58th Annual IEEE Symposium on Foundations of Computer Science

A characterization of testable hypergraph properties [Extended Abstract]

Felix Joos, Jaehoon Kim, Daniela Kühn and Deryk Osthus
School of Mathematics
University of Birmingham, Birmingham, United Kingdom
Email: {f.joos,j.kim.3,d.kuhn,d.osthus}@bham.ac.uk

***Abstract*—We provide a combinatorial characterization of all testable properties of k-graphs (i.e. k-uniform hypergraphs). Here, a k-graph property $\mathbf{P}$ is testable if there is a randomized algorithm which makes a bounded number of edge queries and distinguishes with probability $2/3$ between k-graphs that satisfy $\mathbf{P}$ and those that are far from satisfying $\mathbf{P}$. For the 2-graph case, such a combinatorial characterization was obtained by Alon, Fischer, Newman and Shapira. Our results for the k-graph setting are in contrast to those of Austin and Tao, who showed that for the somewhat stronger concept of local repairability, the testability results for graphs do not extend to the 3-graph setting.**

***Keywords*-property testing, hypergraphs, regularity lemma**

I. Introduction

The universal question in the area of property testing is the following: By considering a small (random) sample S of a combinatorial object $\mathcal{O}$, can we distinguish (with high probability) whether $\mathcal{O}$ has a specific property $\mathbf{P}$ or whether it is far from satisfying $\mathbf{P}$? In this paper we answer this question for k-uniform hypergraphs, where a hypergraph H is *k-uniform* if all edges of H have size $k \in \mathbb{N}$. For brevity, we usually refer to k-uniform hypergraph as *k-graphs* (so 2-graphs are graphs).

We now formalize the notion of testability (throughout, we consider only properties $\mathbf{P}$ which are decidable). For this, we say that two k-graphs G and H on vertex set V with $|V| = n$ are *α-close* if $|G \triangle H| \leq \alpha\binom{n}{k}$, and *$\alpha$-far* otherwise[1]. We say that H is *α-close to satisfying a property* $\mathbf{P}$ if there exists a k-graph G that satisfies $\mathbf{P}$ and is α-close to H, and we say that H is *α-far from satisfying* $\mathbf{P}$ otherwise.

Definition 1 (Testability)**.** *Let $k \in \mathbb{N} \setminus \{1\}$ be fixed and let $q_k : (0,1) \to \mathbb{N}$ be a function. A k-graph property $\mathbf{P}$ is* testable with query complexity at most q_k *if for every $n \in \mathbb{N}$ and every $\alpha \in (0,1)$ there are an integer $q'_k = q'_k(n,\alpha) \leq q_k(\alpha)$ and a randomized algorithm $\mathbf{T} = \mathbf{T}(n,\alpha)$ that can distinguish with probability at least $2/3$ between n-vertex k-graphs satisfying $\mathbf{P}$ and n-vertex k-graphs that are α-far from satisfying $\mathbf{P}$, while making q'_k edge queries:*

(i) *if H satisfies $\mathbf{P}$, then $\mathbf{T}$ accepts H with probability at least $2/3$,*

(ii) *if H is α-far from satisfying $\mathbf{P}$, then $\mathbf{T}$ rejects H with probability at least $2/3$.*

In this case, we say $\mathbf{T}$ is a tester, *or (n,α)*-tester *for $\mathbf{P}$. We also say that $\mathbf{T}$ has query complexity q'_k. The property $\mathbf{P}$ is* testable *if it is testable with query complexity at most q_k for some function $q_k : (0,1) \to \mathbb{N}$.*

Property testing was introduced by Rubinfeld and Sudan [1]. In the graph setting, the earliest systematic results were obtained in a seminal paper of Goldreich, Goldwasser and Ron [2]. These included k-colourability, max-cut and more general graph partitioning problems. (In fact, these results are preceded by the famous triangle removal lemma of Ruzsa and Szemerédi [3], which can be rephrased in terms of testability of triangle-freeness.) This list of problems was greatly extended (e.g. via a description in terms of first order logic by Alon, Fischer, Krivelevich, and Szegedy [4]) and generalized first to monotone properties (which are closed under vertex and edge deletion) by Alon and Shapira [5] and then to hereditary properties (which are closed under vertex deletion), again by Alon and Shapira [6]. Examples of non-testable properties include some properties which are closed under edge deletion [7] and the property of being isomorphic to a given graph G [8], [9], provided the local structure of G is sufficiently 'complex' (e.g. G is obtained as a binomial random graph). This sequence of papers culminated in the result of Alon, Fischer, Newman and Shapira [8] who obtained a combinatorial characterisation of all testable graph properties. This solved a problem posed already by [2], which was regarded as one of the main open problems in the area.

The characterisation proved in [8] states that a 2-graph property $\mathbf{P}$ is testable if and only if it is 'regular reducible'. Roughly speaking, the latter means that $\mathbf{P}$ can be characterised by being close to one of a bounded number of (weighted) Szemerédi-partitions (which arise from an application of Szemerédi's regularity lemma). Our main theorem (Theorem 2) shows that this can be extended to hypergraphs of higher uniformity. Our characterisation is based on the concept of (strong) hypergraph regularity, which was introduced in the ground-breaking work of Rödl et al. [10], [11], [12], [13], Gowers [14], see also Tao [15]. We defer the precise definition of regular reducibility for k-graphs to Section II-E, as the concept of (strong) hypergraph regularity

[1]We identify hypergraphs with their edge set and for two sets A, B we denote by $A \triangle B$ the symmetric difference of A and B.

DOI 10.1109/FOCS.2017.84

IEEE computer society

involves additional features compared to the graph setting (in particular, one needs to consider an entire (suitably nested) family of regular partitions, one for each $j \in [k]$). Accordingly, our argument relies on the so-called 'regular approximation lemma' due to Rödl and Schacht [12], which can be viewed as a powerful variant of the hypergraph regularity lemma. In turn, we derive a strengthening of this result which may have further applications.

Instead of testing whether H satisfies $\mathbf{P}$ or is α-far from $\mathbf{P}$, it is natural to consider the more general task of estimating the distance between H and $\mathbf{P}$: given $\alpha > \beta > 0$, is H $(\alpha - \beta)$-close to satisfying $\mathbf{P}$ or is H α-far from satisfying $\mathbf{P}$? In this case we refer to $\mathbf{P}$ as being *estimable*. The formal definition is similar to that in Definition 1. We show that testability and estimability are in fact equivalent. For graphs this goes back to Fischer and Newman [16].

Theorem 2. *Suppose $k \in \mathbb{N} \setminus \{1\}$ and suppose $\mathbf{P}$ is a k-graph property. Then the following three statements are equivalent:*

(a) $\mathbf{P}$ *is testable.*
(b) $\mathbf{P}$ *is estimable.*
(c) $\mathbf{P}$ *is regular reducible.*

In Section V, we illustrate how Theorem 2 can be used to prove testability of a given property: firstly to test the injective homomorphism density of a given subgraph (which includes the classical example of H-freeness) and secondly to test the size of a maximum ℓ-way cut (which includes testing ℓ-colourability).

Previously, the most general result on hypergraph property testing was the testability of hereditary properties, which was proved by Rödl and Schacht [17], [18], based on deep results on hypergraph regularity. In fact, they showed that hereditary k-graph properties can be even tested with one-sided error (which means that the '2/3' is replaced by '1' in Definition 1(i)). This generalized earlier results in [19], [20].

The result of Alon and Shapira on the testability of hereditary graph properties was strengthened by Austin and Tao [21] in another direction: they showed that hereditary properties of graphs are not only testable with one-sided error, but they are also *locally repairable*[2] (one may think of this as a strengthening of testability). On the other hand, they showed that hereditary properties of 3-graphs are not necessarily locally repairable. Note that this is in contrast to Theorem 2.

An intimate connection between property testing and graph limits was established by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi [22]. In particular, they showed that a graph property $\mathbf{P}$ is testable if and only if for all sequences (G_n) of graphs with $|V(G_n)| \to \infty$ and $\delta_\square(G_n, \mathbf{P}) \to 0$, we have $d_1(G_n, \mathbf{P}) \to 0$. Here $\delta_\square(G, \mathbf{P})$ denotes the cut-distance of G and the closest graph satisfying $\mathbf{P}$ and $d_1(G, \mathbf{P})$ is the normalized edit-distance between G and $\mathbf{P}$ (see also [23] for more background and discussion on this). Another characterisation (in terms of localized samples) using the graph limit framework was given by Lovász and Szegedy [24]. Similarly, the result of Rödl and Schacht [17] on testing hereditary hypergraph properties was reproven via hypergraph limits by Elek and Szegedy [25] as well as Austin and Tao [21]. The latter further extended this to directed pre-coloured hypergraphs (none of these results however yield effective bounds on the query complexity).

Lovász and Vesztergombi [26] recently introduced the notion of 'non-deterministic' property testing, where the tester also has access to a 'certificate' for the property $\mathbf{P}$. By considering the graph limit setting, they proved the striking result that any non-deterministically testable graph property is also deterministically testable (one could think of their result as the graph property testing analogue of proving that P = NP). Karpinski and Markó [27] generalized the Lovász-Vesztergombi result to hypergraphs, also via the notion of (hyper-)graph limits. However, these proofs do not give an explicit bounds on the query complexity – this was achieved by Gishboliner and Shapira [28] for graphs and Karpinski and Markó [29] for hypergraphs.

Another direction of research concerns *easily testable properties*, where we require that the size of the sample is bounded from above by a polynomial in $1/\alpha$. (The bounds coming from Theorem 2 can be made explicit but are quite large, as the approach via the (hyper-)graph regularity lemma incurs at least a tower-type dependence on $1/\alpha$, see [30].) For k-graphs, Alon and Shapira [31] as well as Alon and Fox [32] obtained positive and negative results for the property of containing a given k-graph as an (induced) subgraph. For an approach via a 'polynomial' version of the regularity lemma see [33].

Recent progress on property testing includes many questions beyond the hypergraph setting. Instances include property testing of matrices [34], Boolean functions [35], [36], geometric objects [37], and algebraic structures [38], [33], [39]. Moreover, property testing in the sparse (graph) setting gives rise to many interesting results and questions (see e.g. [40], [41]). Little is known for hypergraphs in this case.

The paper is organized as follows. In Section II, we explain the relevant concepts of hypergraph regularity, in particular we introduce the regular approximation lemma of Rödl and Schacht (Theorem 5). In Section III, we describe a version of the induced counting lemma. In Section IV, we sketch the proof of Theorem 2 and in Section V we discuss applications of our main result and illustrate in detail how

[2]Suppose $\mathbf{P}$ is a hereditary graph property and $\varepsilon > 0$. We say that a graph G is locally δ-close to $\mathbf{P}$ if a random sample S satisfies $\mathbf{P}$ with probability at least $1 - \delta$. A result of Alon and Shapira [6] shows that whenever G is locally δ-close to $\mathbf{P}$ for some $\delta(\varepsilon) > 0$, then G is ε-close to $\mathbf{P}$. The concept of being locally repairable strengthens this by requiring a rule that generates $G' \in \mathbf{P}$ only based on S such that $|G \triangle G'| < \varepsilon n^2$ with probability at least $1 - \delta$.

to apply Theorem 2.

II. Hypergraph regularity

The constants in the hierarchies used to state our results have to be chosen from right to left. More precisely, if we claim that a result holds whenever $1/n \ll a \ll b \leq 1$ (where $n \in \mathbb{N}$ is typically the number of vertices of a hypergraph), then this means that there are non-decreasing functions $f : (0,1] \to (0,1]$ and $g : (0,1] \to (0,1]$ such that the result holds for all $0 < a, b \leq 1$ and all $n \in \mathbb{N}$ with $a \leq f(b)$ and $1/n \leq g(a)$. We say a set E is an *i-set* if $|E| = i$.

A. Basic hypergraph notation

In the following we introduce several concepts about a hypergraph H. We typically refer to $V = V(H)$ as the vertex set of H and usually let $n := |V|$. Given a hypergraph H and a set $Q \subseteq V(H)$, we denote by $H[Q]$ the hypergraph induced on H by Q.

For a partition $\{V_1, \dots, V_\ell\}$ of V and $k \in [\ell]$, we denote by $K_\ell^{(k)}(V_1, \dots, V_\ell)$ the *complete ℓ-partite* k-graph with vertex classes $V_1, \dots, V_\ell$. If $|V_i| = |V_j| \pm 1$ for all $i, j \in [\ell]$, then an *(ℓ, k)-graph* H on $\{V_1, \dots, V_\ell\}$ is a spanning subgraph of $K_\ell^{(k)}(V_1, \dots, V_\ell)$. For notational convenience, we consider the vertex partition $\{V_1, \dots, V_\ell\}$ as an $(\ell, 1)$-graph.

If $2 \leq k \leq i \leq \ell$ and H is an (ℓ, k)-graph, we denote by $\mathcal{K}_i(H)$ the family of all i-element subsets I of $V(H)$ for which $H[I] \cong K_i^{(k)}$, where $K_i^{(k)}$ denotes the complete k-graph on i vertices. If $H^{(1)}$ is an $(\ell, 1)$-graph on $\{V_1, \dots, V_\ell\}$ and $i \in [\ell]$, we denote by $\mathcal{K}_i(H^{(1)})$ the family of all i-element subsets I of $V(H^{(1)})$ which 'cross' the partition $\{V_1, \dots, V_\ell\}$; that is, $I \in \mathcal{K}_i(H^{(1)})$ if and only if $|I \cap V_s| \leq 1$ for all $s \in [\ell]$.

We will consider hypergraphs of different uniformity on the same vertex set. Given an $(\ell, k-1)$-graph $H^{(k-1)}$ and an (ℓ, k)-graph $H^{(k)}$ on $\{V_1, \dots, V_\ell\}$, we say $H^{(k-1)}$ *underlies* $H^{(k)}$ if $H^{(k)} \subseteq \mathcal{K}_k(H^{(k-1)})$; that is, for every edge $e \in H^{(k)}$ and every $(k-1)$-subset f of e, we have $f \in H^{(k-1)}$. If we have an entire cascade of underlying hypergraphs we refer to this as a complex. More precisely, let $\ell \geq k \geq 1$ be integers. An *(ℓ, k)-complex* $\mathcal{H}$ on $\{V_1, \dots, V_\ell\}$ is a collection of (ℓ, j)-graphs $\{H^{(j)}\}_{j=1}^k$ on $\{V_1, \dots, V_\ell\}$ such that $H^{(j-1)}$ underlies $H^{(j)}$ for every $j \in [k] \setminus \{1\}$.

B. Hypergraph regularity

In this subsection we introduce ε-regularity for hypergraphs. Suppose $\ell \geq k \geq 2$. Let $H^{(k)}$ be an (ℓ, k)-graph on $\{V_1, \dots, V_\ell\}$, let $\{i_1, \dots, i_k\} \in \binom{[\ell]}{k}$, and let $H^{(k-1)}$ be a $(k, k-1)$-graph on $\{V_{i_1}, \dots, V_{i_k}\}$. Suppose $\varepsilon > 0$ and $d \geq 0$. We say $H^{(k)}$ is *(ε, d)-regular with respect to* $H^{(k-1)}$ if for all $Q^{(k-1)} \subseteq H^{(k-1)}$ with

$$|\mathcal{K}_k(Q^{(k-1)})| \geq \varepsilon |\mathcal{K}_k(H^{(k-1)})|, \text{ we have}$$
$$|H^{(k)} \cap \mathcal{K}_k(Q^{(k-1)})| = (d \pm \varepsilon)|\mathcal{K}_k(Q^{(k-1)})|.$$

We say $H^{(k)}$ is *ε-regular with respect to* $H^{(k-1)}$ if it is (ε, d)-regular with respect to $H^{(k-1)}$ for some $d \geq 0$. We say an (ℓ, k)-graph $H^{(k)}$ on $\{V_1, \dots, V_\ell\}$ is *(ε, d)-regular with respect to an $(\ell, k-1)$-graph* $H^{(k-1)}$ on $\{V_1, \dots, V_\ell\}$ if for every $\Lambda \in \binom{[\ell]}{k}$, the k-graph $H^{(k)}$ is (ε, d)-regular with respect to the restriction $H^{(k-1)}[\bigcup_{\lambda \in \Lambda} V_\lambda]$.

Let $\mathbf{d} = (d_2, \dots, d_k) \in \mathbb{R}_{\geq 0}^{k-1}$. We say an (ℓ, k)-complex $\mathcal{H} = \{H^{(j)}\}_{j=1}^k$ is *$(\varepsilon, \mathbf{d})$-regular* if $H^{(j)}$ is (ε, d_j)-regular with respect to $H^{(j-1)}$ for every $j \in [k] \setminus \{1\}$.

C. Partitions of hypergraphs and the regular approximation lemma

The regular approximation lemma of Rödl and Schacht implies that for all k-graphs H, there exists a k-graph G which is very close to H and so that G has a very 'high quality' partition into ε-regular subgraphs. To state this formally we need to introduce further concepts involving partitions of hypergraphs.

We start with the concept of a polyad. Roughly speaking, given a vertex partition $\mathscr{P}^{(1)}$, an i-polyad is an i-graph which arises from a partition $\mathscr{P}^{(i)}$ of the complete partite i-graph $\mathcal{K}_i(\mathscr{P}^{(1)})$. The $(i+1)$-cliques spanned by all the i-polyads give rise to a partition $\mathscr{P}^{(i+1)}$ of $\mathcal{K}_{i+1}(\mathscr{P}^{(1)})$ (see Definition 3). Such a 'family of partitions' then provides a suitable framework for describing a regularity partition (see Definition 4).

Suppose we have a vertex partition $\mathscr{P}^{(1)} = \{V_1, \dots, V_\ell\}$ and $\ell \geq k$. Recall that $\mathcal{K}_j(\mathscr{P}^{(1)})$ is the family of all crossing j-sets with respect to $\mathscr{P}^{(1)}$. Suppose that for all $i \in [k] \setminus \{1\}$, we have partitions $\mathscr{P}^{(i)}$ of $\mathcal{K}_i(\mathscr{P}^{(1)})$ such that each part of $\mathscr{P}^{(i)}$ is an (i,i)-graph with respect to $\mathscr{P}^{(1)}$. By definition, for each i-set $I \in \mathcal{K}_i(\mathscr{P}^{(1)})$, there exists exactly one i-graph $P^{(i)} = P^{(i)}(I) \in \mathscr{P}^{(i)}$ so that $I \in P^{(i)}$. Consider $j \in [\ell]$ and any $J \in \mathcal{K}_j(\mathscr{P}^{(1)})$. For each $1 \leq i \leq \max\{j, k-1\}$, the *$i$-polyad* $\hat{P}^{(i)}(J)$ of J is defined by

$$\hat{P}^{(i)}(J) := \bigcup \left\{ P^{(i)}(I) : I \in \binom{J}{i} \right\}.$$

Thus $\hat{P}^{(i)}(J)$ is an (j,i)-graph with respect to $\mathscr{P}^{(1)}$. For $j \in [k-1]$, let

$$\hat{\mathscr{P}}^{(j)} := \left\{ \hat{P}^{(j)}(J) : J \in \mathcal{K}_{j+1}(\mathscr{P}^{(1)}) \right\}.$$

We note that $\hat{\mathscr{P}}^{(1)}$ is the set consisting of all $(2,1)$-graphs with vertex classes V_s, V_t (for distinct $s, t \in [\ell]$). Moreover, note that if $\hat{P}^{(j)} \in \hat{\mathscr{P}}^{(j)}$, it follows that there is a set $J \in \mathcal{K}_{j+1}(\mathscr{P}^{(1)})$ such that $\hat{P}^{(j)} = \hat{P}^{(j)}(J)$. Since $J \in \mathcal{K}_{j+1}(\hat{P}^{(j)}(J))$, we obtain that $\mathcal{K}_{j+1}(\hat{P}^{(j)}) \neq \emptyset$ for any $\hat{P}^{(j)} \in \hat{\mathscr{P}}^{(j)}$.

The above definitions apply to arbitrary partitions $\mathscr{P}^{(i)}$ of $\mathcal{K}_i(\mathscr{P}^{(1)})$. However, it will be useful to consider partitions with more structure. Suppose $A \supseteq B$ are finite sets, $\mathscr{A}$ is a partition of A, and $\mathscr{B}$ is a partition of B. We say $\mathscr{A}$ *refines*

$\mathscr{B}$ and write $\mathscr{A} \prec \mathscr{B}$ if for every $\mathcal{A} \in \mathscr{A}$ there either exists $\mathcal{B} \in \mathscr{B}$ such that $\mathcal{A} \subseteq \mathcal{B}$ or $\mathcal{A} \subseteq A \setminus B$.

Definition 3 (Family of partitions)**.** *Suppose $k \in \mathbb{N}\setminus\{1\}$ and $\mathbf{a} = (a_1, \ldots, a_{k-1}) \in \mathbb{N}^{k-1}$. We say $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a}) = \{\mathscr{P}^{(1)}, \ldots, \mathscr{P}^{(k-1)}\}$ is a* family of partitions *on V if it satisfies the following:*

(i) *$\mathscr{P}^{(1)}$ is a partition of V into $a_1 \geq k$ nonempty classes,*

(ii) *for each $j \in [k-1] \setminus \{1\}$, the set $\mathscr{P}^{(j)}$ is a partition of $\mathcal{K}_j(\mathscr{P}^{(1)})$ into nonempty j-graphs such that*
- *$\mathscr{P}^{(j)} \prec \{\mathcal{K}_j(\hat{P}^{(j-1)}) : \hat{P}^{(j-1)} \in \hat{\mathscr{P}}^{(j-1)}\}$ and*
- *$|\{P^{(j)} \in \mathscr{P}^{(j)} : P^{(j)} \subseteq \mathcal{K}_j(\hat{P}^{(j-1)})\}| = a_j$ for every $\hat{P}^{(j-1)} \in \hat{\mathscr{P}}^{(j-1)}$.*

We now extend the concept of ε-regularity to families of partitions.

Definition 4 (Equitable family of partitions)**.** *Let $k \in \mathbb{N} \setminus \{1\}$. Suppose $\eta > 0$ and $\mathbf{a} = (a_1, \ldots, a_{k-1}) \in \mathbb{N}^{k-1}$. Let V be a vertex set of size n. We say a family of partitions $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a})$ on V is* $(\eta, \varepsilon, \mathbf{a})$-equitable *if it satisfies the following:*

(i) *$1/a_1 \leq \eta$,*

(ii) *$\mathscr{P}^{(1)} = \{V_i : i \in [a_1]\}$ satisfies $|V_i| = |V_j| \pm 1$ for all $i, j \in [a_1]$, and*

(iii) *if $k \geq 3$, then for every k-set $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$ the collection $\{\hat{P}^{(j)}(K)\}_{j=1}^{k-1}$ is an $(\varepsilon, \mathbf{d})$-regular $(k, k-1)$-complex, where $\mathbf{d} = (1/a_2, \ldots, 1/a_{k-1})$.*

Having introduced the necessary notation, we are now ready to state the regular approximation lemma due to Rödl and Schacht. It states that for every k-graph H, there is a k-graph G that is close to H and that has very good regularity properties.

Theorem 5 (Regular approximation lemma [12])**.** *Let $k \in \mathbb{N} \setminus \{1\}$. For all $\eta, \nu > 0$ and every function $\varepsilon : \mathbb{N}^{k-1} \to (0, 1]$, there are integers $t_0 := t_5(\eta, \nu, \varepsilon)$ and $n_0 := n_5(\eta, \nu, \varepsilon)$ so that the following holds:*

For every k-graph H on at least $n \geq n_0$ vertices, there exists a k-graph G on $V(H)$ and a family of partitions $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a}^{\mathscr{P}})$ on $V(H)$ so that

(i) *$\mathscr{P}$ is $(\eta, \varepsilon(\mathbf{a}^{\mathscr{P}}), \mathbf{a}^{\mathscr{P}})$-equitable and $\mathbf{a}^{\mathscr{P}} \in [t_0]^{k-1}$,*

(ii) *G is $\varepsilon(\mathbf{a}^{\mathscr{P}})$-regular with respect to $\hat{P}^{(k-1)}$ for every $\hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)}$, and*

(iii) *$|G \triangle H| \leq \nu\binom{n}{k}$.*

The crucial point here is that in applications we may apply Theorem 5 with a function ε such that $\varepsilon(\mathbf{a}^{\mathscr{P}}) \ll \|\mathbf{a}^{\mathscr{P}}\|_\infty^{-1}$. This is in contrast to other versions (see e.g. [14], [13]) where (roughly speaking) in (iii) we have $G = H$ but in (ii) one has to pay for this by incurring an error parameter ε' which may be large compared to $\|\mathbf{a}^{\mathscr{P}}\|_\infty^{-1}$.

D. The address space

Later on, we will need to explicitly refer to the densities arising in Theorem 5(ii). For this (and other reasons) it is convenient to consider the 'address space'. Roughly speaking the address space consists of a collection of vectors where each vector identifies a polyad.

To define the address space, let us write $\binom{[a_1]}{\ell}_< := \{(\alpha_1, \ldots, \alpha_\ell) \in [a_1]^\ell : \alpha_1 < \cdots < \alpha_\ell\}$. Suppose $k, \ell, p \in \mathbb{N}$, $\ell \geq k$, and $p \geq \max\{k-1, 1\}$, and $\mathbf{a} = (a_1, \ldots, a_p) \in \mathbb{N}^p$. We define

$$\hat{A}(\ell, k-1, \mathbf{a}) := \binom{[a_1]}{\ell}_< \times \prod_{j=2}^{k-1} [a_j]^{\binom{\ell}{j}}$$

to be the (ℓ, k)-*address space*. Observe that $\hat{A}(1, 0, \mathbf{a}) = [a_1]$ and $\hat{A}(2, 1, \mathbf{a}) = \binom{[a_1]}{2}_<$. Note that if $k > 1$, then each $\hat{\mathbf{x}} \in \hat{A}(\ell, k-1, \mathbf{a})$ can be written as $\hat{\mathbf{x}} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)})$, where $\mathbf{x}^{(1)} \in \binom{[a_1]}{\ell}_<$ and $\mathbf{x}^{(j)} \in [a_j]^{\binom{\ell}{j}}$ for each $j \in [k-1] \setminus \{1\}$. Thus each entry of the vector $\mathbf{x}^{(j)}$ corresponds to (i.e. is indexed by) a subset of $\binom{[\ell]}{j}$. For a vector $\mathbf{x} = (\alpha_1, \ldots, \alpha_\ell)$, we let $\mathbf{x}_* := \{\alpha_1, \ldots, \alpha_\ell\}$. We order the elements of both $\binom{[\ell]}{j}$ and $\binom{\mathbf{x}_*^{(1)}}{j}$ lexicographically and consider the bijection $g : \binom{\mathbf{x}_*^{(1)}}{j} \to \binom{[\ell]}{j}$ which preserves this ordering. For each $\Lambda \in \binom{\mathbf{x}_*^{(1)}}{j}$ and $j \in [k-1]$, we denote by $\mathbf{x}_\Lambda^{(j)}$ the entry of $\mathbf{x}^{(j)}$ which corresponds to the set $g(\Lambda)$.

1) Basic properties of the address space: Let $k \in \mathbb{N}\setminus\{1\}$ and let V be a vertex set of size n. Let $\mathscr{P}(k-1, \mathbf{a})$ be a family of partitions on V. For each crossing k-set $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$, the address space allows us to identify (and thus refer to) the set of polyads 'supporting' K. We will achieve this by defining a suitable operator $\hat{\mathbf{x}}(K)$ which maps K to the address space.

To do this, write $\mathscr{P}^{(1)} = \{V_i : i \in [a_1]\}$. Recall from Definition 3(ii) that for each $j \in [k-1] \setminus \{1\}$, we partition $\mathcal{K}_j(\hat{P}^{(j-1)})$ of every $(j-1)$-polyad $\hat{P}^{(j-1)} \in \hat{\mathscr{P}}^{(j-1)}$ into a_j nonempty parts in such a way that $\mathscr{P}^{(j)}$ is the collection of all these parts. Thus, there is a labeling $\phi^{(j)} : \mathscr{P}^{(j)} \to [a_j]$ such that for every polyad $\hat{P}^{(j-1)} \in \hat{\mathscr{P}}^{(j-1)}$, the restriction of $\phi^{(j)}$ to $\{P^{(j)} \in \mathscr{P}^{(j)} : P^{(j)} \subseteq \mathcal{K}_j(\hat{P}^{(j-1)})\}$ is injective. The set $\mathbf{\Phi} := \{\phi^{(2)}, \ldots, \phi^{(k-1)}\}$ is called an $\mathbf{a}$-*labeling* of $\mathscr{P}(k-1, \mathbf{a})$. For a given set $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$, we denote $\mathrm{cl}(K) := \{i : V_i \cap K \neq \emptyset\}$.

For every k-set $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$, we define a vector $\hat{\mathbf{x}}(K) = (\mathbf{x}^{(1)}(K), \ldots, \mathbf{x}^{(k-1)}(K))$ by

- $\mathbf{x}^{(1)}(K) := (\alpha_1, \ldots, \alpha_k)$, where $\alpha_1 < \ldots < \alpha_k$ and $K \cap V_{\alpha_i} = \{v_{\alpha_i}\}$,
- and for $i \in [k-1] \setminus \{1\}$ we set

$$\mathbf{x}^{(i)}(K) := \left(\phi^{(i)}(P^{(i)}) : \{v_\lambda : \lambda \in \Lambda\} \in P^{(i)}, P^{(i)} \in \mathscr{P}^{(i)}\right)_{\Lambda \in \binom{\mathrm{cl}(K)}{i}}.$$

Here, we order $\binom{\mathrm{cl}(K)}{i}$ lexicographically. In particular, $\mathbf{x}^{(i)}(K)$ is a vector of length $\binom{k}{i}$ and $\hat{\mathbf{x}}(K) \in \hat{A}(k, k-1, \mathbf{a})$ for every $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$.

Recall that $\mathcal{K}_k(\hat{P}^{(k-1)}) \neq \emptyset$ for any $\hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)}$, and note that $\hat{\mathbf{x}}(K) = \hat{\mathbf{x}}(K')$ for all $K, K' \in \mathcal{K}_k(\hat{P}^{(k-1)})$ and all $\hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)}$. Hence, for each $\hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)}$ we can define

$$\hat{\mathbf{x}}(\hat{P}^{(k-1)}) := \hat{\mathbf{x}}(K) \text{ for some } K \in \mathcal{K}_k(\hat{P}^{(k-1)}). \quad (1)$$

Let

$$\begin{aligned}\hat{A}(k, k-1, \mathbf{a})_{\neq\emptyset} := \{&\hat{\mathbf{x}} \in \hat{A}(k, k-1, \mathbf{a}) : \exists \hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)} \\ &\text{such that } \hat{\mathbf{x}}(\hat{P}^{(k-1)}) = \hat{\mathbf{x}}\} \\ = \{&\hat{\mathbf{x}} \in \hat{A}(k, k-1, \mathbf{a}) : \exists K \in \mathcal{K}_k(\mathscr{P}^{(1)}) \\ &\text{such that } \hat{\mathbf{x}} = \hat{\mathbf{x}}(K)\}.\end{aligned}$$

Clearly (1) gives rise to a bijection from $\hat{\mathscr{P}}^{(k-1)}$ to $\hat{A}(k, k-1, \mathbf{a})_{\neq\emptyset}$. Thus for each $\hat{\mathbf{x}} \in \hat{A}(k, k-1, \mathbf{a})_{\neq\emptyset}$, we can define the polyad of $\hat{\mathbf{x}}$ by

$$\hat{P}^{(k-1)}(\hat{\mathbf{x}}) := \hat{P}^{(k-1)} \text{ such that } \hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)} \text{ with } \hat{\mathbf{x}} = \hat{\mathbf{x}}(\hat{P}^{(k-1)}).$$

Note that for any $K \in \mathcal{K}_k(\mathscr{P}^{(1)})$, we have $\hat{P}^{(k-1)}(\hat{\mathbf{x}}(K)) = \hat{P}^{(k-1)}(K)$. One can show that if $\mathscr{P}$ is an $(\eta, \varepsilon, \mathbf{a})$-equitable family of partitions and ε is small enough, then $\hat{A}(k, k-1, \mathbf{a})_{\neq\emptyset} = \hat{A}(k, k-1, \mathbf{a})$ and thus (1) gives actually rise to a bijection between $\hat{\mathscr{P}}^{(k-1)}$ and $\hat{A}(k, k-1, \mathbf{a})$.

2) Density functions of address spaces: We say a function $d_{\mathbf{a},k} : \hat{A}(k, k-1, \mathbf{a}) \to [0, 1]$ is a *density function* of $\hat{A}(k, k-1, \mathbf{a})$. Suppose we are given a density function $d_{\mathbf{a},k}$, a real $\varepsilon > 0$, and a k-graph H. We say a family of partitions $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a})$ on $V(H)$ is an $(\varepsilon, \mathbf{a}, d_{\mathbf{a},k})$-*equitable partition* of H if $\mathscr{P}$ is $(1/a_1, \varepsilon, \mathbf{a})$-equitable (as specified in Definition 4) and if for every $\hat{\mathbf{x}} \in \hat{A}(k, k-1, \mathbf{a})$ the hypergraph H is $(\varepsilon, d_{\mathbf{a},k}(\hat{\mathbf{x}}))$-regular with respect to $\hat{P}^{(k-1)}(\hat{\mathbf{x}})$. Thus if $\mathscr{P}$ and G are as obtained by Theorem 5 and $\hat{P}^{(k-1)}(\cdot) : \hat{A}(k, k-1, \mathbf{a}^{\mathscr{P}}) \to \hat{\mathscr{P}}^{(k-1)}$ is a bijection, then there exists a density function $d_{\mathbf{a}^{\mathscr{P}},k}$ such that $\mathscr{P}$ is an $(\varepsilon(\mathbf{a}^{\mathscr{P}}), \mathbf{a}^{\mathscr{P}}, d_{\mathbf{a}^{\mathscr{P}},k})$-equitable partition of G.

E. Regularity instances

A regularity instance R encodes an address space, an associated density function and a regularity parameter. Roughly speaking, a regularity instance can be thought of as encoding a weighted 'reduced multihypergraph' obtained from an application of the regularity lemma for hypergraphs. To formalize this, let $\varepsilon_6(\cdot, \cdot) : \mathbb{N} \times \mathbb{N} \to (0, 1]$ be a function which satisfies the following.

- $\varepsilon_6(\cdot, k)$ is a decreasing function for any fixed $k \in \mathbb{N}$ with $\lim_{x\to\infty} \varepsilon_6(x, k) = 0$,
- $\varepsilon_6(x, \cdot)$ is a decreasing function for any fixed $x \in \mathbb{N}$,
- $\varepsilon_6(t, k) \ll 1/t, 1/k$.

The choice of ε_6 is made more explicit in the full version of the paper. The main constraint is that it is small enough to apply an appropriate version of the hypergraph counting lemma.

Definition 6 (Regularity instance)**.** *Let $k \in \mathbb{N} \setminus \{1\}$. A* regularity instance *$R = (\varepsilon, \mathbf{a}, d_{\mathbf{a},k})$ is a triple, where $\mathbf{a} = (a_1, \ldots, a_{k-1}) \in \mathbb{N}^{k-1}$, $0 < \varepsilon \leq \varepsilon_6(\|\mathbf{a}\|_\infty, k)$, and $d_{\mathbf{a},k}$ is a density function of $\hat{A}(k, k-1, \mathbf{a})$. A k-graph H satisfies the regularity instance R if there exists a family of partitions $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a})$ such that $\mathscr{P}$ is an $(\varepsilon, \mathbf{a}, d_{\mathbf{a},k})$-equitable partition of H. The* complexity *of R is $1/\varepsilon$.*

Since ε_6 depends only on $\|\mathbf{a}\|_\infty$ and k, it follows that for given r and fixed k, the number of vectors $\mathbf{a}$ which could belong to a regularity instance R with complexity r is bounded by a function of r.

Definition 7 (Regular reducible)**.** *A k-graph property $\mathbf{P}$ is* regular reducible *if for any $\beta > 0$, there exists an $r = r_7(\beta, \mathbf{P})$ such that for any integer $n \geq k$, there is a family $\mathcal{R} = \mathcal{R}(n, \beta, \mathbf{P})$ of at most r regularity instances, each of complexity at most r, such that the following hold for every $\alpha > \beta$ and every n-vertex k-graph H:*

- *If H satisfies $\mathbf{P}$, then there exists $R \in \mathcal{R}$ such that H is β-close to satisfying R.*
- *If H is α-far from satisfying $\mathbf{P}$, then for any $R \in \mathcal{R}$ the hypergraph H is $(\alpha - \beta)$-far from satisfying R.*

Thus a property is regular reducible if it can be (approximately) encoded by a bounded number of regularity instances of bounded complexity. Note that if we apply the regular approximation lemma (Theorem 5) to H to obtain G and $\mathscr{P}$, then $\mathbf{a}^{\mathscr{P}}$ and the densities of G with respect to the polyads in $\hat{\mathscr{P}}^{(k-1)}$ naturally give rise to a regularity instance R where G satisfies R and H is close to satisfying R.

Note that different choices of ε_6 lead to a different definition of regularity instances and thus might lead to a different definition of being regular reducible. However, our main result implies that for *any* appropriate choice of ε_6, being regular reducible and testability are equivalent. In particular, if a property is regular reducible for an appropriate choice of ε_6, then it is regular reducible for all appropriate choices of ε_6, and so 'regular reducibility' is well defined.

III. A COUNTING LEMMA

Suppose H is a (large) k-graph satisfying a regularity instance $(\varepsilon, \mathbf{a}, d_{\mathbf{a},k})$ and F is a (small) k-graph. In this section we show how to express the number of copies of F in H in terms of the parameters ε, $\mathbf{a}$, and $d_{\mathbf{a},k}$.

Suppose $k, \ell \in \mathbb{N} \setminus \{1\}$ such that $\ell \geq k$ and suppose $\mathbf{a} \in \mathbb{N}^{k-1}$. Suppose that $d_{\mathbf{a},k} : \hat{A}(k, k-1, \mathbf{a}) \to [0, 1]$ is a density function. Suppose F is a k-graph on ℓ vertices and let $A(F)$ be the size of the automorphism group of F. Suppose $\hat{\mathbf{x}} \in \hat{A}(\ell, k-1, \mathbf{a})$ and suppose in the following that $\sigma : V(F) \to \mathbf{x}_*^{(1)}$ is a bijection. Given $\hat{\mathbf{y}} \in \hat{A}(k, k-1, \mathbf{a})$, we write $\hat{\mathbf{y}} \leq_{k,k-1} \hat{\mathbf{x}}$ if

- $\mathbf{y}_*^{(1)} \subseteq \mathbf{x}_*^{(1)}$ and
- $\mathbf{x}_\Lambda^{(j)} = \mathbf{y}_\Lambda^{(j)}$ for any $\Lambda \in \binom{\mathbf{y}_*^{(1)}}{j}$ and $j \in [k-1] \setminus \{1\}$.

Let

$$IC(F,d_{\mathbf{a},k}) := \frac{1}{\binom{a_1}{\ell}A(F)} \sum_{\hat{\mathbf{x}}\in\hat{A}(\ell,k-1,\mathbf{a})} \sum_{\sigma} \prod_{\substack{\hat{\mathbf{y}}\le_{k,k-1}\hat{\mathbf{x}},\\ \mathbf{y}_*^{(1)}\in\sigma(F)}} d_{\mathbf{a},k}(\hat{\mathbf{y}})\times \prod_{\substack{\hat{\mathbf{y}}\le_{k,k-1}\hat{\mathbf{x}},\\ \mathbf{y}_*^{(1)}\notin\sigma(F)}} (1-d_{\mathbf{a},k}(\hat{\mathbf{y}})) \prod_{j=2}^{k-1} a_j^{-\binom{\ell}{j}}.$$

Given an n-vertex k-graph H, we define $\mathbf{Pr}(F,H)$ such that $\mathbf{Pr}(F,H)\binom{n}{\ell}$ equals the number of induced copies of F in H. Lemma 8 implies that if H satisfies a regularity instance $(\varepsilon,\mathbf{a},d_{\mathbf{a},k})$, then $IC(F,d_{\mathbf{a},k})$ is a very accurate estimate for $\mathbf{Pr}(F,H)$. The same is true if F is replaced by a finite family $\mathcal{F}$ of k-graphs, where we define

$$\mathbf{Pr}(\mathcal{F},H) := \sum_{F\in\mathcal{F}} \mathbf{Pr}(F,H) \quad \text{and}$$
$$IC(\mathcal{F},d_{\mathbf{a},k}) := \sum_{F\in\mathcal{F}} IC(F,d_{\mathbf{a},k}).$$

Lemma 8. *Suppose $0<1/n\ll\varepsilon\ll 1/t,1/a_1\ll\gamma,1/k,1/\ell$ with $2\le k\le\ell$. Let $\mathcal{F}$ be a collection of k-graphs on ℓ vertices. Suppose H is an n-vertex k-graph satisfying a regularity instance $R=(\varepsilon,\mathbf{a},d_{\mathbf{a},k})$, where $\mathbf{a}\in[t]^{k-1}$. Then $\mathbf{Pr}(\mathcal{F},H)=IC(\mathcal{F},d_{\mathbf{a},k})\pm\gamma$.*

We derive Lemma 8 from a counting lemma for cliques in ε-regular k-graphs due to Kohayakawa, Rödl and Skokan (Theorem 6.5 in [42]).

IV. Proof sketch

In the following, we describe the main steps leading to the proof of Theorem 2. While the general strategy emulates that of [8], the hypergraph setting leads to many additional challenges.

A. Testable properties are regular reducible

We first show (a)⇒(c) in Theorem 2. Goldreich and Trevisan [7] proved that every testable graph property is also testable in some canonical way (and their results translate to the hypergraph setting in a straightforward way). Thus we may restrict ourselves to such canonical testers. More precisely, an (n,α)-tester $\mathbf{T}=\mathbf{T}(n,\alpha)$ is *canonical* if, given an n-vertex k-graph H, it chooses a set Q of $q_k'=q_k'(n,\alpha)$ vertices of H uniformly at random, queries all k-sets in Q, and then accepts or rejects H (deterministically) according to (the isomorphism class of) $H[Q]$. In particular, $\mathbf{T}$ has query complexity $\binom{q_k'}{k}$. Moreover, every canonical tester is non-adaptive.

Let $\mathbf{P}$ be a testable k-graph property. Thus there exists a function $q_k:(0,1)\to\mathbb{N}$ such that for every $n\in\mathbb{N}$ and $\alpha\in(0,1)$, there exists a canonical (n,α)-tester $\mathbf{T}=\mathbf{T}(n,\alpha)$ for $\mathbf{P}$ with query complexity at most $q_k(\alpha)$. So $\mathbf{T}$ samples a set Q of $q\le q_k(\alpha)$ vertices, considers $H[Q]$, and then deterministically accepts or rejects H based on $H[Q]$. Let $\mathcal{Q}$ be the set of all the k-graphs on q vertices such that $\mathbf{T}$ accepts H if and only if there is $Q'\in\mathcal{Q}$ that is isomorphic to $H[Q]$.

As $\mathbf{T}$ is an (n,α)-tester, $\mathbf{Pr}(\mathcal{Q},H)\ge 2/3$ if H satisfies $\mathbf{P}$ and $\mathbf{Pr}(\mathcal{Q},H)\le 1/3$ if H is α-far from $\mathbf{P}$. The strategy is now to use Lemma 8. To this end, for a suitable small $\varepsilon>0$ and all $\mathbf{a}\in\mathbb{N}^{k-1}$ in a specified range (in terms of α, $q_k(\alpha)$ and k), we define a 'discretized' set $\mathbf{I}$ of regularity instances $(\varepsilon,\mathbf{a},d_{\mathbf{a},k})$ such that $d_{\mathbf{a},k}(\hat{\mathbf{x}})$ only attains a bounded number of possible values for all $\hat{\mathbf{x}}\in\hat{A}(k,k-1,\mathbf{a})$. Now setting $\mathcal{R}(n,\alpha):=\{R\in\mathbf{I}: IC(\mathcal{Q},d_{\mathbf{a},k})\ge 1/2\}$ leads to the desired result, as Lemma 8 implies $IC(\mathcal{Q},d_{\mathbf{a},k})\sim\mathbf{Pr}(\mathcal{Q},H)$ if H satisfies $(\varepsilon,\mathbf{a},d_{\mathbf{a},k})$. (In the actual argument, we consider some k-graph G obtained from the regular approximation lemma (Theorem 5) rather than H itself.)

B. Satisfying a regularity instance is testable

In this subsection we sketch how we prove that the property of satisfying a particular regularity instance is testable. This forms the main part of the proof of Theorem 2. Suppose H is a k-graph and Q is a subset of the vertices chosen uniformly at random. First we show that if H satisfies a regularity instance R, then with high probability $H[Q]$ is close to satisfying R. Also the converse is true: if H is far from satisfying R, then with high probability $H[Q]$ is also far from satisfying R.

The main tool for this is Lemma 9. It implies that a family of partitions not only transfers from a hypergraph to its random samples with high probability, but also vice versa. Crucially, in both directions these transfer results allow only a small additive increase in the regularity parameters.

Lemma 9. *Suppose $0<1/n<1/q\ll c\ll\delta\ll\varepsilon_0\le 1$ and $k\in\mathbb{N}\setminus\{1\}$. Suppose $R=(2\varepsilon_0/3,\mathbf{a},d_{\mathbf{a},k})$ is a regularity instance. Suppose H is a k-graph on vertex set V with $|V|=n$. Let $Q\in\binom{V}{q}$ be chosen uniformly at random. Then with probability at least $1-e^{-cq}$ the following hold.*

- *If there exists an $(\varepsilon_0,\mathbf{a},d_{\mathbf{a},k})$-equitable partition $\mathscr{O}_1$ of H, then there exists an $(\varepsilon_0+\delta,\mathbf{a},d_{\mathbf{a},k})$-equitable partition $\mathscr{O}_2$ of $H[Q]$.*
- *If there exists an $(\varepsilon_0,\mathbf{a},d_{\mathbf{a},k})$-equitable partition $\mathscr{O}_2$ of $H[Q]$, then there exists an $(\varepsilon_0+\delta,\mathbf{a},d_{\mathbf{a},k})$-equitable partition $\mathscr{O}_1$ of H.*

The key ingredient in the proof of Lemma 9 is Lemma 10, which, roughly speaking, states the following. Suppose there are two k-graphs H_1,H_2 with vertex set V_1,V_2, respectively, and there are two ε-equitable families of partitions of these k-graphs which have the same parameters. Suppose further that there is another ε_0-equitable family of partitions $\mathscr{O}_1$ for H_1. Then there is an equitable family of partitions $\mathscr{O}_2$ of H_2 which has (almost) the same parameters as $\mathscr{O}_1$ provided $\varepsilon\ll\varepsilon_0$. Even more loosely, the result says that if two hypergraphs share a single regularity partition, then they share any regularity partition.

For the proof of Lemma 10 we strengthen the regular approximation lemma (Theorem 5), but we omit the corresponding statement here.

Lemma 10. *Suppose* $0 < 1/n, 1/m \ll \varepsilon \ll 1/T, 1/a_1^{\mathscr{Q}} \ll \delta \ll \varepsilon_0 \leq 1$ *and* $k \in \mathbb{N} \setminus \{1\}$*. Suppose* $\mathbf{a}^{\mathscr{Q}} \in [T]^{k-1}$*. Suppose that* $R = (\varepsilon_0/2, \mathbf{a}^{\mathscr{O}}, d_{\mathbf{a}^{\mathscr{O}},k})$ *is a regularity instance. Suppose* V_1, V_2 *are sets of size* n, m*, and* H_1, H_2 *are* k*-graphs on* V_1, V_2*, respectively. Suppose*

- $\mathscr{Q}_1 = \mathscr{Q}_1(k-1, \mathbf{a}^{\mathscr{Q}})$ *is an* $(\varepsilon, \mathbf{a}^{\mathscr{Q}}, d_{\mathbf{a}^{\mathscr{Q}},k})$*-equitable partition of* H_1*,*
- $\mathscr{Q}_2 = \mathscr{Q}_2(k-1, \mathbf{a}^{\mathscr{Q}})$ *is an* $(\varepsilon, \mathbf{a}^{\mathscr{Q}}, d_{\mathbf{a}^{\mathscr{Q}},k})$*-equitable partition of* H_2*, and*
- $\mathscr{O}_1 = \mathscr{O}_1(k-1, \mathbf{a}^{\mathscr{O}})$ *is an* $(\varepsilon_0, \mathbf{a}^{\mathscr{O}}, d_{\mathbf{a}^{\mathscr{O}},k})$*-equitable partition of* H_1*.*

Then there exists an $(\varepsilon_0 + \delta, \mathbf{a}^{\mathscr{O}}, d_{\mathbf{a}^{\mathscr{O}},k})$*-equitable partition* $\mathscr{O}_2$ *of* H_2*.*

It is not difficult to deduce the following result from Lemma 9.

Theorem 11. *For all* $k \in \mathbb{N}\setminus\{1\}$ *and all regularity instances* $R = (\varepsilon, \mathbf{a}, d_{\mathbf{a},k})$*, the property of satisfying* R *is testable.*

C. The final step

We now aim to use Theorem 11 to show that (c)⇒(a) in Theorem 2, i.e. to prove that a regular reducible property $\mathbf{P}$ is also testable. As $\mathbf{P}$ is regular reducible, we can decide whether H satisfies $\mathbf{P}$ if we can test whether H is close to some regularity instance in a certain set $\mathcal{R}$. We strengthen Theorem 11 to show that the property of satisfying a given regularity instance R is actually estimable (the equivalence (a)⇔(b) is a by-product of this argument). Having proved this, it is straightforward to construct a tester for $\mathbf{P}$ by appropriately combining $|\mathcal{R}|$ estimators which estimate the distance of H and a given $R \in \mathcal{R}$.

V. Applications

In this section we illustrate how Theorem 2 can be applied. We first show how to test the (injective) homomorphism density, where a homomorphism of a k-graph F into a k-graph H is a function $f : V(F) \to V(H)$ that maps edges onto edges. Let $\mathrm{inj}(F, H)$ be the number of (vertex-)injective homomorphisms from F into H and let $t_{\mathrm{inj}}(F, H) := \mathrm{inj}(F, H)/(n)_{|V(F)|}$.

Corollary 12. *Suppose* $p, \delta \in (0, 1)$*,* $k \in \mathbb{N} \setminus \{1\}$*, and* F *is a* k*-graph. Let* $\mathbf{P}$ *be the property that a* k*-graph* H *satisfies* $t_{\mathrm{inj}}(F, H) = p \pm \delta$*. Then* $\mathbf{P}$ *is testable.*

Before we continue with the proof of Corollary 12, we state two simple propositions. We leave the proofs to the reader.

Proposition 13. *Suppose* $0 < 1/n \ll \nu, 1/k, 1/\ell$ *and* $\nu \ll \alpha, 1-\alpha$*. Let* F *be an* ℓ*-vertex* k*-graph and* H *be an* n*-vertex* k*-graph. If* $t_{\mathrm{inj}}(F, H) = \alpha \pm \nu$ *for some* $\alpha \in (0, 1)$*, then there exists an* n*-vertex* k*-graph* G *with* $t_{\mathrm{inj}}(F, G) = \alpha \pm 1/n$ *and* $|G \triangle H| \leq (\frac{2\nu}{\min\{\alpha, 1-\alpha\}})^{1/\ell} \binom{n}{k}$*.*

Proposition 14. *Suppose* $n, k, \ell \in \mathbb{N}$ *with* $k \leq \ell \leq n$ *and* G *and* H *are* n*-vertex* k*-graphs on vertex set* V *and* $\mathcal{F}$ *is a collection of* ℓ*-vertex* k*-graphs. If* $|G \triangle H| \leq \nu\binom{n}{k}$*, then*

$$\mathbf{Pr}(\mathcal{F}, G) = \mathbf{Pr}(\mathcal{F}, H) \pm \ell^k \nu.$$

Proof of Corollary 12: Let $\ell := |V(F)|$. We may assume that $|F| > 0$ as otherwise $t_{\mathrm{inj}}(F, H) = 1$ for every n-vertex graph H with $n \geq \ell$. By Theorem 2, it suffices to verify that $\mathbf{P}$ is regular reducible.

Suppose $\beta > 0$. We may assume that $\beta \ll p - \delta, 1/\ell$ if $p - \delta > 0$ and $\beta \ll 1 - (p + \delta), 1/\ell$ if $p + \delta < 1$. We write $\beta' := \beta^{\ell+1}$ and $\beta'' := 2^{-\binom{\ell}{k}}\beta'$. We fix some function $\overline{\varepsilon} : \mathbb{N}^{k-1} \to (0, 1)$ such that $\overline{\varepsilon}(\mathbf{a}) \ll \|\mathbf{a}\|_\infty^{-k}$ for all $(a_1, \ldots, a_{k-1}) = \mathbf{a} \in \mathbb{N}^{k-1}$. We choose constants ε, η, and $n_0, T \in \mathbb{N}$ such that $1/n_0 \ll \varepsilon \ll 1/T \ll \eta \ll \beta, 1/k, 1/\ell$. In particular, we have $n_0 \geq n_5(\eta, \beta''\ell^{-k}/2, \overline{\varepsilon})$, $T \geq t_5(\eta, \beta''\ell^{-k}/2, \overline{\varepsilon})$ and $\varepsilon \ll \overline{\varepsilon}(\mathbf{a})$ for all $\mathbf{a} \in [T]^{k-1}$. For simplicity, we consider only n-vertex k-graphs H with $n \geq n_0$.

Let $\mathbf{I}$ be the collection of regularity instances $R = (\varepsilon'', \mathbf{a}, d_{\mathbf{a},k})$ such that

- $\varepsilon'' \in \{\varepsilon, 2\varepsilon, \ldots, \lceil (\overline{\varepsilon}(\mathbf{a}))^{1/2}\varepsilon^{-1}\rceil\varepsilon\}$,
- $\|\mathbf{a}\|_\infty \leq T$ and $a_1 > \eta^{-1}$, and
- $d_{\mathbf{a},k}(\hat{\mathbf{x}}) \in \{0, \varepsilon^2, 2\varepsilon^2, \ldots, 1\}$ for every $\hat{\mathbf{x}} \in \hat{A}(k, k-1, \mathbf{a})$.

Observe that by construction $|\mathbf{I}|$ is bounded by a function of β, k and ℓ. We define

$$\mathcal{R} := \left\{(\varepsilon'', \mathbf{a}, d_{\mathbf{a},k}) \in \mathbf{I} : \sum_{J\,:\,|V(J)|=\ell} \mathrm{inj}(F, J) \cdot IC(J, d_{\mathbf{a},k})/\ell! = p \pm (\delta + \beta')\right\}.$$

First, suppose that an n-vertex k-graph H satisfies $\mathbf{P}$. Then

$$\frac{1}{\ell!} \sum_{J\,:\,|V(J)|=\ell} \mathrm{inj}(F, J) \cdot \mathbf{Pr}(J, H) = t_{\mathrm{inj}}(F, H) = p \pm \delta. \quad (2)$$

By applying the regular approximation lemma (Theorem 5) with $H, \eta, \beta''\ell^{-k}/2, \overline{\varepsilon}$ playing the roles of $H, \eta, \nu, \varepsilon$, we obtain a k-graph G and a family of partitions $\mathscr{P} = \mathscr{P}(k-1, \mathbf{a}^{\mathscr{P}})$ such that

(I) $\mathscr{P}$ is $(\eta, \overline{\varepsilon}(\mathbf{a}^{\mathscr{P}}), \mathbf{a}^{\mathscr{P}})$-equitable for some $\mathbf{a}^{\mathscr{P}} \in [T]^{k-1}$,
(II) G is $\overline{\varepsilon}(\mathbf{a}^{\mathscr{P}})$-regular with respect to $\hat{P}^{(k-1)}$ for every $\hat{P}^{(k-1)} \in \hat{\mathscr{P}}^{(k-1)}$, and
(III) $|G \triangle H| \leq \beta''\ell^{-k}\binom{n}{k}/2$.

Let $\varepsilon' := \overline{\varepsilon}(\mathbf{a}^{\mathscr{P}})$. By the choice of $\overline{\varepsilon}$ and η, we conclude that $0 < \varepsilon' \ll 1/\|\mathbf{a}^{\mathscr{P}}\|_\infty \leq 1/a_1^{\mathscr{P}} \ll \beta, 1/k, 1/\ell$ and by the choice of ε, we obtain $\varepsilon \ll \varepsilon'$. Note that if a k-graph J is (ε', d)-regular with respect to a $(k-1)$-graph J', then J is (ε'', d')-regular with respect to J' for some $d' \in \{0, \varepsilon^2, 2\varepsilon^2, \dots, 1\}$ and $\varepsilon'' \in \{\varepsilon, 2\varepsilon, \dots, \lceil \varepsilon'^{1/2}\varepsilon^{-1} \rceil \varepsilon\} \cap [2\varepsilon', 3\varepsilon']$. Thus there exists

$$R_G = (\varepsilon'', \mathbf{a}^{\mathscr{P}}, d^G_{\mathbf{a}^{\mathscr{P}},k}) \in \mathbf{I} \tag{3}$$

such that G satisfies R_G.

For every ℓ-vertex k-graph J, Proposition 14 with (III) and Lemma 8 imply that

$$IC(J, d^G_{\mathbf{a}^{\mathscr{P}},k}) = \mathbf{Pr}(J,G) \pm \beta''/2 = \mathbf{Pr}(J,H) \pm \beta''. \tag{4}$$

Hence

$$\begin{aligned} &\sum_{J:\,|V(J)|=\ell} \mathrm{inj}(F,J) IC(J, d^G_{\mathbf{a}^{\mathscr{P}},k})/\ell! \\ &\overset{(4)}{=} \sum_{J:\,|V(J)|=\ell} \mathrm{inj}(F,J)(\mathbf{Pr}(J,H) \pm \beta'')/\ell! \\ &\overset{(2)}{=} p \pm (\delta + \beta'). \end{aligned} \tag{5}$$

By the definition of $\mathcal{R}$ and (3), this implies that $R_G \in \mathcal{R}$ and so H is indeed β-close to a graph G satisfying R_G, one of the regularity instances of $\mathcal{R}$.

Now we show that if $\alpha > \beta$ and H is α-far from $\mathbf{P}$, then H is $(\alpha - \beta)$-far from all $R \in \mathcal{R}$. We prove this by verifying the following statement: if H is $(\alpha - \beta)$-close to some $R \in \mathcal{R}$, then it is α-close to $\mathbf{P}$.

Suppose H is $(\alpha - \beta)$-close to some $R = (\varepsilon'', \mathbf{a}, d_{\mathbf{a},k}) \in \mathcal{R}$. Then there exists a k-graph G_R such that G_R satisfies R and $|H \triangle G_R| \leq (\alpha - \beta)\binom{n}{k}$. By the definition of $\mathcal{R}$, we have $\sum_{J:\,|V(J)|=\ell} \mathrm{inj}(F,J) \cdot IC(J, d_{\mathbf{a},k})/\ell! = p \pm (\delta + \beta')$. Similarly to the calculations leading to (5), we obtain $t_{\mathrm{inj}}(F, G_R) = p \pm (\delta + 2\beta')$.

By Proposition 13, there exists a k-graph G such that $t_{\mathrm{inj}}(F,G) = p \pm \delta$ and $|G \triangle G_R| \leq (\beta/2) \cdot \binom{n}{k}$. Therefore, G satisfies $\mathbf{P}$ and $|H \triangle G| \leq |H \triangle G_R| + |G_R \triangle G| < \alpha\binom{n}{k}$ which implies that H is α-close to satisfying $\mathbf{P}$. Thus, $\mathbf{P}$ is indeed regular reducible. ■

We proceed with another corollary of Theorem 2. For a given n-vertex k-graph H, we define the following parameter measuring the size of a largest ℓ-partite subgraph:

$$\mathrm{maxcut}_\ell(H) := \binom{n}{k}^{-1} \max_{\substack{\{V_1,\dots,V_\ell\} \text{ is a} \\ \text{partition of } V(H)}} \{|\mathcal{K}_k(V_1,\dots,V_\ell) \cap H|\}.$$

We let

$$c_{\ell,k}(n) := \binom{n}{k}^{-1} \sum_{\Lambda \in \binom{[\ell]}{k}} \prod_{\lambda \in \Lambda} \left\lfloor \frac{n + \lambda - 1}{\ell} \right\rfloor.$$

Thus $c_{\ell,k}(n)\binom{n}{k}$ is the number of edges of the complete ℓ-partite k-graph on n vertices whose vertex class sizes are as equal as possible. In particular, any n-vertex k-graph H satisfies $\mathrm{maxcut}_\ell(H) \leq c_{\ell,k}(n)$.

Corollary 15. *Suppose $\ell, k \in \mathbb{N} \setminus \{1\}$ and $c = c(n)$ is such that $0 \leq c \leq c_{\ell,k}(n)$. Let $\mathbf{P}$ be the property that an n-vertex k-graph H satisfies $\mathrm{maxcut}_\ell(H) \geq c$. Then $\mathbf{P}$ is testable.*

Note that since the property of having a given edge density is trivially testable, it follows from Corollary 15 that the property of being strongly ℓ-colourable is also testable (in a strong colouring, we require all vertices of an edge to have distinct colours). A proof of Corollary 15 can be found in the full version of the paper. Finally, a natural question arising from Corollary 15 is whether $\mathbf{P}$ is in fact easily testable.

Acknowledgment

The research leading to these results was partially supported by the EPSRC, grant nos. EP/M009408/1 (F. Joos, D. Kühn and D. Osthus), and EP/N019504/1 (D. Kühn). The research was also partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013) / ERC Grant 306349 (J. Kim and D. Osthus).

References

[1] R. Rubinfeld and M. Sudan, "Robust characterizations of polynomials with applications to program testing," *SIAM J. Comput.*, vol. 25, pp. 252–271, 1996.

[2] O. Goldreich, S. Goldwasser, and D. Ron, "Property testing and its connection to learning and approximation," *J. ACM*, vol. 45, pp. 653–750, 1998.

[3] I. Z. Ruzsa and E. Szemerédi, "Triple systems with no six points carrying three triangles," in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, ser. Colloq. Math. Soc. János Bolyai. North-Holland, Amsterdam-New York, 1978, vol. 18, pp. 939–945.

[4] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, "Efficient testing of large graphs," *Combinatorica*, vol. 20, pp. 451–476, 2000.

[5] N. Alon and A. Shapira, "Every monotone graph property is testable," *SIAM J. Comput.*, vol. 38, pp. 505–522, 2008.

[6] ——, "A characterization of the (natural) graph properties testable with one-sided error," *SIAM J. Comput.*, vol. 37, pp. 1703–1727, 2008.

[7] O. Goldreich and L. Trevisan, "Three theorems regarding testing graph properties," *Random Structures Algorithms*, vol. 23, pp. 23–57, 2003.

[8] 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.*, vol. 39, pp. 143–167, 2009.

[9] E. Fischer, "The difficulty of testing for isomorphism against a graph that is given in advance," *SIAM J. Comput.*, vol. 34, pp. 1147–1158, 2005.

[10] P. Frankl and V. Rödl, "The uniformity lemma for hypergraphs," *Graphs Combin.*, vol. 8, pp. 309–312, 1992.

[11] V. Rödl and M. Schacht, "Regular partitions of hypergraphs: counting lemmas," *Combin. Probab. Comput.*, vol. 16, pp. 887–901, 2007.

[12] ——, "Regular partitions of hypergraphs: regularity lemmas," *Combin. Probab. Comput.*, vol. 16, pp. 833–885, 2007.

[13] V. Rödl and J. Skokan, "Regularity lemma for k-uniform hypergraphs," *Random Structures Algorithms*, vol. 25, pp. 1–42, 2004.

[14] W. T. Gowers, "Hypergraph regularity and the multidimensional Szemerédi theorem," *Ann. of Math.*, vol. 166, pp. 897–946, 2007.

[15] T. Tao, "A variant of the hypergraph removal lemma," *J. Combin. Theory Ser. A*, vol. 113, pp. 1257–1280, 2006.

[16] E. Fischer and I. Newman, "Testing versus estimation of graph properties," *SIAM J. Comput.*, vol. 37, pp. 482–501, 2007.

[17] V. Rödl and M. Schacht, "Property testing in hypergraphs and the removal lemma," in *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*. ACM, New York, 2007, pp. 488–495.

[18] ——, "Generalizations of the removal lemma," *Combinatorica*, vol. 29, pp. 467–501, 2009.

[19] C. Avart, V. Rödl, and M. Schacht, "Every monotone 3-graph property is testable," *SIAM J. Discrete Math.*, vol. 21, pp. 73–92, 2007.

[20] Y. Kohayakawa, B. Nagle, and V. Rödl, "Efficient testing of hypergraphs (extended abstract)," in *Automata, languages and programming*, ser. Lecture Notes in Comput. Sci. Springer, Berlin, 2002, vol. 2380, pp. 1017–1028.

[21] T. Austin and T. Tao, "Testability and repair of hereditary hypergraph properties," *Random Structures Algorithms*, vol. 36, pp. 373–463, 2010.

[22] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy, and K. Vesztergombi, "Graph limits and parameter testing," in *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. ACM, New York, 2006, pp. 261–270.

[23] L. Lovász, *Large networks and graph limits*, ser. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2012, vol. 60.

[24] L. Lovász and B. Szegedy, "Testing properties of graphs and functions," *Israel J. Math.*, vol. 178, pp. 113–156, 2010.

[25] G. Elek and B. Szegedy, "A measure-theoretic approach to the theory of dense hypergraphs," *Adv. Math.*, vol. 231, pp. 1731–1772, 2012.

[26] L. Lovász and K. Vesztergombi, "Non-deterministic graph property testing," *Combin. Probab. Comput.*, vol. 22, pp. 749–762, 2013.

[27] M. Karpinski and R. Markó, "On the complexity of nondeterministically testable hypergraph parameters," *arXiv:1503.07093*, 2015.

[28] G. Lior and A. Shapira, "Deterministic vs non-deterministic graph property testing," *Israel J. Math.*, vol. 204, pp. 397–416, 2014.

[29] M. Karpinski and R. Markó, "Explicit bounds for nondeterministically testable hypergraph parameters," *arXiv:1509.03046*, 2015.

[30] W. T. Gowers, "Lower bounds of tower type for Szemerédi's uniformity lemma," *Geom. Funct. Anal.*, vol. 7, pp. 322–337, 1997.

[31] N. Alon and A. Shapira, "Linear equations, arithmetic progressions, and hypergraph property testing," *Theory Comput.*, vol. 1, pp. 177–216, 2005.

[32] N. Alon and J. Fox, "Easily testable graph properties," *Combin. Probab. Comput.*, vol. 24, pp. 646–657, 2015.

[33] J. Fox, J. Pach, and A. Suk, "A polynomial regularity lemma for semialgebraic hypergraphs and its applications in geometry and property testing," *SIAM J. Comput.*, vol. 45, pp. 2199–2223, 2016.

[34] N. Alon and O. Ben-Eliezer, "Removal lemmas for matrices," *arXiv:1609.04235*, 2016.

[35] N. Alon and E. Blais, "Testing Boolean function isomorphism," in *Approximation, randomization, and combinatorial optimization*, ser. Lecture Notes in Comput. Sci. Springer, Berlin, 2010, vol. 6302, pp. 394–405.

[36] N. Alon and A. Shapira, "Testing satisfiability," *J. Algorithms*, vol. 47, pp. 87–103, 2003.

[37] N. Alon, S. Dar, M. Parnas, and D. Ron, "Testing of clustering," *SIAM J. Discrete Math.*, vol. 16, pp. 393–417, 2003.

[38] M. Ben Or, D. Coppersmith, M. Luby, and R. Rubinfeld, "Non-abelian homomorphism testing, and distributions close to their self-convolutions," *Random Structures Algorithms*, vol. 32, pp. 49–70, 2008.

[39] K. Friedl, G. Ivanyos, and M. Santha, "Efficient testing of groups," in *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*. ACM, New York, 2005, pp. 157–166.

[40] I. Benjamini, O. Schramm, and A. Shapira, "Every minor-closed property of sparse graphs is testable," *Adv. Math.*, vol. 223, pp. 2200–2218, 2010.

[41] I. Newman and C. Sohler, "Every property of hyperfinite graphs is testable," *SIAM J. Comput.*, vol. 42, pp. 1095–1112, 2013.

[42] Y. Kohayakawa, V. Rödl, and J. Skokan, "Hypergraphs, quasi-randomness, and conditions for regularity," *J. Combin. Theory Ser. A*, vol. 97, pp. 307–352, 2002.