_i}(x)\mathbb{I}_{Q_j}(y)\mathrm{d}x\mathrm{d}y \leq \Delta/2. \tag{3.4}$$

We can reformulate the above by putting

$$J = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and defining the tensor product $B_A = A \otimes J$, so that $B_{i,j}^{\alpha,\beta} = A_{ij} J_{\alpha,\beta}$. The first matrix J correspond to the partition $(S \cap T, S \setminus T, T \setminus S, [0,1] \setminus (S \cup T)) = (T_1, T_2, T_3, T_4)$ generated by a pair (S, T) of measurable subsets of $[0,1]$ so that for any function W it holds that

$$\sum_{i,j=1}^{4} J_{ij} \int\limits_{[0,1]^2} W(x,y) \mathbb{I}_{T_i}(x) \mathbb{I}_{T_j} \mathrm{d}x\mathrm{d}y = \int\limits_{S \times T} W(x,y) \mathrm{d}x\mathrm{d}y.$$

It follows that the inequality (3.4) is equivalent to saying

$$\max_{A \in \mathcal{A}} \max_{\hat{\mathcal{Q}}} \sum_{i,j=1}^{t_{\mathcal{P}}} \sum_{\alpha,\beta=1}^{4} (B_A)_{i,j}^{\alpha,\beta} \int\limits_{[0,1]^2} (W_G - V)(x,y) \mathbb{I}_{Q_i^\alpha}(x) \mathbb{I}_{Q_j^\beta}(y) \mathrm{d}x\mathrm{d}y \leq \Delta/2, \tag{3.5}$$

where the second maximum goes over all n-partitions $\hat{\mathcal{Q}} = (Q_i^\alpha)_{\substack{i \in [t] \\ \alpha \in [4]}}$ into $4t$ classes. Let us substitute an arbitrary graphon W for $W_G - V$ in (3.5) and define

$$h_{A,\hat{\mathcal{Q}}}(U) = \sum_{\substack{1 \leq i,j \leq t_{\mathcal{P}} \\ 1 \leq \alpha,\beta \leq 4}} (B_A)_{i,j}^{\alpha,\beta} \int\limits_{[0,1]^2} U(x,y) \mathbb{I}_{Q_i^\alpha}(x) \mathbb{I}_{Q_j^\beta}(y) \mathrm{d}x\mathrm{d}y$$

and

$$h_A(U) = \max_{\hat{\mathcal{Q}}} h_{A,\hat{\mathcal{Q}}}(U)$$

as the expression whose optima is sought.

For notational convenience only lower indeces will be used when referring to the entries of B_A. We introduce a relaxed version of the above function $h_{A,\hat{\mathcal{Q}}}$ by replacing the requirement on $\hat{\mathcal{Q}}$ being an n-partition, instead we define

$$h_{A,f}(U) = \sum_{1 \leq i,j \leq 4t} (B_A)_{i,j} \int\limits_{[0,1]^2} U(x,y) f_i(x) f_j(y) \mathrm{d}x\mathrm{d}y$$

with $f = (f_i)_{i \in [4t]}$ being a fractional n-partition into $4t$ classes, that is, each component of f is a non-negative n-function, and their sum is the constant 1 function.

It is easy to see that

$$h_A(U) = \max_f h_{A,f}(U),$$

where f runs over all fractional n-partitions.

Denote $U' = W_{\mathbb{G}(q,U)}$. We wish to upper bound the probability that the deviation $|h_A(U) - h_A(U')|$ exceeds $\Delta/2$, for each $A \in \mathcal{A}$ simultaneously. To do be able to this we will require following tools, the proofs are not given here, we direct the reader to [2], [4], and [10] respectively for a reference. The technique we employ is at some places cited as the cut decomposition method.

The first result is a variant of the regularity lemma with an additional bound on the factors used to construct the approximating function.

Lemma 3.5. *[2] Let $\Delta > 0$ arbitrary, and $n \geq 1$. For any bounded measurable n-function $U\colon [0,1]^2 \to [0,1]$ there exist an $s \leq \frac{1}{\Delta^2}$, measurable n-sets $S_i, T_i \subset [0,1]$ with $i = 1, \ldots, s$, and real numbers $d_1, \ldots, d_s$ so that with $B = \sum_{i=1}^{s} d_i \mathbb{I}_{S_i \times T_i}$ it holds that*

(i) $\|U\|_2 \geq \|U - B\|_2$,

(ii) $\|U - B\|_\square < \Delta$, *and*

(iii) $\sum_{i=1}^{s} |d_i| \leq \frac{1}{\Delta}$.

The next lemma tells us that the original cut norm is preserved under uniform sampling.

Lemma 3.6. *[4] For any $\Delta > 0$ and bounded measurable function $U\colon [0,1]^2 \to [-1,1]$ we have that*

$$\mathbb{P}\left(\left|\|\mathbb{H}(q,U)\|_\square - \|U\|_\square\right| > \Delta\right) < 2\exp\left(-\frac{10}{\Delta^2}\right),$$

where $q \geq \frac{1000}{\Delta^4}$.

The following lemma asserts that if we sample from a linear program a certain way to get a linear program of bounded complexity, then the optimal value of the objective function cannot deviate too much when the right scaling is applied.

Lemma 3.7. *[2], [10] Let n be a positive integer, $c_m\colon [n] \to \mathbb{R}$, $U_{i,m}\colon [n] \to \mathbb{R}$ for $i = 1, \ldots, s$, $m = 1, \ldots, q$, $u \in \mathbb{R}^{s \times q}$, $\alpha \in \mathbb{R}$. If the optimum of the linear program*

$$\begin{aligned}
&\text{maximize} && \frac{1}{n}\sum_{t=1}^{n}\sum_{m=1}^{q} y_{t,m} c_m(t) && \\
&\text{subject to} && \frac{1}{n}\sum_{t=1}^{n} y_{t,m} U_{i,m}(t) \leq u_{i,m} && \text{for } i \in [s] \text{ and } m \in [q] \\
& && 0 \leq y_{t,m} \leq 1 && \text{for } t \in [n] \text{ and } m \in [q] \\
& && \sum_{m=1}^{q} y_{t,m} = 1 && \text{for } t \in [n]
\end{aligned}$$

is less than α, then for any $\Delta > 0$ and $k \leq n$ and the independent uniform random sample $X_1, \ldots, X_k$ of $[n]$ with replacement the optimum of the sampled linear program

$$\begin{aligned} &\text{maximize} && \sum_{j=1}^{k}\sum_{m=1}^{q} \frac{1}{k} x_{j,m} c_m(X_j) && \\ &\text{subject to} && \sum_{j=1}^{k} \frac{1}{k} x_{j,m} U_{i,m}(X_j) \leq u_{i,m} - \Delta \|U\|_\infty && \text{for } i \in [s] \text{ and } m \in [q] \\ &&& 0 \leq x_{j,m} \leq 1 && \text{for } j \in [k] \text{ and } m \in [q] \\ &&& \sum_{m=1}^{q} x_{j,m} = 1 && \text{for } j \in [k] \end{aligned}$$

is less than $\alpha + \Delta$ with probability at least $1 - \exp(-\frac{\Delta^2 k}{2\|c\|_\infty^2})$.

Lemma 3.5 ensures the existence of an integer $s \leq \frac{32t^4}{\Delta^2}$, measurable sets $S_i, T_i \subset [0,1]$ with $i = 1, \ldots, s$, and real numbers $d_1, \ldots, d_s$ such that $\|U - \sum_{i=1}^{s} d_i \mathbb{I}_{S_i \times T_i}\|_\square \leq \Delta/32t^2$ and $\sum_i |d_i| \leq \frac{32t^2}{\Delta}$. The h_A value of this weighted sum of indicator functions approximates $h_{A,f}(U)$ sufficiently well for any fractional n-partition f. Let $D = \sum_{i=1}^{s} d_i \mathbb{I}_{S_i \times T_i}$. Then

$$\begin{aligned} |h_{A,f}(U) - h_{A,f}(D)| &= \left| \sum_{1 \leq i,j \leq 4t} (B_A)_{i,j} \int_{[0,1]^2} U(x,y) f_i(x) f_j(y) \mathrm{d}x \mathrm{d}y \right| \\ &\leq 4t^2 \|U - D\|_\square \leq \Delta. \end{aligned}$$

In the same manner one can introduce a low complexity approximation on the sample $\mathbb{H}(k, U)$, we will show that the image of D mapped via the sampling process is suitable. To do this we only need to define the subsets $[k] \supset \hat{S}_i = \{m : X_m \in S_i\}$ and $[k] \supset \hat{T}_i = \{m : X_m \in T_i\}$. Let $\hat{D} = \sum_{i=1}^{s} d_i \mathbb{I}_{\hat{S}_i \times \hat{T}_i}$. First we condition on the event from Lemma 3.6, call this event E_1, that is $\left| \|\mathbb{H}(k,U) - \hat{D}\|_\square - \|U - D\|_\square \right| < \Delta/32t^2$. Set $D' = W_{\hat{D}}$, then on E_1 it follows that for any $\hat{\mathcal{Q}}$ that is an n-partition into $4t^2$ classes we have

$$\begin{aligned} |h_{A,\hat{\mathcal{Q}}}(U') - h_{A,\hat{\mathcal{Q}}}(D')| &\leq 16t^2 \|\mathbb{H}(k,U) - \hat{D}\|_\square \\ &\leq 16t^2 \|U - D\|_\square + \Delta/2 \leq \Delta. \end{aligned}$$

Let $\mathcal{S} = \{S_i : 1 \leq i \leq 2s\}$ denote the set of measurable sets that occur in the sum that defines D, and let $\mathcal{S}'$ stand for the corresponding set of sets on the sample, note that $|\mathcal{S}| \leq \frac{64t^4}{\Delta^2}$. Define the sets

$$I(b) = \{f : 1 \leq i \leq s, j = 1, \ldots, 4t : |\int_{S_i} f_j - b_{i,j}^{(1)}| \leq \frac{\Delta}{16t^2} \text{ and } |\int_{T_i} f_j - b_{i,j}^{(2)}| \leq \Delta\},$$

and

$$I'(b) = \{x : 1 \leq i \leq s, j = 1, \ldots, 4t : |\frac{1}{q} \sum_{j \in S'_i} x_{j,i} - b^{(1)}_{i,j}| \leq \Delta/2 \text{ and } |\frac{1}{q} \sum_{j \in T'_i} x_{j,i} - b^{(2)}_{i,j}| \leq \frac{\Delta}{32t^2}\}.$$

for each $b \in [0,1]^{8st}$.

We will use the grid points $\mathcal{B} = \{(b^{(\alpha)}_{i,j}) : \forall i, j, \alpha : b^{(\alpha)}_{i,j} \in [0,1] \cap (\Delta/2)\mathbb{Z}\}$.

On every set $I(b)$ we can produce a linear approximation of $h_{A,f}(D)$ (linearity is meant whit respect to the components of f) which carries through to a linear approximation in $I'(b)$ of $h_{A,x}(D')$ via sampling. The precise description of this is given in the next auxiliary result.

Fix $b \in \mathcal{B}$ and define the b dependent real number

$$l_0 = \sum_{i,j=1}^{4t} \sum_{k=1}^{s} (B_A)_{i,j} d_k b^{(1)}_{i,k} b^{(2)}_{j,k},$$

and the n-functions $l_1, l_2, \ldots, l_{4t} \colon [0,1] \to \mathbb{R}$ with

$$l_m(x) = \sum_{j=1}^{4t} \sum_{k=1}^{s} d_k \left[(B_A)_{m,j} b^{(2)}_{j,k} \mathbb{I}_{S_k}(x) + (B_A)_{j,m} b^{(1)}_{j,k} \mathbb{I}_{T_k}(x) \right].$$

Then it is not hard to check (see also [2] and [10]) that the following holds. For every $f \in I(b)$ we have that

$$\left| h_{A,f}(D) - l_0 - \int_0^1 \sum_{m=1}^{4t} f_m(t) l_m(t)(x) \mathrm{d}x \right| < \Delta/2,$$

and for every $\mathrm{x} \in I'(b)$ it is true that

$$\left| h_{A,f}(D') - l_0 - \sum_{n=1}^{q} \sum_{m=1}^{4t} \frac{1}{q} x_{m,n} l_m(X_n) \right| < \Delta/2.$$

Additionally we obtain that $l_1, l_2, \ldots, l_{4t}$ are in the supremum norm bounded from above by $\frac{32t^2}{\Delta}$.

Lemma 3.7 tells us that the event $E_2(A, b)$ comprising the implication that if the linear program

$$\begin{array}{lll} \text{maximize} & l_0 + \sum_{n=1}^{q} \sum_{m=1}^{4t} \frac{1}{q} x_{n,m} l_m(U_n) & \\ \text{subject to} & \mathrm{x} \in I'(b) & \\ & 0 \leq x_{n,m} \leq 1 & \text{for } m = 1, \ldots, q \text{ and } m = 1, \ldots, 4t \end{array}$$

$$\sum_{m=1}^{4t} x_{n,m} = 1 \qquad \text{for } m = 1, \ldots, 4t$$

has optimal value α, then the continuous linear program

$$\begin{aligned} &\text{maximize} && l_0 + \int_0^1 \sum_{m=1}^{4t} l_m(y) f_m(y) \mathrm{d}y \\ &\text{subject to} && f \in I(b) \\ & && 0 \le f_m(y) \le 1 && \text{for } y \in [0,1] \text{ and } m = 1, \ldots, 4t \\ & && \sum_{m=1}^{4t} f_m(y) = 1 && \text{for } y \in [0,1] \end{aligned}$$

has optimal value at least $\alpha - \Delta/2$ has probability at least $1 - \exp(-\frac{\Delta^4 q}{10^5 t^6})$.

Condition on E_1 and E_2, where the second event is the simultaneous occurrence of $E_2(A, b)$ for each $A \in \mathcal{A}$ and $b \in \mathcal{B}$. E_2 has failure probability at most $\exp(-t)$ whenever $q \ge t^{14}$.

Let $A \in \mathcal{A}$ and also let x be an arbitrary fractional n-partition such that $f \in I'(b_0)$ for some $b_0 \in \mathcal{B}$. Then there exists a fractional n-partition $g \in I(b_0)$ such that

$$\begin{aligned} h_{A,x}(U') - h_{A,f}(U) &\le h_{A,x}(D') - h_{A,f}(D) + \Delta/8 \\ &\le \sum_{n=1}^{q} \sum_{m=1}^{4t} \frac{1}{q} x_{m,n} l_m(X_n) - \int_0^1 \sum_{m=1}^{4t} f_m(t) l_m(t)(x) \mathrm{d}x + \Delta/4 \\ &\le \Delta/2. \end{aligned}$$

This shows eventually that with probability at least $1 - \exp(-t)$ we have that

$$\max_{A \in \mathcal{A}} \max_{\hat{\mathcal{Q}}} \sum_{i,j=1}^{t_{\mathcal{P}}} \sum_{\alpha,\beta=1}^{4} (B_A)_{i,j}^{\alpha,\beta} \int_{[0,1]^2} (W_{\mathbb{G}(q, W_G - V)})(x,y) \mathbb{I}_{Q_i^\alpha}(x) \mathbb{I}_{Q_j^\beta}(y) \mathrm{d}x \mathrm{d}y \le \Delta. \tag{3.6}$$

This however is equivalent to saying that for every $\mathcal{Q}$ partition into $t_{\mathcal{Q}}$ classes $t_{\mathcal{Q}} \le t_{\mathcal{P}}$ it is true that

$$\|W_{\mathbb{G}(q, W_G - V)}\|_{\square \mathcal{Q}} \le \Delta/2. \tag{3.7}$$

The second estimate we require concerns the closeness of the step function V and its sample. Our aim is to overlay these two functions via measure preserving permutations of $[0,1]$, such that the measure of the subset of $[0,1]^2$ where they differ is as small as possible.

Let $V' = W_{\mathbb{H}(q,V)}$, this n-function is well-defined this way and is a step function with steps forming the n-partition $\mathcal{P}'$. This latter n-partition of $[0,1]$ is the image of $\mathcal{P}$ induced by the sample and the map $i \mapsto [\frac{i-1}{q}, \frac{i}{q})$. Let ψ be a measure preserving n-permutation of $[0,1]$ that satisfies that for each $i \in [t]$ the volumes $\lambda(P_i \triangle \psi(P_i')) = |\lambda(P_i) - \lambda(P_i')|$. Let

$\mathcal{P}''$ denote the partition with classes $\mathcal{P}_i'' = \psi(P_i')$ and $V'' = (V')^\psi$ (note that V'' and V' are equivalent as graphons), furthermore let N_V be the subset of $[0,1]^2$ where the two functions V and V'' differ. Then

$$\mathbb{E}[\lambda(N_V)] \leq 2\mathbb{E}[\sum_{i=1}^{t'} |\lambda(P_i) - \lambda(P_i')|]. \tag{3.8}$$

The random variables $\lambda(P_i')$ for each i can be interpreted as the average positive outcome of q independent Bernoulli trials with success probability $\lambda(P_i)$. It follows that

$$\mathbb{E}[\sum_{i=1}^{t} |\lambda(P_i) - \lambda(P_i')|] \leq \sqrt{t'\mathbb{E}[\sum_{i=1}^{t} (\lambda(P_i) - \lambda(P_i'))^2]} \leq \sqrt{\frac{t}{q}}. \tag{3.9}$$

This calculation yields that $\mathbb{E}[\lambda(N_V)] \leq \sqrt{\frac{4t}{q}}$. Standard concentration result gives us that $\lambda(N_V)$ is also small in probability if q is chosen large enough. For convenience, define the martingale $M_l = \mathbb{E}[\lambda(N_V)|X_1, \ldots, X_l]$ for $1 \leq l \leq q$, and recognize that the martingale differences are uniformly bounded: $|M_l - M_{l-1}| \leq \frac{4}{q}$. Azuma's inequality then yields that

$$\mathbb{P}(\lambda(N_V) \geq \sqrt{\frac{4t}{q}} + \alpha) \leq \mathbb{P}(\lambda(N_V) \geq \mathbb{E}[\lambda(N_V)] + \alpha) \leq \exp(-\alpha^2 q/32). \tag{3.10}$$

Define the event E_3 that holds whenever $\lambda(N_V) \leq \sqrt{\frac{4t}{q}} + q^{-1/4}$ and condition on it in addition to the above events E_1 and E_2. It follows from (3.10) that the failure probability of E_3 is at most $\exp(-t)$.

It follows that there exists an n-permutation of $[0,1]$ denoted by ϕ such that $\|(V')^\phi - V\|_1 \leq \Delta/4$. Now employing the triangle inequality and the bound (3.7) we get for all $\mathcal{Q}$ n-partitions into t parts that

$$\|W_G - (W_F)^\phi\|_{\square \mathcal{Q}} \leq \|W_G - V\|_{\square \mathcal{Q}} + \|V - (V')^\phi\|_1 + \|(V')^\phi - (W_F)^\phi\|_{\square \phi(\mathcal{Q})} \leq \Delta \tag{3.11}$$

Now let $U \in \mathcal{M}$ be arbitrary, and let $\mathcal{P}_U$ denote the partition consisting of the steps of U. Let ψ be the n permutation of $[0,1]$ that delivers $d_{U,\mathcal{P}_U}(G) = \|U - (W_G)^\psi\|_{\square \mathcal{P}_U}$. Then

$$d_{U,\mathcal{P}_U}(G) - d_{U,\mathcal{P}_U}(F) \leq \|U - (W_G)^\psi\|_{\square \mathcal{P}_U} - \|U - (W_F)^{\phi\psi}\|_{\square \mathcal{P}_U} \tag{3.12}$$

$$\leq \|W_G - (W_F)^\phi\|_{\square \psi^{-1}(\mathcal{P}_U)} \leq \Delta. \tag{3.13}$$

The lower bound on the above difference can be handled in a similar way, therefore we have that $|d_{U,\mathcal{P}_U}(G) - d_{U,\mathcal{P}_U}(F)| \leq \Delta$ for every $U \in \mathcal{M}$.

We conclude the proof with mentioning that the failure probability of the three events E_1, E_2, and E_3 taking place simultaneously is at most $3\exp(-t)$. □

We are now ready to conduct the proof of the main result.

Proof of Theorem 1.5.

Let us fix $\epsilon > 0$ and the simple graph G. To establish the lower bound on $f(\mathbb{G}(q_f, G))$ not much effort is required: we pick a coloring $\mathbf{G}$ of G that certifies the value $f(G)$, that is, $g(\mathbf{G}) = f(G)$. Then the coloring of $\mathbb{G}(G, q_g)$ induced by $\mathbf{G}$, which we call $\mathbf{F}$, satisfies $g(\mathbf{F}) \geq g(\mathbf{G}) - \epsilon/2$ with probability at least $1 - \epsilon/2$, simply due to the testability condition on g, which in turn implies $f(\mathbb{G}(q_g, G)) \geq f(G) - \epsilon/2$ with probability at least $1 - \epsilon/2$. So the condition $q_f(\epsilon) \geq q_g(\epsilon/2)$ is sufficient for this part.

The problem concerning the upper bound in terms of q on $f(\mathbb{G}(q, G))$ is the difficult part of the proof, the rest of it deals with this case. We introduce the error parameter $\Delta > 0$, that is an explicit function of $\epsilon > 0$, the precise connection will be stated later.

Let us condition on the event in the statement of Lemma 3.4. Let $\mathcal{N}$ be the set of all k-colored digraphs $\mathbf{W}$ that are step functions with at most $t_k(\Delta/2)$ equal canonical steps $\mathcal{P}$, and for $U = \sum_{(\alpha,\beta)\in M} W^{(\alpha,\beta)}$ we have $d_{U,\mathcal{P}}(G) \leq 2\Delta$.

Our main step in the proof will be that, conditioned on the aforementioned event, we can find for each (k, m)-coloring of F a corresponding coloring of G so that the g values of the two colored instances are sufficiently close. We will make this argument precise in the following.

Let us fix an arbitrary (k, m)-coloring of F denoted by $\mathbf{F}$. Lemma 2.13 implies that there exist a $\mathbf{W}$ that is a step function with at most $t_k(\Delta/2)$ equal canonical steps $\mathcal{P}$ whose values are integer multiples of $\Delta/2$ between 0 and 1, such that there exists a ϕ measure-preserving permutation of $[0, 1]$ such that $d_{\square\mathcal{P}}((\mathbf{W}_\mathbf{F})^\phi, \mathbf{W}) \leq \Delta$. Therefore with $U = \sum_{(\alpha,\beta)\in M} W^{(\alpha,\beta)}$ we have $d_{U,\mathcal{P}}(F) \leq \Delta$ with $U \in \mathcal{M}$, which in turn implies by the conditioned event that $d_{U,\mathcal{P}}(G) \leq 2\Delta$, and consequently $\mathbf{W} \in \mathcal{N}$. It follows from Lemma 3.2 that there exists a (k, m)-coloring of G denoted by $\mathbf{G}$ such that $d_\square(\mathbf{W}, (\mathbf{W}_\mathbf{G})^\psi) \leq 2k^2\Delta$ for some ψ.

Therefore we get that $\delta_\square(\mathbf{G}, \mathbf{F}) \leq (2k + 1)\Delta$. Now we have to choose Δ small enough so that by Lemma 3.1 we can assert that $|g(\mathbf{G}) - g(\mathbf{F})| \leq \epsilon/2$, $\Delta = \frac{2^{-q_g^2(\epsilon/2)\log k}}{2k+1}$ will do. This finishes our argument, as $\mathbf{F}$ was arbitrary and the sample size was chosen in a way that $q \leq \exp^{(2)}(O(1/\Delta^2)) \leq \exp^{(3)}(O(q_g^2(\epsilon/2)))$, where the big-O hides also the role of k.

□

4 Generalizations and special cases

We will extract three possible directions of further research specifically with respect to the framework of the current paper, and will provide partial answers to questions posed by the authors of [14].

First we will introduce an even more restrictive notion of nondeterminism (the definition used in the current paper's previous section is a special case of the notion used commonly in the complexity theory), relying on this we are able to improve on the sample complexity upper bound of weakly nondeterministically testable graph parameters using a simplified version of our approach applied in the proof of Theorem 1.5 without significant alterations. Secondly, we will take an outlook on nondeterministically testable graph parameters whose witness parameter has polynomial sample complexity in $\frac{1}{\epsilon}$, where ϵ is the error parameter, and will compare the approach with the case of MAXCUT, whose sample complexity is known to be polynomial in $\frac{1}{\epsilon}$. The third point will concern a generalization of the current framework to uniform hypergraphs of higher order.

4.1 Weak nondeterminism

We formulate the definition of a stronger property than the previously defined non-deterministic testability. The notion itself may seem at first more complicated, but in fact it only corresponds to the case, where the witness parameter g of f for a graph G is evaluated only on the set of node-colorings of G instead of edge-colorings in order to define the maximum expression. This modification will enable us to rely only on the cut-norm and the corresponding regularity lemma instead of the cut-$\mathcal{P}$-norm, thus leading us to better upper bounds on the sample complexity of f with respect to that of g. This time we only treat the case of undirected graph colorings in detail, the directed case is analogous.

We will introduce the set of colorings of G called node-(k, m)-colorings. Let $\mathcal{T} = (T_1, \dots, T_k)$ be a partition of $V(G)$ and $\mathcal{D} = (D_1, \dots, D_m)$, $\mathcal{D}' = (D'_1, \dots, D'_m)$ be two partitions of $[t]^2$, together they induce two partitions, $\mathcal{C} = (C_1, \dots, C_m)$ and $\mathcal{C}' = (C'_1, \dots, C'_m)$, of $V(G)^2$ such that each class is of the form $C_\alpha = \cup_{(i,j)\in D_\alpha} T_i \times T_j$ and $C'_\alpha = \cup_{(i,j)\in D'_\alpha} T_i \times T_j$ respectively. A node-(k, m)-coloring of G is defined by some $\mathcal{C}$ of the previous form and is the $2m$-tuple of simple graphs $\mathbf{G} = (G_1, \dots, G_m, \tilde{G}_1, \dots, \tilde{G}_m)$ with $G_\alpha = G[C_\alpha]$ and $\tilde{G}_\alpha = G^c[C'_\alpha]$. Here G^c stands for the complement of G (the union of G and its complement is the directed complete graph with all loops present), and $G[C_\alpha]$ is the union of induced subgraphs of G between T_i and T_j for each $(i, j) \in D_\alpha$ for $i \neq j$, in the case of $i = j$ the term in the union is the induced subgraph of G on the node set T_i.

Definition 4.1. *The graph parameter f is weakly non-deterministically testable if there exist integers m and k with $m \leq k$ and a testable edge-k-colored directed graph parameter g such that for any simple graph G the value $f(G) = \max_{\mathbf{G}'=G} g(\mathbf{G})$, where the maximum goes over the set of node-(k, m)-colorings of G.*

The following lemma is the analogous result to Lemma 3.2 that can be employed in the proof of Theorem 1.5 in the special case of weakly non-deterministically testable graph parameters.

Lemma 4.2. *Let $\epsilon > 0$, be a U and V be arbitrary graphons with $\|U - V\|_\square \leq \epsilon$, and also let $k \geq 1$. For any $\mathbf{U} = (U^{(1)}, \dots, U^{(m)}, \tilde{U}^{(1)}, \dots, \tilde{U}^{(m)})$ node-(k, m)-coloring of U there exists a node-(k, m)-coloring of V denoted by $\mathbf{V} = (V^{(1)}, \dots, V^{(k)}, \tilde{V}^{(1)}, \dots, \tilde{V}^{(m)})$ so that $d_\square(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^m \|U^{(i)} - V^{(i)}\|_\square + \sum_{i=1}^m \|\tilde{U}^{(i)} - \tilde{V}^{(i)}\|_\square \leq 2k^2\epsilon$. If $V = W_G$ for some simple graph G on n nodes and each $U^{(i)}$ is an n step function then there is a coloring $\mathbf{G}$ of G such that $d_\square(\mathbf{U}, \mathbf{W_G}) \leq 2k^2\epsilon$.*

Proof. Our approach is quite elementary: consider the partition $\mathcal{T}$ of $[0, 1]$ and $\mathcal{C}$ of $[0, 1]^2$ that provide $\mathbf{U}$ and define $V^{(i)} = V\mathbb{I}_{C_i}$ and $\tilde{V}^{(i)} = (1 - V)\mathbb{I}_{C_i}$ for each $i \in [m]$. Then

$$\|U^{(i)} - V^{(i)}\|_\square \leq \sum_{(\alpha,\beta)\in D_i} \|(U - V)\mathbb{I}_{T_\alpha \times T_\beta}\|_\square \leq \epsilon |D_i| \tag{4.1}$$

for each $i \in [m]$, and the same upper bound applies to $\|\tilde{U}^{(i)} - \tilde{V}^{(i)}\|_\square$. Summing up over i gives the result stated in the lemma.

The argument showing the part regarding simple graphs is identical. $\square$

Note that in Lemma 3.2 we required U and V to be close in the cut-$\mathcal{P}$-norm for some partition $\mathcal{P}$ and U to be a $\mathcal{P}$ step function to guarantee for each $\mathbf{U}$ the existence of $\mathbf{V}$ that is close to it in the cut distance of k-colored digraphons. Using the fact that in the weakly non-deterministic framework cut-closeness of instances implies the cut-closeness of the sets of their node-(k, m)-colorings we can formulate the next corollary of Theorem 1.5 that is the main result of this subsection.

Corollary 4.3. *Let f be a testable graph parameter with weak non-deterministic witness parameter g of node-(k, m)-colored graphs, and let the corresponding sample complexity functions be q_f and q_g. Then there for any $\epsilon > 0$ we have $q_f(\epsilon) \leq \exp^{(2)}(\mathrm{poly}(q_g(\epsilon/2)))$.*

Proof. We will give only a sketch of the proof, as it is almost identical to that of Theorem 1.5, and we automatically refer to that, including the notation used in the proof, if not noted otherwise. The part concerning the lower bound of $f(\mathbb{G}(q, G))$ is completely identical. For the upper bound we have to replace $\mathcal{M}$ by its subset $\mathcal{M}'$ that consists only of the step functions with at most $t'_k(\Delta/2)$ steps.

Let $\epsilon > 0$ be arbitrary, and let $\Delta > 0$ to be specified later as a functions of $\epsilon > 0$. We condition now on the event that $\delta_\square(G, \mathbb{G}(q, G) \leq \Delta$, whose probability is sufficiently small due to Lemma 2.17. Now we select an arbitrary (k, m)-coloring $\mathbf{F}$ of $\mathbb{G}(q, G)$ and apply the Weak Regularity Lemma for k-colored graphons, Lemma 2.9, in the n step function case with error parameter $\Delta/(2k^2 + 1)$ to get a tuple of n-step functions $\mathbf{U}$ with $t'_k(\Delta/(2k^2 + 1))$ steps. We define the n-graphon $U = \sum_{i=1}^m U_i$ and observe that our condition implies that $\delta(G, U) \leq 2\Delta$. To finish of the proof apply Lemma 4.2, it implies the existence of a coloring $\mathbf{G}$ of G so that $\delta_\square(\mathbf{G}, \mathbf{F}) \leq (2k^2 + 1)\Delta$. Setting Δ to $\exp(-cq_q^2(\epsilon))$ and applying Lemma 3.1 delivers the desired result. $\square$

4.2 Polynomially non-deterministically testable graph parameters

This subsection deals with the special case of non-deterministically testable parameters whose witness parameter is testable with sample complexity that is polynomial in $\frac{1}{\epsilon}$. The aim in this setting would be naturally to investigate if polynomial testability of the witness implies the polynomial testabilty of the parameter in consideration. We were not able to provide any improvement in general to this issue, but propose the following reduction of a certain special case.

We stick to the previously presented framework of weak nondeterminism using some variant of weak regularity, therefore we impose the additional condition that witness parameters should obey α-Hölder-continuity in the cut metric for some fixed $\alpha > 0$. That means that if we find an Δ-approximating graphon step function W to our graph G in the cut metric (its existence is provided by the Weak Regularity Lemma) such that their g values are ϵ-close, then it suffices to take a sample whose size that is $\mathrm{poly}(1/\Delta)$, and therefore $\mathrm{poly}(1/\epsilon)$, so that the sampled graphon W' and $\mathbb{G}(q, G)$ are cut-(2Δ)-close with high probability, and therefore their g values are $(2^{1/\alpha}\epsilon)$-close, also with high probability. Here we assumed that the parameter g is defined also for graphons, this assumption spares us some technical difficulties that otherwise would have been to overcome. We also want to use the version of the Weak Regularity Lemma that has the least conditions on the approximating step function (without equal sizes of classes), but the number of steps of W, s, still exponential in $1/\Delta^2$ in general, although W can also be regarded as the weighted sum of $4/\Delta^2$ indicator functions of the form $\mathbb{I}_{S\times T}$.

Let $W = \sum_{i=1}^{s} d_i \mathbb{I}_{S_i\times T_i}$ and $W' = \sum_{i=1}^{s} d_i \mathbb{I}_{S'_i\times T'_i}$. Notice that a priori W' has no explicit form and can be transformed via measure-preserving permutations of $[0,1]$ without changing its graphon equivalence class, therefore we only fix one of these representations. The only thing that we can rely on that the sizes of the atoms defined by the sets $\{S'_1,\dots,S'_s,T'_1,\dots,T'_s\}$ (which are random variables, actually average outcomes of certain Bernoulli trials) have to respect the sampling procedure. Roughly said, if the sample size is at least of square order of the number of atoms, that is, it is at least $2^{8/\Delta^2}$, than each atom size can be approximated well, so W and W' can be arranged by measure-preserving permutations in a way that the L^1-norm of their difference is small, and therefore also their cut distance is small.

The most important problem is the following: Is there a way to get rid of the necessity of two step functions (of the previous special form) being close in the cut metric to achieve closeness in their g values by relying on the fact that the sample complexity of g is much smaller than the number of steps (atoms) of the aforementioned step functions, but is comparable to the number of sets defining these atoms? We finish with the exact formulation of this to open problem whose solution would shed more light on the sample complexity of parameters that are weakly nondeterministically testable.

Let $(\Omega,\mathcal{A},\mu)$ be a probability space and h be a parameter of random variables $\{X|X : \Omega\to[-d,d] \text{ measurable}\}$ with $d>0$ being some bound, that is, the h value of two random variables is identical, whenever they have the same distribution function. Let $X_1,\dots,X_q$ be independent samples according to the distribution of some random variable X, and define the random variable $X[q]:\Omega\to\mathbb{R}$ that takes the values $X_1,\dots,X_q$ with probability proportional to the frequency of their appearance. Suppose that there is a function $q_g:\mathbb{R}^+\to\mathbb{R}^+$ such that $|h(X)-\mathbb{E}[h(X[q_g(\epsilon)])]|<\epsilon$ for every X, and that h is L^1-continuous.

Question 4.4. *Let X and Y be two random variables and $s>0$, $d_1,\dots d_s$, such that $X=\sum_{i=1}^{s} d_iB_i$ and $Y=\sum_{i=1}^{s} d_iC_i$, where $B_1,\dots,B_s,C_1,\dots,C_s$ are Bernoulli random variables (they are not assumed to be independent, but Y can be thought of as a random variable in Ω' that is a copy of Ω). Does a function t exist that does not depend on s, but $\log(t(\epsilon))=\Theta(\log(q_g(\epsilon)))$, such that for each $\epsilon>0$, $t'\le t(\epsilon)$ and $i_1,\dots,i_{t'}$, if $|\mathbb{E}[B_{i_1}\dots B_{i_{t'}}]-\mathbb{E}[C_{i_1}\dots C_{i_{t'}}]|\le\epsilon^{st'}$ holds, then $|h(X)-h(Y)|\le\epsilon$?*

Acknowledgement

We thank Laci Lovász for interesting discussions connected to the subject of this paper.

References

[1] N. Alon, E. Fischer, I. Newman, and A. Shapira: A Combinatorial Characterization of the Testable Graph Properties: It's All About Regularity. SIAM J. Comput. 39(1): 143-167 (2009)

[2] N. Alon , W. F. de la Vega , R. Kannan , M. Karpinski: Random Sampling and Approximation of MAX-CSPs, Journal of Computer and System Sciences, 67 (2) (2003), 212-243.

[3] S. Arora, D. Karger and M. Karpinski: Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems, Proc. 13th ACM STOC (1995), 193-210; also in Journal of Computer and System Sciences 58 (1999), 193-210.

[4] C. Borgs, J. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi: Convergent Sequences of Dense Graphs I: Subgraph Frequencies, Metric Properties and Testing, Advances in Math. 219 (2008), 1801-1851.

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

[6] P. Diaconis and S. Janson: Graph Limits and Exchangeable Random Graphs, Rendiconti di Matematica, Serie VII 28, (2008), 33-61.

[7] G. Elek and B. Szegedy: A measure-theoretic approach to the theory of dense hypergraphs Advances in Mathematics 231 (2012), 1731-1772.

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

[9] L. Gishboliner and A. Shapira: Deterministic vs Non-deterministic Graph Property Testing. Electronic Colloquium on Computational Complexity 20 (2013), 59.

[10] M. Karpinski and R. Markó: Limits of CSP Problems and Efficient Parameter Testing, arXiv: 1406.3514v1 (2014).

[11] L. Lovász and B. Szegedy: Limits of dense graph sequences, J. Comb. Theory B 96 (2006), 933-957.

[12] L. Lovász and B. Szegedy: Limits of compact decorated graphs (2010) arXiv: 1010.5155v1 (2010).

[13] L.Lovász and B. Szegedy: Szemerédi's lemma for the analyst, Geometric and Functional Analysis, Vol. 17 (1) (2007), 252-270.

[14] L. Lovász and K. Vesztergombi: Non-Deterministic Graph Property Testing. Combinatorics, Probability & Computing 22(5) (2013), 749-762.

[15] L. Lovász: Large networks and graph limits (American Mathematical Society) (2012).