Complexity of Nondeterministic Graph Parameter Testing (Revised Version)

Marek Karpinski[*]   Roland Markó[†]

Abstract

We study the sample complexity of nondeterministically testable graph parameters and improve existing bounds on it by several orders of magnitude. The technique used would be also of independent interest. We also discuss the special case of weak nondeterministic testing for uniform hypergraphs of arbitrary order.

1 Introduction

In this paper we investigate the estimation of graph parameters by means of uniform vertex sampling. We focus on the novel concept of nondeterministic graph parameter and property testing.

We call a non-negative function on the set of labeled simple graphs a *graph parameter* if it is invariant under graph isomorphism, i.e. relabeling of vertices. We define parameters of edge-k-colored directed graphs, which will be considered in this paper as loop-free and complete in the sense that each directed edge is present and carries exactly one color, and introduce analogously the concept also for graphons, the limit objects of dense graphs [3], [9]. The central characteristic of parameters investigated in the current paper is whether it is the possible to estimate a given value of the parameter up to a desired accuracy via uniform sampling of bounded size that is independent from the size of the input graph, in short, is the parameter testable. If the answer to this question is positive, then we can ask for the smallest sample size that is sufficient for this purpose. For a graph G (directed and k-colored possibly) the expression $\mathbb{G}(q, G)$ denotes the random induced subgraph of G with the vertex set chosen uniformly among all subsets of $V(G)$ that have cardinality q.

Definition 1.1. *The graph parameter f is testable if for any $\varepsilon > 0$ there exists a positive integer $q_0(\varepsilon)$ such that for any $q \geq q_0(\varepsilon)$ and simple graph G with at least $q_0(\varepsilon)$ nodes*

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

[*]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. 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 smallest function q_0 satisfying the previous inequality is called the sample complexity of f and is denoted by q_f. The testability of parameters of k-colored directed graphs and uniform hypergraphs is defined analogously.

From now on colored means edge-colored if not noted otherwise and we generally assume that $q_f(\varepsilon) \geq \max\{1, 1/\varepsilon\}$. An a priori weaker notion than testability is the second cornerstone of the current work, it was introduced in [12].

Definition 1.2. *The graph parameter f is non-deterministically testable if there exist integers $k \geq m$ and a testable k-colored directed 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, m)-colorings of G. The edge-k-colored directed graph $\mathbf{G}$ is a (k, m)-coloring of G, if after erasing all edges of $\mathbf{G}$ colored with an element of $[m+1, \ldots, k]$ and discarding the orientation, coloring, and multiplicity of the remaining edges we end up with G. We say in this case that G is the shadow of $\mathbf{G}$.*

The corresponding definition for r-uniform hypergraphs (in short, r-graphs) is analogous. The choice of maximizing over the g-values in Definition 1.2 is somewhat arbitrary, in a more general sense we could have $f(G) = g(\mathrm{argmax}_{\mathbf{G}} L(g(\mathbf{G})))$ for any α-Hölder continuous function L from $\mathbb{R}$ to $\mathbb{R}$. Also, stronger formulations of being a witness can be employed, such as permitting only undirected instances or imposing $k = 2m$.

1.1 Previous work

The problem regarding the relationship of the class of parameters that are testable and those who are non-deterministically testable was first studied in the framework of dense graph limits and property testing by Lovász and Vesztergombi [12] in the spirit of the general "P vs. NP" question, that is a central problem in theoretical computer science. Using the particular notion of nondeterminism above 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. *[12] Every non-deterministically testable graph property $\mathcal{P}$ is testable. The same holds in parameter testing.*

However, no explicit relationship was provided between on one hand, the sample size required for estimating the f value, and on the other, the two factors, the number of colors k and m, and the sample complexity of the witness g. The reason for the non-efficient characteristic of the result is that the authors exploited various consequences of the next remarkable fact.

Graphons are bounded symmetric measurable functions on the unit square, their cut norm $\|.\|_\square$ given in Definition 2.4 below. At this point we wish to stress that it is weaker than the L^1-norm, and the $\delta_\square$-distance induced by it combined with an optimal overlay has a compact unit ball, this is not the case for the δ_1-distance generated analogously by the L^1-norm.

2

Fact. If $(W_n)_{n \geq 1}$ is a sequence of graphons and $\|W_n\|_\square \to 0$ when n tends to infinity, then 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 point-wise.

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 nondeterministically testable property $\mathcal{P}$ and its colored witness $\mathcal{Q}$ was analyzed by Gishboliner and Shapira [7] relying on Szemerédi's Regularity Lemma and its connections to graph property testing unveiled by Alon, Fischer, Newman, and Shapira [2]. In the upper bound given in [7] the height of the exponential tower was not bounded and growing as a function of the inverse of the accuracy, $1/\varepsilon$, the main result of [7] for parameters can be rephrased as follows.

Theorem 1.4. *[7] Every non-deterministically testable graph parameter f is testable. If the sample complexity of the witness parameter g for each $\varepsilon > 0$ is $q_g(\varepsilon)$, then the sample complexity of f for each $\varepsilon > 0$ is at most $\mathrm{tf}(cq_g(\varepsilon/2))$ for some universal constant $c > 0$, where $\mathrm{tf}(t)$ is the exponential tower of twos of height t.*

1.2   Our contribution

In the current paper we improve on the result of [7] by using a weaker type of regularity approach which eliminates the tower-type dependence on the sample complexity of the witness parameter. The function $\exp^{(t)}$ stands for the t-fold iteration of the exponential function $(\exp^{(0)} = \mathrm{id})$. Our main result is the following.

Theorem 1.5. *Let f be a nondeterministically testable simple graph parameter with witness parameter g of k-colored digraphs, and let the corresponding sample complexity be q_g. Then f is testable with sample complexity q_f, and there exists a constant $c > 0$ only depending on k but not on f or g such that for any $\varepsilon > 0$ the inequality $q_f(\varepsilon) \leq \exp^{(3)}(cq_g^2(\varepsilon/2))$ holds.*

We also investigate the case where only node colored graphs and their parameters serve as nondeterministic certificate, and improve on the above upper bound in that setting. Moreover, we extend the method to be able to deal with hypergraphs of higher rank, see below for the precise formulation.

1.3   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, our 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 a special case of the non-deterministic testing notion applied in the current paper.

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 preserve binary relationships, that is for each $u, v \in V(F)$ we have $\phi(u)\phi(v) \in E(G)$ if and only if when $uv \in E(F)$, in particular, a homomorphism has to be injective. In some previous works only the presence of edges had to be sustained by ϕ. Furthermore, let $t(F, G) = \frac{\hom(F,G)}{|V(G)|^{|V(F)|}}$ denote the subgraph density of F in G. The density $t(\mathbf{F}, \mathbf{G})$ in the case of k-colored digraphs is defined analogously.

Definition 2.1. *[9] Let $(G_n)_{n \geq 1}$ be a sequence of simple graphs. It is said to be convergent if for every simple graph 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 digraphs analogously.*

We now describe the space of limit objects of simple graphs in the sense of Lovász and Szegedy [9]. Let $\mathcal{W}_0$ be the set of all bounded measurable functions $W\colon [0,1] \times [0,1] \to \mathbb{R}$, these objects are called *kernels*. The subspace $\mathcal{W}$ of $[0,1]$-valued elements of $\mathcal{W}_0$ that are symmetric in the sense that $W(x, y) = W(y, x)$ for all $x, y \in [0,1]$ is the space of *graphons*. The space of k-colored directed graphons can be described in a similar, though more involved way. Let $\mathcal{W}_0^{(k)}$ be the set of $k \times k$-tuples $\mathbf{W} = (W^{(\alpha,\beta)})_{\alpha,\beta \in [k]}$ of kernels referred to as k-*colored dikernels*. The subspace $\mathcal{W}^{(k)}$ of $\mathcal{W}_0^{(k)}$ whose components obey a symmetry in the sense that $W^{(\alpha,\beta)}(x, y) = W^{(\beta,\alpha)}(y, x)$ for each $x, y \in [0,1]$, are non-negative, and additionally satisfy $\sum_{\alpha,\beta \in [k]} W^{(\alpha,\beta)}(x, y) = 1$ for each $x, y \in [0,1]$ is referred to as the space of k-*colored directed graphons*. Note that for each set $[m]$ with $m \leq k$ the function $W'(x, y) = \sum_{\alpha,\beta \in [m]} W^{(\alpha,\beta)}(x, y)$ is a graphon 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.

A step function is a real-valued function on $[0,1]^2$ that is constant on the product sets $P_i \times Q_j$ for some pair of partitions $\mathcal{P}$ and $\mathcal{Q}$ of $[0,1]$ into the same number of classes referred to as steps, the step function is proper when $\mathcal{P} = \mathcal{Q}$. For a partition $\mathcal{P}$ the integer $t_{\mathcal{P}}$ denotes the number of its classes. We call a partition $\mathcal{P}$ of $[0,1]$ a *canonical n-partition* if its classes are the intervals $P_i = [\frac{i-1}{n}, \frac{i}{n})$ for each $i \in [n]$. If the canonical n-partition refines a partition, then we speak of an $\mathcal{I}_n$-partition, and an $\mathcal{I}_n$-set is the union of some classes of the canonical n-partition. Further, a measure-preserving map from $[0,1]$ to $[0,1]$ is referred to as an $\mathcal{I}_n$-permutation if it corresponds to a permutation of the classes of the canonical n-partition. Functions on $[0,1]$ and $[0,1]^2$ are called $\mathcal{I}_n$-functions if they are constant on the classes of the canonical n-partition and products of those, respectively. We require these concepts to be able to relate graphs on different vertex sets to each other in a simple and computationally efficient way.

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 n-partition and taking the 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 and set $W_{\mathbf{G}}^{(\alpha,\beta)}$ 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. Here on the diagonal cubes $P_i \times P_i$ we require the special color ι (standing for undefined color), the overall measure of the diagonal cubes is $O(1/|V(\mathbf{G})|)$.

Next we define the sampling process for the objects in consideration.

Definition 2.2. *Let $q \geq 1$, 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 analogously, same applies for r-uniform hypergraphs for arbitrary r.*

Let W be a graphon and $q \geq 1$, furthermore, $(X_i)_{i \in [q]}$ and $(Y_{ij})_{ij \in \binom{[q]}{2}}$ be 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 the vertices i and j if $Y_{ij} \geq W(X_i, X_j)$. The random k-colored directed graph $\mathbb{G}(q, \mathbf{W})$ has also vertex set $[q]$, and further, conditioned on the choice of $(X_i)_{i \in [q]}$, the colors for the edges in the two directions are chosen independently for all pairs $ij \in \binom{[q]}{2}$ of vertices, the event that (i, j) carries the color α and at the same time (j, i) carries the color β has 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 as random objects.

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

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

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

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

where the product in the integral is taken over all unordered pairs $ij \in \binom{[k]}{2}$.

The next theorem, first proven in [9], states that the graphons truly represent the limit space of simple graphs. For the proof of the more general cases, see [5], [8], and [11].

Theorem 2.3. *[9], [11] 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 k-colored digraph $\mathbf{F}$ it holds that $t(\mathbf{F}, \mathbf{G}_n) \to t(\mathbf{F}, \mathbf{W})$.*

We proceed by enumerating the norms and distances that are relevant for the current work and are related to the graph limit theory and parameter testing.

5

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

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

where $A(S,T) = \sum_{s \in S, t \in T} A(s,t)$.
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 kernel W is

$$\|W\|_\square = \max_{S,T \subset [0,1]} \left| \int_{S \times T} W(x,y) \mathrm{d}x \mathrm{d}y \right|, \tag{2.1}$$

where maximum is taken over all pairs of measurable sets S and T. We speak of the n-cut norm of kernels when the maximum in (2.1) is only taken over pairs of $\mathcal{I}_n$-sets, it is denoted by $\|W\|_\square^{\langle n \rangle}$. The cut norm of a $k \times k$-tuple of kernels $\mathbf{W} = (W^{(\alpha,\beta)})_{\alpha,\beta=1}^k$ is

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

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 pairs of measure-preserving map from $[0,1]$ to $[0,1]$, 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

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

where $F[t]$ is the t-fold equitable blow up of F and the minimum goes over all node relabellings ϕ of G. In the case $n = m$ we omit the upper index and use $\hat{\delta}_\square$.

In fact, $\hat{\delta}_\square$ and $\delta_\square$ define only pseudometrics, graphs have distance zero whenever they have equitable blow-ups that are isomorphic. For graphons we introduce the term graphon equivalence for the case whenever the $\delta_\square$ distance is 0, but will refer to the above with a slight abuse of notation as proper distances.

Observe that for two graphs F and G on the common node set $[n]$ the distance $d_\square(F, G) = \|W_F - W_G\|_\square = \|W_F - W_G\|_\square^{\langle n \rangle}$. Also note that in general for F and G with identical vertex cardinalities $\delta_\square(F, G)$ is not necessarily equal to $\delta_\square^{\langle n \rangle}(F, G)$, however in [3] it was demonstrated that $\delta_\square(F, G) \leq \hat{\delta}_\square(F, G) \leq 32(\delta_\square(F, G))^{1/67}$.

An important property of the distances introduced above is that subgraph densities are uniformly continuous in the topology defined by them.

Lemma 2.5. *[3] Let U and W be two graphons. Then for every simple graph F on q vertices we have*

$$|t(F, W) - t(F, U)| \leq \binom{q}{2} \delta_\square(U, W).$$

The analogous result holds for k-colored digraphons.

The connection to graph limits is given in the next theorem from [3].

Theorem 2.6. *[3] A graph sequence $(G_n)_{n \geq 1}$ (a 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.*

A remarkable feature of the $\delta_\square$ distance is that the deviation of a sampled graph from the original graph or graphon can be upper bounded by a function that decreases logarithmically in the inverse of the sample size. Originally, this result was established to verify Theorem 2.6.

Lemma 2.7. *[3] Let $\varepsilon > 0$ and let U be a graphon. Then for $q \geq 2^{100/\varepsilon^2}$ we have*

$$\mathbb{P}\left(\delta_\square(U, \mathbb{G}(q, U)) \geq \varepsilon\right) \leq \exp\left(-4^{100/\varepsilon^2} \frac{\varepsilon^2}{50}\right). \tag{2.2}$$

We turn our attention to the continuous formulation of the Regularity Lemma of Frieze and Kannan [6] in the graphon space. For a partition $\mathcal{P}$ of $[0, 1]$ and a kernel W we obtain $W_\mathcal{P}$ from W by averaging on every rectangle given by product sets from $\mathcal{P}$.

Lemma 2.8 (Weak Regularity Lemma for kernels). *[6], [10] For every $\varepsilon > 0$ and $W \in \mathcal{W}_0$ there exists a partition $\mathcal{P} = (P_1, \ldots, P_m)$ of $[0, 1]$ into $m \leq 2^{\frac{8}{\varepsilon^2}}$ parts, such that*

$$\|W - W_\mathcal{P}\|_\square \leq \varepsilon \|W\|_2. \tag{2.3}$$

By slight adaptation of the original proof in [6] of the above result we obtain the version for k-colored digraphons.

7

Lemma 2.9 (Weak Regularity Lemma for k-colored directed graphons)**.** *For every $\varepsilon > 0$ and k-colored digraphon $\mathbf{W}$ there exists a partition $\mathcal{P} = (P_1, \ldots, P_m)$ of $[0,1]$ into $m \leq 2^{k^4 \frac{8}{\varepsilon^2}} = t'_k(\varepsilon)$ parts, such that*

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

When $\mathbf{W} = \mathbf{W_G}$ for a k-colored digraph $\mathbf{G}$ with vertex cardinality n, then one can require in the above statement that $\mathcal{P}$ is an $\mathcal{I}_n$-partition.

The following kernel norm shares some useful properties with the cut-norm. Most prominently it admits a regularity lemma that outputs a partition whose number of classes is considerably below the tower-type magnitude in the desired accuracy. On the other hand, it does not admit a straight-forward definition of a related distance by calculating the norm of the difference of two optimally overlayed objects as in Definition 2.4. This is the result of the general assumption that the partition $\mathcal{P}$ below involved in the definition always belongs to one of the graphons whose deviation we wish to estimate. Therefore a relabeling of this graphon should also act on $\mathcal{P}$, hence symmetry fails. Its advantages in comparison to the cut norm will become clearer in the proof of the main result below.

Definition 2.10. *Let W be a kernel and $\mathcal{P} = (P_1, \ldots, P_t)$ 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 kernels U and W let $d_{W,\mathcal{P}}(U)$ denote the cut-$\mathcal{P}$-deviation 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 on $[n]$ 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 $\mathbf{W} = (W^{(1,1)}, \ldots, W^{(k,k)})$ be a $k \times k$ tuple of kernels and $\mathcal{P} = (P_1, \ldots, P_t)$ a partition of $[0,1]$. Then the cut-$\mathcal{P}$-norm of $\mathbf{W}$ is*

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

8

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

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

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 in fact a norm on the space where we identify two kernels when they differ only on a set of measure 0. From the definition it follows directly that for arbitrary kernels U and W, and any partition $\mathcal{P}$ we have $\|W\|_{\square} \leq \|W\|_{\square\mathcal{P}} \leq \|W\|_1$ and $\delta_{\square}(U, W) \leq d_{W,\mathcal{P}}(U) \leq \delta_1(U, W)$, the same is true for the k-colored directed version.

Remark 2.12. We present a different description of the cut-$\mathcal{P}$-norm of W and $\mathbf{W}$ respectively that will allow us to rely on results concerning the cut-norm of Definition 2.4 more directly. For a partition $\mathcal{P}$ with t classes and $A = (A_{j,l})_{j,l=1}^{t} \in \{-1, +1\}^{t \times t}$, let $W^A(x, y) = A_{j,l}W(x, y)$ ($\mathbf{W}^A$ is given by $(W^{(\alpha,\beta)})^A(x, y) = A_{j,l}W^{(\alpha,\beta)}(x, y)$ respectively) for $x \in P_j$ and $y \in P_l$. Then $\|W\|_{\square\mathcal{P}} = \max_{A} \|W^A\|_{\square}$ and $\|\mathbf{W}\|_{\square\mathcal{P}} = \max_{A} \|\mathbf{W}^A\|_{\square}$.

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

Lemma 2.13. *For every* $\varepsilon > 0$, $m_0 \colon [0,1] \to \mathbb{N}$, $k \geq 1$ *and k-colored directed graphon* $\mathbf{W} = (W^{(\alpha,\beta)})_{\alpha,\beta \in [k]}$ *there exists a partition* $\mathcal{P} = (P_1, \ldots, P_t)$ *of* $[0,1]$ *into* $t \leq \frac{(16m_0(\varepsilon))^{2\frac{k^4}{\varepsilon^2}}}{4} = t_k(\varepsilon, m_0(\varepsilon))$ *parts, such that for any partition* $\mathcal{Q}$ *of* $[0,1]$ *into at most* $\max\{t, m_0(\varepsilon)\}$ *classes we have*

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{P}}\|_{\square\mathcal{Q}} \leq \varepsilon. \qquad (2.10)$$

If $\mathbf{W} = \mathbf{W}_{\mathbf{G}}$ *for some k-colored* $\mathbf{G}$ *with* $|V(\mathbf{G})| = n$, *then one can require that* $\mathcal{P}$ *is an* $\mathcal{I}_n$-*partition. If we want the parts to have equal measure (almost equal in the graph case), then the upper bound on the number of classes is modified to* $\frac{((3k)^{12}m_0(\varepsilon)/\varepsilon^4)^{2^{2k^4/\varepsilon^2}}}{(3k)^6/\varepsilon^2}$.

Proof. Fix an arbitrary function m_0, a $\varepsilon > 0$, and a $\mathbf{W} \in \mathcal{W}^{(k)}$. We construct a sequence of partitions $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_m$ such that $\mathcal{R}_0 = [0,1]$ and each $\mathcal{R}_{i+1}$ refines the preceding $\mathcal{R}_i$. The integer m is a priori undefined.

The construction is sequential in the sense that we assume that we have already constructed $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_{i-1}$ before considering the ith step of the construction.

If for $i \geq 1$ there exists a partition $\mathcal{Q} = (Q_1, \ldots, Q_{t_{\mathcal{Q}}})$ of $[0,1]$ into at most $\max\{t_{\mathcal{R}_{i-1}}, m_0(\varepsilon)\}$ classes such that

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{R}_{i-1}}\|_{\square\mathcal{Q}} > \varepsilon, \qquad (2.11)$$

then we proceed to the construction of $\mathcal{R}_i$. In the case of $i = 1$ we choose $\mathcal{Q}$ to have exactly m_0 parts of positive measure, this can be achieved since for any refinement $\mathcal{Q}'$ of $\mathcal{Q}$ we have $\|\mathbf{W} - \mathbf{W}_{\mathcal{R}_{i-1}}\|_{\square \mathcal{Q}'} \geq \|\mathbf{W} - \mathbf{W}_{\mathcal{R}_{i-1}}\|_{\square \mathcal{Q}}$. The inequality (2.11) implies that there are $\alpha_0, \beta_0 \in [k]$ and measurable subsets S and T of $[0, 1]$ such that

$$\sum_{i,j=1}^{t_\mathcal{Q}} \left| \int_{(S \cap Q_i) \times (T \cap Q_j)} W^{(\alpha_0, \beta_0)}(x, y) - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}(x, y) \mathrm{d}x \mathrm{d}y \right| > \varepsilon / k^2. \tag{2.12}$$

In this case we define $\mathcal{R}_i$ to be the coarsest common refinement of $\mathcal{R}_{i-1}$, $\mathcal{Q}$, and $\{S, T\}$ for some arbitrary choice of the latter partition and sets satisfying (2.12).

Set $S_j = S \cap Q_j$ and $T_j = T \cap Q_j$ for $j \in [t_\mathcal{Q}]$ and $U = W^{(\alpha_0, \beta_0)} - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}$, further, define the step function $V = \sum_{j, l \in [t_\mathcal{Q}]} \mathrm{sgn}(\int_{S_j \times T_l} U) \mathbb{I}_{S_j \times T_l}$

In this case

$$\|\mathbf{W}_{\mathcal{R}_i}\|_2^2 - \|\mathbf{W}_{\mathcal{R}_{i-1}}\|_2^2 = \sum_{\alpha, \beta = 1}^{k} \langle W^{(\alpha, \beta)}_{\mathcal{R}_i}, W^{(\alpha, \beta)}_{\mathcal{R}_i} \rangle - \langle W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}, W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}} \rangle$$

$$= \sum_{\alpha, \beta = 1}^{k} \langle W^{(\alpha, \beta)}_{\mathcal{R}_i} - W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}, W^{(\alpha, \beta)}_{\mathcal{R}_i} - W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}} \rangle \tag{2.13}$$

$$= \sum_{\alpha, \beta = 1}^{k} \|W^{(\alpha, \beta)}_{\mathcal{R}_i} - W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}\|_2^2$$

$$\geq \|W^{(\alpha_0, \beta_0)}_{\mathcal{R}_i} - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}\|_2^2$$

$$\geq \frac{1}{\|V\|_2^2} |\langle W^{(\alpha_0, \beta_0)}_{\mathcal{R}_i} - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}, V \rangle|^2 \tag{2.14}$$

$$= \frac{1}{\|V\|_2^2} |\langle W^{(\alpha_0, \beta_0)} - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}, V \rangle|^2 \tag{2.15}$$

$$\geq |\langle W^{(\alpha_0, \beta_0)} - W^{(\alpha_0, \beta_0)}_{\mathcal{R}_{i-1}}, V \rangle|^2 \tag{2.16}$$

$$> \varepsilon^2 / k^4. \tag{2.17}$$

Here we used first in (2.13) that $\langle W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}, W^{(\alpha, \beta)}_{\mathcal{R}_i} \rangle = \langle W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}, W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}} \rangle$, since $W^{(\alpha, \beta)}_{\mathcal{R}_{i-1}}$ is constant on $\mathcal{R}_{i-1}$ rectangles, and the integral of the two functions is equal on these rectangles. In (2.14) we used the Cauchy-Schwarz inequality, then in (2.15) the fact that $\langle W^{(\alpha_0, \beta_0)}_{\mathcal{R}_i}, V \rangle = \langle W^{(\alpha_0, \beta_0)}, V \rangle$, that is true by V being constant on $\mathcal{R}_i$ rectangles and $W^{(\alpha_0, \beta_0)}_{\mathcal{R}_i}$ and $W^{(\alpha_0, \beta_0)}$ having the same integral value on taken on $\mathcal{R}_i$ rectangles. We concluded the calculation in (2.16) by $\|V\|_2 \leq 1$ and in (2.17) using the condition (2.12).

If for some $i_0 \geq 0$ we have

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{R}_{i_0}}\|_{\square \mathcal{Q}} \leq \varepsilon \tag{2.18}$$

for every partition $\mathcal{Q}$ of $[0,1]$ into at most $\max\{t_{\mathcal{R}_{i_0}}, m_0(\varepsilon)\}$ classes, then we stop the process and set $\mathcal{P} = \mathcal{R}_{i_0}$ and $m = i_0$.

We have $\|\mathbf{W}_{\mathcal{R}_j}\|_2^2 \le \|\mathbf{W}\|_2^2 \le \|\mathbf{W}\|_1^2 \le 1$ for each $j \ge 0$, and at each non-terminating step we showed $\|\mathbf{W}_{\mathcal{R}_i}\|_2^2 - \|\mathbf{W}_{\mathcal{R}_{i-1}}\|_2^2 > \varepsilon^2/k^4$. Therefore by

$$\|\mathbf{W}_{\mathcal{R}_j}\|_2^2 \ge \sum_{i=1}^{j} \|\mathbf{W}_{\mathcal{R}_i}\|_2^2 - \|\mathbf{W}_{\mathcal{R}_{i-1}}\|_2^2,$$

for each $j \ge 1$ we conclude that the process terminates definitely after a finite number of steps and $m \le k^4/\varepsilon^2$. The partition $\mathcal{P}$ satisfies

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{P}}\|_{\square\mathcal{Q}} \le \varepsilon$$

for each $\mathcal{Q}$ with $t_{\mathcal{Q}} \le \max\{t, m_0(\varepsilon)\}$ by the choice of m and the construction of the partition sequence, we are left to verify the upper bound on $t_{\mathcal{P}}$ in the statement of the lemma.

We know that $t_{\mathcal{R}_0} = 1$, and if the partition does not terminate before the first step, then we assume $m_0 \le t_{\mathcal{R}_1}$. This lower bound does not affect generality, the partition $\mathcal{Q}_0$ that certifies that $\mathcal{R}_0$ is not suitable for the choice of the partition $\mathcal{P}$ in the statement of the lemma is selected to have $t_{\mathcal{Q}_0} = m_0$. For this particular choice of $\mathcal{Q}_0$ and $\mathcal{R}_1$ we can reformulate the condition that the terminating partition $\mathcal{R}_m$ has to fulfill as

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{R}_m}\|_{\square\mathcal{Q}} \le \varepsilon$$

for every $\mathcal{Q}$ partition of $[0,1]$ into at most $t_{\mathcal{R}_m}$ classes, since $t_{\mathcal{R}_i} \ge m_0$ for every $i \ge 1$.

We set $s(0) = 1$, $s(1) = 4m_0$, and further define $s(i+1) = 4s(i)^2$ for each $i \ge 1$. We claim that for each $i \ge 0$ we have $t_{\mathcal{R}_i} \le s(i)$, this can be easily verified by induction, since at each step $\mathcal{R}_{i+1}$ is the coarsest common refinement of two partitions with $t_{\mathcal{R}_i}$ classes and two additional sets.

Further, for each $i \ge 1$ we have now $\log 4s(i+1) = 2\log 4s(i)$, therefore $s(i) = \frac{(16m_0)^{2^{i-1}}}{4}$, and consequently $s(m) \le \frac{(16m_0)^{2^{\frac{k^4}{\varepsilon^2}-1}}}{4}$.

The case regarding $\mathbf{W} = \mathbf{W_G}$ for a k-colored directed graph $\mathbf{G}$ of vertex cardinality n follows completely identically, at each step of the construction of the partitions $\mathcal{R}_i$ the partition $\mathcal{Q}_0$ and the sets S and T can be chosen to be an $\mathcal{I}_n$-partition and sets, respectively. Hence, $\mathcal{P}$ is an $\mathcal{I}_n$-partition, the upper bound on $t_{\mathcal{P}}$ is identical to the one in the general case.

In a similar way we can achieve that $\mathcal{P}$ is an equiv-partition, or a $\mathcal{I}_n$-partition with classes of almost equal size in the graph case respectively. Fix $\varepsilon > 0$ and a $\mathbf{W} \in \mathcal{W}^{(k)}$. For this setup we define the partition sequence somewhat differently, in particular each element is an equiv-partition. Set $\mathcal{R}_0 = [0,1]$ and each $\mathcal{R}_{i+1}$ refines the preceding $\mathcal{R}_i$.

If for $i \ge 1$ there exists a partition $\mathcal{Q} = (Q_1, \dots, Q_{t_{\mathcal{Q}}})$ of $[0,1]$ into at most $\max\{t_{\mathcal{R}_{i-1}}, m_0\}$ classes such that

$$\|\mathbf{W} - \mathbf{W}_{\mathcal{R}_{i-1}}\|_{\square\mathcal{Q}} > \varepsilon, \tag{2.19}$$

11

then we proceed to the construction of $\mathcal{R}_i$, otherwise we stop, as above, and set $\mathcal{P} = \mathcal{R}_{i-1}$, and $m = i - 1$. Assume that we are facing the first case. Let $\mathcal{R}'_{i-1}$ be the coarsest common refinement of $\mathcal{R}_{i-1}$, $\mathcal{Q}$, and $\{S, T\}$, where the sets S and T certify (2.19) as above. Then $\|\mathbf{W}_{\mathcal{R}'_{i-1}}\|_2^2 - \|\mathbf{W}_{\mathcal{R}_{i-1}}\|_2^2 > \varepsilon^2/k^4$. Let $\mathcal{R}''_{i-1} = (R_1^2, \ldots, R_l^2)$ be the partition that is obtained from the classes of $\mathcal{R}'_{i-1} = (R_1^1, \ldots, R_l^1)$, such that the measure of each of the classes of $\mathcal{R}''_{i-1}$ is an integer multiple of $\varepsilon^2/(14k^6 t_{\mathcal{R}'_{i-1}})$ with $\lambda(R_i^1 \triangle R_i^2) \leq \varepsilon^2/(14k^6 t_{\mathcal{R}'_{i-1}})$ for each $i \in [l]$. (We disregard the technical difficulty of $1/\varepsilon^2$ not being an integer to facilitate readability.)

Claim 2.14. For any kernel $W \colon [0,1]^2 \to \mathbb{R}$ and partitions $\mathcal{P} = (P_1, \ldots, P_t)$ and $\mathcal{S} = (S_1, \ldots, S_l)$ we have

$$\|W_{\mathcal{P}} - W_{\mathcal{S}}\|_1 \leq 7 \sum_{i=1}^{l} \lambda(P_i \triangle S_i). \tag{2.20}$$

To see this, let $T_i = P_i \cap S_i$, $N_i = P_i \setminus S_i$, and $M_i = S_i \setminus P_i$ for each $i \in [l]$, and let $\mathcal{T}_1 = (T_1, \ldots, T_l, N_1, \ldots, N_l)$ and $\mathcal{T}_2 = (T_1, \ldots, T_l, M_1, \ldots, M_l)$ be two partitions of $[0,1]$. Then

$$\|W_{\mathcal{P}} - W_{\mathcal{T}_1}\|_1 \leq \sum_{i,j=1}^{l} \left| \int_{T_i \times T_j} \frac{\int_{P_i \times P_j} W}{\lambda(P_i)\lambda(P_j)} - \frac{\int_{T_i \times T_j} W}{\lambda(T_i)\lambda(T_j)} \right| + 2\sum_{i=1}^{l} \lambda(N_i)$$

$$\leq \sum_{i,j=1}^{l} \frac{1}{\lambda(P_i)\lambda(P_j)\lambda(T_i)\lambda(T_j)} \left| \int_{T_i \times T_j} \left[\lambda(T_i)\lambda(T_j) \left(\int_{N_i \times T_j} W + \int_{T_i \times N_j} W + \int_{N_i \times N_j} W \right) \right. \right.$$

$$\left. \left. - (\lambda(N_i)\lambda(T_j) + \lambda(T_i)\lambda(N_j) + \lambda(N_i)\lambda(N_j)) \int_{T_i \times T_j} W \right] \right| + 2\sum_{i=1}^{l} \lambda(N_i)$$

$$\leq \sum_{i,j=1}^{l} 2\|W\|_\infty \frac{\lambda^2(T_i)\lambda^2(T_j)\left[\lambda(N_i)\lambda(T_j) + \lambda(T_i)\lambda(N_j) + \lambda(N_i)\lambda(N_j)\right]}{\lambda(P_i)\lambda(P_j)\lambda(T_i)\lambda(T_j)}$$

$$+ 2\sum_{i=1}^{l} \lambda(N_i)$$

$$\leq 2\sum_{i,j=1}^{l} \lambda(N_i)\lambda(P_j) + \lambda(N_j)\lambda(P_i) + 2\sum_{i=1}^{l} \lambda(N_i)$$

$$= 6\sum_{i=1}^{l} \lambda(N_i).$$

Similarly,

$$\|W_{\mathcal{S}} - W_{\mathcal{T}_2}\|_1 \leq 6\sum_{i=1}^{l} \lambda(M_i),$$

12

and also

$$\|W_{\mathcal{T}_2} - W_{\mathcal{T}_1}\|_1 \le 2\sum_{i=1}^{l} \lambda(M_i),$$

which implies the claim.

By Claim 2.14 it follows that

$$\left|\|\mathbf{W}_{\mathcal{R}'_{i-1}}\|_2^2 - \|\mathbf{W}_{\mathcal{R}''_{i-1}}\|_2^2\right| = \left|\sum_{\alpha,\beta=1}^{k} \int_{[0,1]^2} (W^{(\alpha,\beta)}_{\mathcal{R}'_{i-1}})^2(x,y) - (W^{(\alpha,\beta)}_{\mathcal{R}''_{i-1}})^2(x,y)\mathrm{d}x\mathrm{d}y\right|$$

$$\le \sum_{\alpha,\beta=1}^{k} \left|\int_{[0,1]^2} (W^{(\alpha,\beta)}_{\mathcal{R}'_{i-1}}(x,y) - W^{(\alpha,\beta)}_{\mathcal{R}''_{i-1}}(x,y))(W^{(\alpha,\beta)}_{\mathcal{R}'_{i-1}}(x,y) + W^{(\alpha,\beta)}_{\mathcal{R}''_{i-1}}(x,y))\mathrm{d}x\mathrm{d}y\right|$$

$$\le \sum_{2} \|\mathbf{W}\|_\infty \alpha, \beta = 1^k \|W^{(\alpha,\beta)}_{\mathcal{R}'_{i-1}} - W^{(\alpha,\beta)}_{\mathcal{R}''_{i-1}}\|_1$$

$$\le 28k^2 \sum_{i=1}^{l} \lambda(R_i^1 \triangle R_i^2)$$

$$\le \varepsilon^2/(2k^4).$$

We finish with the construction of $\mathcal{R}_i$ by refining $\mathcal{R}''_{i-1}$ into $\varepsilon^2/(28k^6 t_{\mathcal{R}'_{i-1}})$ sets in total of equal measure so that the resulting partition refines $\mathcal{R}''_{i-1}$. It follows that

$$\|\mathbf{W}_{\mathcal{R}_i}\|_2^2 - \|\mathbf{W}_{\mathcal{R}''_{i-1}}\|_2^2 \ge 0, \tag{2.21}$$

hence

$$\|\mathbf{W}_{\mathcal{R}_i}\|_2^2 - \|\mathbf{W}_{\mathcal{R}_{i-1}}\|_2^2 \ge \varepsilon^2/(2k^4). \tag{2.22}$$

The construction of the partitions terminates after at most $2k^4/\varepsilon^2$ steps. The partition $\mathcal{P}$ satisfies the norm conditions of the lemma, we are left to check whether it has the right number of classes. Similarly as above, let $s(0) = 1$ and $s(1) = m_0(3k)^6/\varepsilon^2$, and further for $i \ge 1$ let $s(i+1) = (3k)^6 s^2(i)/\varepsilon^2$. It is clear from the construction that $t_{\mathcal{R}_i} \le s(i)$. Let $a = (3k)^6/\varepsilon^2$, then it is not difficult to see that $s(i) = \frac{(a^2 m_0)^{2^{i-1}}}{a}$. It follows that $t_{\mathcal{P}} \le \frac{((3k)^{12} m_0/\varepsilon^4)^{2^{2k^4/\varepsilon^2}}}{(3k)^6/\varepsilon^2}$.

The graph case also follows analogously to the general partition case we dealt with above.

$\square$

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 and graphon case the above reads as follows.

13

Corollary 2.15. *For every $\varepsilon > 0$ and $W \in \mathcal{W}$ there exists a partition $\mathcal{P} = (P_1, \ldots, P_m)$ of $[0,1]$ into $m \leq 16^{2^{1/\varepsilon^2}}/4$ parts, such that*

$$\|W - W_{\mathcal{P}}\|_{\square \mathcal{Q}} \leq \varepsilon. \tag{2.23}$$

for each partition $\mathcal{Q}$ of $[0,1]$ into at most $t_{\mathcal{P}}$ classes.

    With the additional condition that the partition classes should have the same measure the above is true with $m \leq \frac{(3^{12}/\varepsilon^4)^{2^{(2^4/\varepsilon^2)}}}{3^6/\varepsilon^2}$.

3   Proof of Theorem 1.5

We will exploit the continuity of a testable graph parameter with respect to the cut norm and distance, and the connection of this characteristic to the sample complexity of the parameter. We require two results, the first one quantifies the above continuity. We generally assume that the sample complexity satisfies $q_g(\varepsilon) \geq 1/\varepsilon$, also $\varepsilon \leq 1$ and $k \geq 2$.

Lemma 3.1. *Let g be a testable k-colored digraph parameter with sample complexity at most q_g. Then for any $\varepsilon > 0$ and two graphs, $\mathbf{G}$ and $\mathbf{H}$, with $|V(\mathbf{G})|, |V(\mathbf{H})| \geq \left(\frac{2q_g^2(\varepsilon/4)}{\varepsilon} \right)^{1/(q_g(\varepsilon/4)-1)}$ satisfying $\delta_{\square}(\mathbf{G}, \mathbf{H}) \leq k^{-2q_g^2(\varepsilon/4)}$ we have*

$$|g(\mathbf{G}) - g(\mathbf{H})| \leq \varepsilon.$$

Proof. Let $\varepsilon > 0$, $\mathbf{G}$ and $\mathbf{H}$ be as in the statement, and set $q = q_g(\varepsilon/4)$. Then we have

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

The first and the last term on the right of (3.1) can be each upper bounded by $\varepsilon/4$ with cumulative failure probability $\varepsilon/2$ due to the assumptions of the lemma. To deal with the second term we require the fact that $\mathbb{G}(q, \mathbf{G})$ and $\mathbb{G}(q, \mathbf{W_G})$ have the same distribution conditioned on the event that the X_i variables that define $\mathbb{G}(q, \mathbf{W_G})$ 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 $q^2/2|V(\mathbf{G})|^{q-1}$, which is at most $\varepsilon/4$, analogously for the fourth term. Until this point we have not dealt with the relationship of the two random objects $\mathbb{G}(q, \mathbf{W_G})$ and $\mathbb{G}(q, \mathbf{W_H})$, therefore the above discussion is valid for every coupling of them.

    In order to handle the third term we upper bound the probability that the two random graphs are different by means of an appropriate coupling, since clearly in the event of identity the third term of (3.1) vanishes. 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 $\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) \neq \mathbb{G}(q, \mathbf{W_H})) < 1 - \varepsilon$. 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 $\mathbf{H}$ through the cut distance of these graphs, see Lemma 2.5. In particular,

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

Therefore in our case

$$\sum_{\mathbf{F}} |\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) = \mathbf{F}) - \mathbb{P}(\mathbb{G}(q, \mathbf{W_H}) = \mathbf{F})| \leq k^{2\binom{q}{2}} \binom{q}{2} k^{-2q^2} \leq \varepsilon,$$

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

Since there are only finitely many possible target graphs for the random objects, we can couple $\mathbb{G}(q, \mathbf{W_G})$ and $\mathbb{G}(q, \mathbf{W_H})$ so that in the end we have $\mathbb{P}(\mathbb{G}(q, \mathbf{W_G}) \neq \mathbb{G}(q, \mathbf{W_H})) \leq \varepsilon$. This implies that with positive probability (in fact, with at least $1 - 2\varepsilon$) 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 [12]. It clarifies why the cut-$\mathcal{P}$-norm, (Definition 2.10, Definition 2.11) and the need for the accompanying regularity lemma, Lemma 2.13, are essential for our intent.

Lemma 3.2. *Let $k \geq 2$, $\varepsilon > 0$, U be a step function with steps $\mathcal{P} = (P_1, \ldots, P_t)$ and V be a graphon with $\|U - V\|_{\square\mathcal{P}} \leq \varepsilon$. For any k-colored digraphon $\mathbf{U} = (U^{(1,1)}, \ldots, U^{(k,k)})$ that is a step function with steps from $\mathcal{P}$ and a (k, m)-coloring of U there exists a (k, m)-coloring $\mathbf{V} = (V^{(1,1)}, \ldots, V^{(k,k)})$ of V so that $\|\mathbf{U} - \mathbf{V}\|_{\square} = \sum_{\alpha,\beta=1}^{k} \|U^{(\alpha,\beta)} - V^{(\alpha,\beta)}\|_{\square} \leq k^2\varepsilon$.*

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

Proof. Fix $\varepsilon > 0$, and let U, V, and $\mathbf{U}$ be 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 that is at most m, so we have $\sum_{(\alpha,\beta) \in M} U^{(\alpha,\beta)} = U$ by definition. 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 $U < 1$ 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 $\|U^{(\alpha,\beta)} - V^{(\alpha,\beta)}\|_{\square} \leq \varepsilon$. We will explicitly perform the calculation only for $(\alpha, \beta) \in M$, the other case is analogous. Fix some $S, T \subset [0, 1]$, then

$$\left| \int_{S \times T} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} \right| = \left| \int_{S \times T, U > 0} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} + \int_{S \times T, U = 0} U^{(\alpha,\beta)} - V^{(\alpha,\beta)} \right|$$

15

$$\leq \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j),U>0} \frac{U^{(\alpha,\beta)}}{U}(U-V) + \int_{(S\cap P_i)\times(T\cap P_j),U=0} \frac{1}{k^2-(k-m)^2}(U-V) \right|$$

$$= \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j)} (U-V)\left[\mathbb{1}_{U>0}\frac{U^{(\alpha,\beta)}}{U} + \mathbb{1}_{U=0}\frac{1}{k^2-(k-m)^2} \right] \right|$$

$$\leq \sum_{i,j=1}^{t} \left| \int_{(S\cap P_i)\times(T\cap P_j)} (U-V) \right|$$

$$= \|U-V\|_{\square\mathcal{P}} \leq \varepsilon.$$

The second inequality is a consequence of $\left[\mathbb{1}_{U>0}\frac{U^{(\alpha,\beta)}}{U} + \mathbb{1}_{U=0}\frac{1}{k^2-(k-m)^2} \right]$ 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 partition $\mathcal{P}$ that is an $\mathcal{I}_n$-partition. The general discussion above 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 $\mathbb{P}\left(Z_{ij} = (\alpha,\beta)\right) = n^2 \int_{[\frac{i-1}{n},\frac{i}{n}]\times[\frac{j-1}{n},\frac{j}{n}]} V^{(\alpha,\beta)}(x,y)\mathrm{d}x\mathrm{d}y$. 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. It remains to conduct the analysis of the deviation in the statement of the lemma, we will show that this is small with high probability with respect to the randomization, which in turn implies the existence. We have

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

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

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 [3]. This implies for $n \geq 16/\varepsilon^2$ the existence of a suitable coloring, which in turn finishes the proof of the lemma.

$\square$

Remark 3.3. Actually we can perform the same proof to verify the existence of a k-coloring $\mathbf{V}$ such that $d_{\square\mathcal{P}}(\mathbf{U},\mathbf{V}) \leq k^2\varepsilon$. 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 $\mathbf{V}$, for example in the case when the number of steps of $\mathbf{U}$ is exponential in $1/\|U-V\|_\square$.

We proceed towards the proof of the main statement of the paper. Before we can outline that we require yet another specific lemma.

Let $\mathcal{M}_{\Delta,n}$ denote the set of $\mathcal{I}_n$-step functions U that have steps $\mathcal{P}_U$ with $|\mathcal{P}_U| \leq t_k(\Delta, 1)$ classes, and values between 0 and 1, where t_k is the function from Lemma 2.13. In order to verify Theorem 1.5 we will condition on the event that is formulated in the following lemma. Recall Definition 2.10 for the deviation $d_{W,\mathcal{P}}(V)$.

Lemma 3.4. *Let G be a simple graph on n vertices and $\Delta > 0$. Then for $q \geq 2^{2^{(2k^4/\Delta^2)+4}}$ we have*

$$|d_{U,\mathcal{P}_U}(G) - d_{U,\mathcal{P}_U}(\mathbb{G}(q,G))| \leq \Delta, \tag{3.2}$$

for each $U \in \mathcal{M} = \mathcal{M}_{\Delta,n}$ simultaneously with probability at least $1 - \exp(-\frac{\Delta^2 q}{2^7})$, whenever $n \geq 4q/\Delta$.

Proof. Let G and $\Delta > 0$ be arbitrary, and q be such that it satisfies the conditions of the lemma. For technical convenience we assume that n is an integer multiple of q, let us introduce the quantity $t_1 = t_k(\Delta, 1) = 2^{2^{\frac{k^4}{\Delta^2}+2}-2}$, and denote $\mathbb{G}(q,G)$ by F. For the case when q is not a divisor of n, then we just add at most q isolated vertices to G to achieve the above condition, by this operation $d_{U,\mathcal{P}}(G)$ is changed by at most q/n. Also, we can couple in a way such that $d_{U,\mathcal{P}}(\mathbb{G}(q,G))$ remains unchanged with probability at least $1 - q/n$.

We will show that there exists an $\mathcal{I}_n$-permutation ϕ of $[0,1]$ such that $\|W_G - W_F^\phi\|_{\square\mathcal{Q}} < \Delta$ for any $\mathcal{I}_n$-partition $\mathcal{Q}$ of $[0,1]$ into at most t_1 classes with high probability. Applying Lemma 2.13 with the error parameter $\Delta/4$ and $m_0(\Delta) = t_k(\Delta, 1)$ for approximating W_G by a step function we can assert that there exists an $\mathcal{I}_n$-partition $\mathcal{P}$ of $[0,1]$ into $t_\mathcal{P}$ classes with $t_\mathcal{P} \leq t_2$ with $t_2 = t_k(\Delta/4, t_k(\Delta, 1)) = 2^{2^{(2k^4/\Delta^2)+2}+2^{(k^4/\Delta^2)+1}} \leq 2^{2^{(2k^4/\Delta^2)+3}}$ such that for every $\mathcal{I}_n$-partition $\mathcal{Q}$ into $t_\mathcal{Q}$ classes $t_\mathcal{Q} \leq \max\{t_\mathcal{P}, t_1\}$ it holds that

$$\|W_G - (W_G)_\mathcal{P}\|_{\square\mathcal{Q}} \leq \Delta/4.$$

We only need here

$$\sup_{\mathcal{Q}:t_\mathcal{Q}\leq t_1} \|W_G - (W_G)_\mathcal{P}\|_{\square\mathcal{Q}} \leq \Delta/4. \tag{3.3}$$

This property is by Remark 2.12 equivalent to stating that

$$\max_\mathcal{Q} \max_{A\in\mathbb{A}} \max_{S,T\subset[0,1]} \sum_{i,j=1}^{t_1} A_{i,j} \int_{S\times T} (W_G - (W_G)_\mathcal{P})(x,y)\mathbb{I}_{Q_i}(x)\mathbb{I}_{Q_j}(y)\mathrm{d}x\mathrm{d}y \leq \Delta/4, \tag{3.4}$$

where $\mathbb{A}$ is the set of all $t_1 \times t_1$ matrices with -1 or $+1$ entries.

We can reformulate the above expression (3.4) by putting

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

17

and defining the tensor product $B_A = A \otimes J$, so that $B_{i,j}^{\alpha,\beta} = A_{ij} J_{\alpha,\beta}$ for each $A \in \mathbb{A}$. The first matrix J corresponds to the $\mathcal{I}_n$-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 $\mathcal{I}_n$-sets of $[0,1]$ so that for any function $U \colon [0,1]^2 \to \mathbb{R}$ it holds that

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

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

$$\max_{A \in \mathbb{A}} \max_{\hat{Q}} \sum_{i,j=1}^{t_1} \sum_{\alpha,\beta=1}^{4} (B_A)_{i,j}^{\alpha,\beta} \int_{[0,1]^2} (W_G - (W_G)_{\mathcal{P}})(x,y) \mathbb{I}_{Q_i^\alpha}(x) \mathbb{I}_{Q_j^\beta}(y) \mathrm{d}x\mathrm{d}y \leq \Delta/4, \qquad (3.5)$$

where the second maximum goes over all $\mathcal{I}_n$-partitions $\hat{Q} = (Q_i^\alpha)_{\substack{i \in [t_1] \\ \alpha \in [4]}}$ into $4t_1$ classes. Let us substitute an arbitrary graphon U for $W_G - (W_G)_{\mathcal{P}}$ in (3.5) and define

$$\hat{h}_{A,\hat{Q}}(U) = \sum_{\substack{1 \leq i,j \leq t_1 \\ 1 \leq \alpha,\beta \leq 4}} (B_A)_{i,j}^{\alpha,\beta} \int_{[0,1]^2} U(x,y) \mathbb{I}_{Q_i^\alpha}(x) \mathbb{I}_{Q_j^\beta}(y)\mathrm{d}x\mathrm{d}y$$

and

$$\hat{h}_A(U) = \max_{\hat{Q}} h_{A,\hat{Q}}(U)$$

as the expression whose optima is sought for a fixed $A \in \mathbb{A}$.

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

$$h_{A,f}(U) = \sum_{1 \leq i,j \leq 4t_1} (B_A)_{i,j} \int_{[0,1]^2} U(x,y) f_i(x) f_j(y)\mathrm{d}x\mathrm{d}y$$

with $f = (f_i)_{i \in [4t_1]}$ being a fractional $\mathcal{I}_n$-partition into $4t_1$ classes, that is, each component of f is a non-negative $\mathcal{I}_n$-function, and their sum is the constant 1 function. Further, we define $h_A(U) = \max_f h_{A,f}(U)$, where f runs over all fractional $\mathcal{I}_n$-partitions into $4t_1$ parts. It is easy to see that

$$|\hat{h}_A(U) - h_A(U)| \leq 1/n,$$

since the two functions coincide when U is 0 on the diagonal blocks. Denote $U' = W_{\mathbb{H}(q,U)}$, where the graphon is given by the increasing order of the sample points $\{ X_i \mid i \in [q] \}$. We

wish to upper bound the probability that the deviation $|h_A(U) - h_A(U')|$ exceeds $\Delta/4$, for some $A \in \mathbb{A}$. Similarly as above, $|\hat{h}_A(U') - h_A(U')| \leq 1/q$.

We remark, that a simple approach would be using a slight variant of the counting lemma Lemma 2.5 that $|h_A(U) - h_A(U')| \leq 16t_1^2\delta_\square(U, U')$ together with a version of Lemma 2.7 for kernels with perhaps negative values, this way we would have to impose a lower bound on q that is exponential in t_1 in order to satisfy the statement of the lemma. We can do slightly better using more involved methods.

We require the notion of ground state energies from [4]. Let

$$\hat{\mathcal{E}}(G, J) = \max_{\mathcal{Q}} \sum_{i,j=1}^{s} J_{i,j} \int_{[0,1]^2} \mathbb{I}_{Q_i}(x)\mathbb{I}_{Q_j}(y)W_G(x, y)\mathrm{d}x\mathrm{d}y,$$

where the maximum runs over all $\mathcal{I}_n$-partitions $\mathcal{Q}$ into s parts when $|V(G)| = n$.

Further,

$$\mathcal{E}(U, J) = \sup_{f} \sum_{i,j=1}^{s} J_{i,j} \int_{[0,1]^2} f_i(x)f_j(y)U(x, y)\mathrm{d}x\mathrm{d}y,$$

where the supremum runs over all fractional partitions f into s parts.

The next result was first proved in [1], subsequently refined in [8].

Theorem 3.5. *[1][8] Let $s \geq 1$, and $\rho > 0$. Then there is an absolute constant $c > 0$ such that for any $s \geq 1$, $\rho > 0$, kernel U, real matrix J, and $q \geq c\Theta^4 \log(\Theta)$ with $\Theta = \frac{s^2}{\rho}$ we have*

$$\mathbb{P}(|\mathcal{E}(U, J) - \hat{\mathcal{E}}(\mathbb{G}(q, U), J)| > \rho\|U\|_\infty) < 2\exp\left(-\frac{\rho^2 q}{32}\right). \tag{3.6}$$

We have seen above that $\hat{h}_A(U) = \hat{\mathcal{E}}(U, B_A)$ and $h_A(U) = \mathcal{E}(U, B_A)$. Since $q \geq t_2^2 \geq 2^{85}t_1^{10}/\Delta^5$ we can apply Theorem 3.5 for each $A \in \mathbb{A}$ with $s = 4t_1$, and $\rho = \Delta/4$.

This shows eventually that with probability at least $1 - 2^{16t_1^2+1}\exp\left(-\frac{\Delta^2 q}{2^9}\right) \leq 1 - \exp\left(-\frac{\Delta^2 q}{2^8}\right)$ we have

$$\max_{A \in \mathbb{A}} \max_{\hat{\mathcal{Q}}} \sum_{i,j=1}^{t_1} \sum_{\alpha,\beta=1}^{4} (B_A)_{i,j}^{\alpha,\beta} \int_{[0,1]^2} (W_{\mathbb{G}(q,U)})(x, y)\mathbb{I}_{Q_i^\alpha}(x)\mathbb{I}_{Q_j^\beta}(y)\mathrm{d}x\mathrm{d}y \leq \Delta/2, \tag{3.7}$$

where the second maximum runs over all $\mathcal{I}_q$-partitions $\hat{\mathcal{Q}}$ of $[0, 1]$ into $4t_1$ parts. Denote this event by E_1.

This however is equivalent to saying that for every $\mathcal{Q}$ partition into $t_{\mathcal{Q}}$ classes $t_{\mathcal{Q}} \leq t_1$ it is true that

$$\|W_{\mathbb{G}(q,U)}\|_{\square\mathcal{Q}} \leq \Delta/2. \tag{3.8}$$

The second estimate we require concerns the closeness of the step function $(W_G)_{\mathcal{P}}$ and its sample $W_{\mathbb{H}(q,(W_G)_{\mathcal{P}})}$. Our aim is to overlay these two functions via measure preserving

19

permutations of $[0,1]$, such that the measure of the subset of $[0,1]^2$ where they differ is as small as possible.

Let $V = W_{\mathbb{H}(q,(W_G)_\mathcal{P})}$, this $\mathcal{I}_n$-function is well-defined this way and is a step function with steps forming the $\mathcal{I}_n$-partition $\mathcal{P}'$. This latter $\mathcal{I}_n$-partition of $[0,1]$ is the image of $\mathcal{P}$ induced by the sample $\{X_1, \ldots, X_q\}$ and the map $i \mapsto [\frac{i-1}{q}, \frac{i}{q})$. Let ψ be a measure preserving $\mathcal{I}_n$-permutation of $[0,1]$ that satisfies that for each $i \in [t_\mathcal{P}]$ the volumes $\lambda(P_i \triangle \psi(P'_i)) = |\lambda(P_i) - \lambda(P'_i)|$. Let $\mathcal{P}''$ denote the partition with classes $P''_i = \psi(P'_i)$ and $V' = (V)^\psi$ (note that V' and V are equivalent as graphons), furthermore let N be the (random) subset of $[0,1]^2$ where the two functions $(W_G)_\mathcal{P}$ and V' differ. Then

$$\mathbb{E}[\lambda(N)] \leq 2\mathbb{E}[\sum_{i=1}^{t_\mathcal{P}} |\lambda(P_i) - \lambda(P'_i)|]. \tag{3.9}$$

The random variables $\lambda(P'_i)$ for each i can be interpreted as the proportion of positive outcomes out of q independent Bernoulli trials with success probability $\lambda(P_i)$. By Cauchy-Schwarz it follows that

$$\mathbb{E}[\sum_{i=1}^{t_\mathcal{P}} |\lambda(P_i) - \lambda(P'_i)|] \leq \sqrt{t_\mathcal{P} \mathbb{E}[\sum_{i=1}^{t_\mathcal{P}} (\lambda(P_i) - \lambda(P'_i))^2]} \leq \sqrt{\frac{t_2}{q}}. \tag{3.10}$$

This calculation yields that $\mathbb{E}[\lambda(N)] \leq \sqrt{\frac{4t_2}{q}} \leq \Delta/8$ by the choice of q, since $q \geq t_2^2$. Standard concentration result gives us that $\lambda(N)$ is also small in probability if q is chosen large enough. For convenience, define the martingale $M_l = \mathbb{E}[\lambda(N)|X_1, \ldots, X_l]$ for $1 \leq l \leq q$, and notice that the martingale differences are uniformly bounded, $|M_l - M_{l-1}| \leq \frac{4}{q}$. The Azuma-Hoeffding inequality then yields

$$\mathbb{P}(\lambda(N) \geq \Delta/4) \leq \mathbb{P}(\lambda(N) \geq \mathbb{E}[\lambda(N)] + \Delta/8) \leq \exp(-\Delta^2 q/2^{11}). \tag{3.11}$$

Define the event E_2 that holds whenever $\lambda(N) \leq \Delta/4$, and condition on E_1 and E_2, the failure probability of each one is at most $\exp(-\frac{\Delta^2 q}{2^{11}})$.

It follows that $\|V^\psi - (W_G)_\mathcal{P}\|_1 \leq \lambda(N) \leq \Delta/4$. Now employing the triangle inequality and the bound (3.8) we get for all $\mathcal{I}_n$-partitions $\mathcal{Q}$ into t parts that

$$\|W_G - (W_F)^\psi\|_{\square\mathcal{Q}} \leq \|W_G - (W_G)_\mathcal{P}\|_{\square\mathcal{Q}} + \|(W_G)_\mathcal{P} - V^\psi\|_1 + \|V^\psi - (W_F)^\psi\|_{\square\psi(\mathcal{Q})} \leq \Delta.$$

Now let $U \in \mathcal{M}_{\Delta,n}$ be arbitrary, and let $\mathcal{P}_U$ denote the partition consisting of the steps of U. Let ϕ be the $\mathcal{I}_n$-permutation of $[0,1]$ that is optimal in the sense that $d_{U,\mathcal{P}_U}(G) = \|U - (W_G)^\phi\|_{\square\mathcal{P}_U}$. Then

$$d_{U,\mathcal{P}_U}(G) - d_{U,\mathcal{P}_U}(F) \leq \|U - (W_G)^\phi\|_{\square\mathcal{P}_U} - \|U - (W_F)^{(\psi \circ \phi)}\|_{\square\mathcal{P}_U}$$
$$\leq \|W_G - (W_F)^\psi\|_{\square\phi^{-1}(\mathcal{P}_U)} \leq \Delta.$$

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 $U \in \mathcal{M}_{\Delta,n}$.

20

We conclude the proof with mentioning that the failure probability of the two events E_1 and E_2 taking place simultaneously is at most $\exp(-\frac{\Delta^2 q}{2^7})$.

$\square$

We are now ready to conduct the proof of the main result of the paper concerning graph parameters.

Proof of Theorem 1.5.

Let us fix $\varepsilon > 0$ and the simple graph G with n vertices. We introduce the error parameter $\Delta = \frac{k^{-2} q_g^2(\varepsilon/2)}{4k^2+1}$ and set $q \geq 2^{2^{(2k^4/\Delta^2)+4}}$. To establish the lower bound on $f(\mathbb{G}(q, G))$ not much effort is required: we pick a (k, m)-coloring $\mathbf{G}$ of G that certifies the value $f(G)$, that is, $g(\mathbf{G}) = f(G)$. Then the (k, m)-coloring of $\mathbf{F} = \mathbb{G}(q, \mathbf{G})$ of $\mathbb{G}(q, G)$ induced by $\mathbf{G}$ satisfies $g(\mathbf{F}) \geq g(\mathbf{G}) - \varepsilon/2$ with probability at least $1 - \varepsilon/2$ since $q \geq q_g(\varepsilon/2)$, due to the testability property of g, which in turn implies $f(\mathbb{G}(q, G)) \geq f(G) - \varepsilon/2$ with probability at least $1 - \varepsilon/2$.

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. Recall that $\mathcal{M}_{\Delta,n}$ denotes the set of the $[0, 1]$-valued proper $\mathcal{I}_n$-step functions that have at most $t_k(\Delta, 1)$ steps. Let us condition on the event in the statement of Lemma 3.4, that is for all $U \in \mathcal{M}_{\Delta,n}$ it holds that $|d_{U, \mathcal{P}_U}(G) - d_{U, \mathcal{P}_U}(\mathbb{G}(q, G))| \leq \Delta$. Let $\mathcal{N}$ be the set of all k-colored digraphons $\mathbf{W}$ that are $\mathcal{I}_n$-step functions with at most $t_k(\Delta, 1)$ steps $\mathcal{P}$, and that satisfy $d_{U, \mathcal{P}}(G) \leq 2\Delta$ for $U = \sum_{(\alpha,\beta) \in M} W^{(\alpha,\beta)}$.

Our main step in the proof is that, conditioned on the aforementioned event, we construct 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 elaborate on this argument in the following.

Let us fix an arbitrary (k, m)-coloring of F denoted by $\mathbf{F}$. According to Lemma 2.13 there exists a $\mathbf{W}$ that is a proper $\mathcal{I}_n$-step function with at most $t_k(\Delta, 1)$ steps $\mathcal{P}_{\mathbf{W}}$ such that there exists an $\mathcal{I}_n$-permutation ϕ of $[0, 1]$ such that $d_{\square \mathcal{P}}((\mathbf{W_F})^\phi, \mathbf{W}) \leq \Delta$. Therefore, by setting $U = \sum_{(\alpha,\beta) \in M} W^{(\alpha,\beta)}$ we have $d_{U, \mathcal{P}}(F) \leq \Delta$ and $U \in \mathcal{M}_{\Delta,n}$. This in turn implies 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 $d_{\square}(\mathbf{W}, (\mathbf{W_G})^\psi) \leq 4k^2 \Delta$ for some ψ that is an $\mathcal{I}_n$-permutation of $[0, 1]$.

Therefore we get that $\delta_{\square}(\mathbf{G}, \mathbf{F}) \leq (4k^2 + 1)\Delta$. By virtue of Lemma 3.1 we can assert that $|g(\mathbf{G}) - g(\mathbf{F})| \leq \varepsilon/2$. This finishes our argument, as $\mathbf{F}$ was arbitrary, and the failure probability of the conditioned event in the analysis of the upper bound is at most $\varepsilon/2$.

$\square$

4 Weak nondeterminism

We introduce an even more restrictive notion of nondeterminism corresponding to node colorings (Definition 1.2 used throughout the paper is a special case of the nondeterminism notion used commonly in complexity theory). Relying on this new concept we are able to improve on the upper bound of the sample complexity using a simplified version of our approach applied in the proof of Theorem 1.5 without significant alterations.

We formulate the definition of a stronger property than the previously defined nondeterministic testability. The notion itself may seem at first more involved, but in fact it only corresponds to the case, where the witness parameter g for a graph G is evaluated only on the set of node-colorings instead of edge-colorings of G in order to determine the f value in the maximum expression. This modification will enable us to rely only on the cut-norm and the corresponding regularity lemmas instead of the cut-$\mathcal{P}$-norm that was employed in the general case, thus leads us to improved 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 $\mathcal{T} = (T_1, \ldots, T_k)$ be a partition of $V(G)$ and $\mathcal{D} = ((D_1, \ldots, D_m), (D'_1, \ldots, D'_m))$ be two partitions of $[k]^2$, together they induce two partitions, $\mathcal{C} = ((C_1, \ldots, C_m), (C'_1, \ldots, C'_m))$, of $V(G)^2$ such that each class is of the form $C_i = \cup_{(\alpha,\beta) \in D_i} T_\alpha \times T_\beta$ and $C'_i = \cup_{(\alpha,\beta) \in D'_i} T_\alpha \times T_\beta$ 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 $\mathbf{G} = (G_1, \ldots, G_m, \tilde{G}_1, \ldots, \tilde{G}_m)$ with $G_i = G[C_i]$ and $\tilde{G}_i = G^c[C'_i]$. Here G^c stands for the complement of G (the union of G and its complement is the undirected complete graph), and $G[C_i]$ is the union of induced labeled subgraphs of G between T_α and T_β for each $(\alpha, \beta) \in D_i$ for $\alpha \neq \beta$, in the case of $\alpha = \beta$ the term in the union is the induced labeled subgraph of G on the node set T_α.

These special edge-$2m$-colored graphs that can serve as node-(k,m)-colorings are given by a triple $(G, \mathcal{T}, \mathcal{D})$, where G is a simple graph, $\mathcal{T}$ is a partition of $V(G)$ into k parts, and $\mathcal{D}$ is a pair of partitions of $[k]^2$ into m parts. In the case of r-uniform hypergraphs for arbitrary $r \geq 2$ a node-(k,m)-colorings is also a triple $(G, \mathcal{T}, \mathcal{D})$, the only difference in comparison to the graph case is that $\mathcal{D}$ is a pair of partitions of $[k]^r$ into m parts, the rest of the description is analogous.

Definition 4.1. *The r-uniform hypergraph parameter f is weakly non-deterministically testable if there exist integers m and k with $m \leq k^r$ and a testable edge-$2m$-colored directed r-graph parameter g such that for any simple r-graph G we have $f(G) = \max_{\mathbf{G}} g(\mathbf{G})$, where the maximum goes over the set of node-(k,m)-colorings of G.*

We present two approaches to handle this variant of the nondeterministic testability. The first method follows the proof framework introduced for the general case for graphs, its adaptation to the current setting results an improvement on the upper bound on the sample complexity to a 2-fold exponential of the sample complexity of the witness a parameter and is also applicable to the corresponding property testing setting. The second idea entails the graph case as well as the r-uniform hypergraph setting for arbitrary rank r of the weak setting. We manage to reduce the upper bound on the sample complexity further to only exponential dependence. This approach does seems to be more problem specific, than the previous one, and it does not directly yield an analogous statement in property testing.

First approach

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

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

Proof. Our approach is quite elementary: consider the partition $\mathcal{T}$ of $[0, 1]$ and $\mathcal{C}$ that is a pair of partitions of $[0, 1]^2$ corresponding to a pair of partitions $\mathcal{D}$ of $[k]^2$ as above that together with U describe $\mathbf{U}$, 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

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

for each $i \in [m]$, and the analogous 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 one of the main results of this subsection.

Corollary 4.3. *Let f be a weakly non-deterministically testable graph parameter with witness parameter g of node-(k, m)-colored graphs with the corresponding sample complexity q_g. Then f is testable with sample complexity q_f, and there exists a $c > 0$ that does depend only on k and not on f so that for any $\varepsilon > 0$ have $q_f(\varepsilon) \leq \exp^{(2)}(cq_g^2(\varepsilon/2))$.*

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 current proof. Let G be a simple graph on n nodes, and let $\varepsilon > 0$ be fixed, $q \geq \exp^{(2)}(cq_g^2(\varepsilon/2))$ for some constant $c > 0$ that will be specified later. The part concerning the lower bound of $f(\mathbb{G}(q, G))$ is completely identical to the general case.

For the upper bound set $\Delta = \exp(-cq_q^2(\varepsilon))$. We condition on the event $\delta_\square(G, \mathbb{G}(q, G)) \leq \Delta$, whose failure probability is sufficiently small due to Lemma 2.7, i.e. for $q \geq 2^{100/\Delta^2}$ it is at

most $\exp\left(-4^{100/\Delta^2}\frac{\Delta^2}{50}\right)$. We define c to be large enough so that the above lower bound on q holds true whenever $q \geq q_f(\varepsilon)$. Now we select an arbitrary node-(k,m)-coloring $\mathbf{F}$ of $\mathbb{G}(q,G)$ and apply the Weak Regularity Lemma for $2m$-colored graphons, Lemma 2.9, in the $\mathcal{I}_n$-step function case with error parameter $\Delta/(2k^2+1)$ (keeping in mind that $m \leq k^2$) to get a tuple of $\mathcal{I}_n$-step functions forming $\mathbf{U}$ with at most $t'_{2k^2}(\Delta/(2k^2+1))$ steps. We define the $\mathcal{I}_n$-step function graphon $U = \sum_{i=1}^m U_i$ and note that our condition implies that $\delta_\square(G,U) \leq 2\Delta$, since $\delta_\square(G,U) \leq \delta_\square(G,\mathbb{G}(q,G)) + \delta_\square(\mathbb{G}(q,G),U)$. To finish the proof we apply Lemma 4.2, it implies the existence of a node-(k,m)-coloring $\mathbf{G}$ of G so that $\delta_\square(\mathbf{G},\mathbf{F}) \leq (2k^2+1)\Delta$. Applying Lemma 3.1 delivers the desired result by establishing that $|g(\mathbf{F}) - g(\mathbf{G})| \leq \varepsilon$.

$\square$

Second approach

Recall the notion of layered ground state energies of r-arrays of [4] and [8] for arbitrary $r \geq 1$.

Let $r, k \geq 1$, and $G = (G^z)_{z\in[k]^r}$ be $[k]^r$-tuple of real r-arrays of size n, and $\mathcal{T} = (T_1, \ldots, T_k)$ a partition of $[n]$ into k parts. Then

$$\mathcal{E}_{\mathcal{T}}(G) = \sum_{z\in[k]^r} \frac{1}{n^r} \sum_{i_1,\ldots,i_r=1}^n G^z(i_1,\ldots,i_r) \prod_{j=1}^r \mathbb{I}_{T_{z_j}}(i_j), \tag{4.2}$$

and

$$\hat{\mathcal{E}}(G) = \max_{\mathcal{T}} \mathcal{E}_{\mathcal{T}}(G), \tag{4.3}$$

where the maximum runs over all integer partitions $\mathcal{T}$ of $[n]$ into k parts.

We will make use of the next generalization of Theorem 3.5 for hypergraphs from [1] and [8] that deals with the testability of layered ground state energies, in particular the dependence of the upper bound on the sample complexity on the dimension r.

Theorem 4.4. *[1, 8] Let $r \geq 1$, $q \geq 1$, and $\varepsilon > 0$. Then for any $[q]^r$-tuple of $([-\|W\|_\infty, \|W\|_\infty], r)$-graphons $W = (W^z)_{z\in[q]^r}$ and $k \geq \Theta^4 \log(\Theta)q^r$ with $\Theta = \frac{2^{r+7}q^r r}{\varepsilon}$ we have*

$$\mathcal{P}(|\mathcal{E}(W) - \hat{\mathcal{E}}(\mathbb{G}(k,W))| > \varepsilon\|W\|_\infty) < \varepsilon. \tag{4.4}$$

We are ready to state and prove the other main result of the section that includes a further improvement fo the upper bound on the sample complexity compared to our first approach in the weak nondeterministic testing setting.

Theorem 4.5. *Let $r \geq 1$ and f be a weakly non-deterministically testable r-graph parameter with witness parameter g of node-(k,m)-colored graphs, and let the corresponding sample complexity functions be q_f and q_g. Then f is testable and there exist a $c_{r,k} > 0$ that does depend only on r and k, but not on f such that for any $\varepsilon > 0$ we have $q_f(\varepsilon) \leq \exp(c_{r,k}q_g(\varepsilon/8))$.*

Proof. Let $r \geq 1$ be arbitrary, and f be a weakly nondeterministically testable r-graph parameter with a certificate specified by the constants k and $m \leq k^r$, and the testable $2m$-colored r-graph parameter g. Then

$$f(G) = \max_{\mathcal{T},\mathcal{D}} g(\mathbf{G}(G,\mathcal{T},\mathcal{D})),$$

where the maximum goes over every pair $(\mathcal{T},\mathcal{D})$, where $\mathcal{T}$ is a partition of $V(G)$ into k parts, and $\mathcal{D}$ is a pair of partitions of $[k]^r$ into m parts, and $\mathbf{G}(G,\mathcal{T},\mathcal{D})$ is the edge $2m$-colored graph defined by its parameters as seen above. Define the for each fixed $\mathcal{D}$ the node-k-colored r-graph (i.e., a simple r-graph together with a k-coloring of its nodes) parameter $g^{\mathcal{D}}(G,\mathcal{T}) = g(\mathbf{G}(G,\mathcal{T},\mathcal{D}))$ and the simple r-graph parameter $f^{\mathcal{D}}(G) = \max_{\mathcal{T}} g^{\mathcal{D}}(G,\mathcal{T})$.

Let $\varepsilon > 0$ be arbitrary, define

$$g^{\varepsilon}(\mathbf{G}(G,\mathcal{P},\mathcal{D})) = \sum_{F,\mathcal{T}} t(\mathbf{G}(F,\mathcal{T},\mathcal{D}), \mathbf{G}(G,\mathcal{P},\mathcal{D})) g(\mathbf{G}(F,\mathcal{T},\mathcal{D})),$$

where the sum goes over all simple r-graphs F on $q_0 = q_g(\varepsilon/8)$ vertices and partitions $\mathcal{T}$ of $[q_0]$ into k parts. By the testability of g we have

$$|g^{\varepsilon}(\mathbf{G}(G,\mathcal{P},\mathcal{D})) - g(\mathbf{G}(G,\mathcal{P},\mathcal{D}))| \leq \varepsilon/4, \tag{4.5}$$

for each permitted tuple $(G,\mathcal{P},\mathcal{D})$. Analogously we define

$$g^{\varepsilon,\mathcal{D}}(G,\mathcal{P}) = \sum_{F,\mathcal{T}} t(\mathbf{G}(F,\mathcal{T},\mathcal{D}), \mathbf{G}(G,\mathcal{P},\mathcal{D})) g(\mathbf{G}(F,\mathcal{T},\mathcal{D})),$$

and

$$f^{\varepsilon,\mathcal{D}}(G) = \max_{\mathcal{T}} g^{\varepsilon,\mathcal{D}}(G,\mathcal{T}).$$

It follows from (4.5) that for any G simple r-graph

$$|f^{\varepsilon,\mathcal{D}}(G) - f^{\mathcal{D}}(G)| \leq \varepsilon/4,$$

and for any $\varepsilon > 0$ and $q \geq 1$ we have

$$|f(G) - f(\mathbb{G}(q,G))| \leq \max_{\mathcal{D}} |f^{\varepsilon,\mathcal{D}}(G) - f^{\varepsilon,\mathcal{D}}(\mathbb{G}(q,G))| + \varepsilon/2. \tag{4.6}$$

For any $\varepsilon > 0$ and $\mathcal{D}$ that is a pair of partitions of $[k]^r$ into m parts the parameter $f^{\varepsilon,\mathcal{D}}$ can be re-written as an energy of q_0-arrays: For G of size $[n]$ let $H = (H^z)_{z \in [k]^{q_0}}$ so that for each $z \in [k]^{q_0}$ the real q_0-array H^z is defined by

$$H^z(i_1, \ldots, i_{q_0}) = g^{\mathcal{D}}(G[(i_1, \ldots, i_{q_0})], \mathcal{P}_z(i_1, \ldots, i_{q_0}))$$

for each $(i_1, \ldots, i_{q_0}) \in [n]^{q_0}$, where $\mathcal{P}_z(i_1, \ldots, i_{q_0}) = (P_1, \ldots, P_k)$ is a partition of $(i_1, \ldots, i_{q_0})$ given by $P_l = \{ i_j \mid z_j = l \}$ for $l \in [k]$. Then for each $\mathcal{T}$ that is a partition of $[n]$ into k parts we can assert that

$$
\begin{aligned}
g^{\varepsilon, \mathcal{D}}(G, \mathcal{T}) &= \sum_{z \in [k]^{q_0}} \frac{1}{n^{q_0}} \sum_{i_1, \ldots, i_r = 1}^{n} g^{\mathcal{D}}(G[(i_1, \ldots, i_{q_0})], \mathcal{P}_z(i_1, \ldots, i_{q_0})) \prod_{j=1}^{r} \mathbb{I}_{T_{z_j}}(i_j) \\
&= \sum_{z \in [k]^{q_0}} \frac{1}{n^{q_0}} \sum_{i_1, \ldots, i_r = 1}^{n} H^z(i_1, \ldots, i_{q_0}) \prod_{j=1}^{r} \mathbb{I}_{T_{z_j}}(i_j) \\
&= \mathcal{E}_{\mathcal{T}}(H),
\end{aligned}
$$

and further

$$
f^{\varepsilon, \mathcal{D}}(G) = \max_{\mathcal{T}} g^{\varepsilon, \mathcal{D}}(G, \mathcal{T}) = \max_{\mathcal{T}} \mathcal{E}_{\mathcal{T}}(H) = \hat{\mathcal{E}}(H).
$$

Analogously it holds for any $q \geq q_0$ that $f^{\varepsilon, \mathcal{D}}(\mathbb{G}(q, G)) = \hat{\mathcal{E}}(\mathbb{G}(q, H))$.

This implies by Theorem 4.4 that for $q \geq \Theta^4 \log(\Theta)$ with $\Theta = \frac{2^{q_0 + 11} k^{q_0} q_0}{\varepsilon}$ and each fixed $\mathcal{D}$ that

$$
\mathbb{P}(|f^{\varepsilon, \mathcal{D}}(G) - f^{\varepsilon, \mathcal{D}}(\mathbb{G}(q, G))| > \varepsilon/2) < 2 \exp\left(-\frac{\varepsilon^2 q}{32 q_0^2}\right).
$$

The probability that the event in the previous formula occurs for some $\mathcal{D}$ is at most $k^{2rk^r} 2 \exp\left(-\frac{\varepsilon^2 q}{32 q_0^2}\right)$, therefore by recalling (4.6) we can conclude that there exists a constant $c_{r,k} > 0$ not depending on other specifics of f such that for each simple graph G and $q \geq \exp(c_{r,k} q_g(\varepsilon/8))$ it holds that

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

$\square$

5   Further research

The sample complexity upper bounds provided in this paper in both the general and the special case are not known to be sharp, moreover, the lower bounds available at the moment are only trivial ones. An interesting open question is to improve both upper and lower bounds in the above setting, perhaps upper bound conditions (such as polynomial testability) for the witness parameter are of relevance here.

Acknowledgement

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

References

[1] Noga Alon, Wenceslas 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] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[4] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012.

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

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

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

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

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

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

[11] László Lovász and Balázs Szegedy. Limits of compact decorated graphs, 2010. preprint, arXiv:1010.5155.

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