A characterization of testable hypergraph properties

Deryk Osthus — **University of Birmingham**
Felix Joos — **University of Birmingham** **Jaehoon Kim** — **University of Birmingham**
Daniela Kühn — **University of Birmingham**

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.

1 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. 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.

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:

1. if H satisfies $\mathbf{P}$, then $\mathbf{T}$ accepts H with probability at least $2/3$,
2. 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 [22]. In the graph setting, the earliest systematic results were obtained in a seminal paper of Goldreich, Goldwasser and Ron [12]. 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 [23], 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 [1]) and generalized first to monotone properties (which are closed under vertex and edge deletion) by Alon and Shapira [6] and then to hereditary properties (which are closed under vertex deletion), again by Alon and Shapira [5]. Examples of non-testable properties include some properties which are closed under edge deletion [13] and the property of being isomorphic to a given graph G [2, 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 [2] who obtained a combinatorial characterization of all testable graph properties. This solved a problem posed already by [12], which was regarded as one of the main open problems in the area.

The characterization proved in [2] 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 characterized 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 1) shows that this can be extended to hypergraphs of higher uniformity. Our characterization is based on the concept of (strong) hypergraph regularity, which was introduced in the ground-breaking work of Rödl et al. [11, 19, 20, 21], Gowers [15], see also Tao [24]. We defer the precise definition of regular reducibility for k-graphs to [16] as the concept of (strong) hypergraph regularity 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 [20], 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 as follows.

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

- if H is $(\alpha - \beta)$-close to satisfying $\mathbf{P}$, then $\mathbf{T}$ accepts H with probability at least $2/3$,
- 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 an *estimator*, or (n, α, β)-*estimator* for $\mathbf{P}$. We also say that $\mathbf{T}$ has query complexity q'_k. The property $\mathbf{P}$ is *estimable* if it is estimable with query complexity at most q_k for some function $q_k : (0,1)^2 \to \mathbb{N}$.

We show that testability and estimability are in fact equivalent. For graphs this goes back to Fischer and Newman [10].

Theorem 1 ([16])**.** *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 [16] we also illustrate how Theorem 1 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 [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 (i) of the definition of testability).

The result of Alon and Shapira on the testability of hereditary graph properties was strengthened by Austin and Tao [7] in another direction: they showed that hereditary properties of graphs are not only testable with one-sided error, but they are also *locally repairable* (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 1.

An intimate connection between property testing and graph limits was established by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi [8]. 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 $\mathbf{P}$. Another characterization (in terms of localized samples) using the graph limit framework was given by Lovász and Szegedy [17] (none of these results however yield effective bounds on the query complexity).

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 1 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 [14].) For k-graphs, Alon and Shapira [4] as well as Alon and Fox [3] obtained positive and negative results for the property of containing a given k-graph as an (induced) subgraph.

References

[1] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), 451–476.

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

[3] N. Alon and J. Fox, *Easily testable graph properties*, Combin. Probab. Comput. **24** (2015), 646–657.

[4] N. Alon and A. Shapira, *Linear equations, arithmetic progressions, and hypergraph property testing*, Theory Comput. **1** (2005), 177–216.

[5] N. Alon and A. Shapira, *A characterization of the (natural) graph properties testable with one-sided error*, SIAM J. Comput. **37** (2008), 1703–1727.

[6] N. Alon and A. Shapira, *Every monotone graph property is testable*, SIAM J. Comput. **38** (2008), 505–522.

[7] T. Austin and T. Tao, *Testability and repair of hereditary hypergraph properties*, Random Structures Algorithms **36** (2010), 373–463.

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

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

[10] E. Fischer and I. Newman, *Testing versus estimation of graph properties*, SIAM J. Comput. **37** (2007), 482–501.

[11] P. Frankl and V. Rödl, *The uniformity lemma for hypergraphs*, Graphs Combin. **8** (1992), 309–312.

[12] O. Goldreich, S. Goldwasser, and D. Ron, *Property testing and its connection to learning and approximation*, J. ACM **45** (1998), 653–750.

[13] O. Goldreich and L. Trevisan, *Three theorems regarding testing graph properties*, Random Structures Algorithms **23** (2003), 23–57.

[14] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), 322–337.

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

[16] F. Joos, J. Kim, D. Kühn and D. Osthus, *A characterization of testable hypergraph properties*, submitted (an extended abstract of this paper appeared at FOCS 2017).

[17] L. Lovász and B. Szegedy, *Testing properties of graphs and functions*, Israel J. Math. **178** (2010), 113–156.

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

[19] V. Rödl and M. Schacht, *Regular partitions of hypergraphs: counting lemmas*, Combin. Probab. Comput. **16** (2007), 887–901.

[20] V. Rödl and M. Schacht, *Regular partitions of hypergraphs: regularity lemmas*, Combin. Probab. Comput. **16** (2007), 833–885.

[21] V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures Algorithms **25** (2004), 1–42.

[22] R. Rubinfeld and M. Sudan, *Robust characterizations of polynomials with applications to program testing*, SIAM J. Comput. **25** (1996), 252–271.

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

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