WEAK HYPERGRAPH REGULARITY AND LINEAR HYPERGRAPHS

Y. KOHAYAKAWA, B. NAGLE, V. RÖDL, AND M. SCHACHT

Abstract. In this note, we consider conditions which allow the embedding of linear hypergraphs of fixed size. In particular, we prove that any k-uniform hypergraph H of positive uniform density contains all linear k-uniform hypergraphs of a given size. The main ingredient in the proof of this result is a *counting lemma* for linear hypergraphs, which establishes that the straightforward extension of graph ε-regularity to hypergraphs suffices for counting linear hypergraphs. We also consider some related problems.

1. Introduction and results

A graph $G=(V,E)$ is said to be (ϱ,d)-*quasirandom* if any subset $U\subseteq V$ of size $|U|\geq\varrho|V|$ induces $(d\pm\varrho)\binom{|U|}{2}$ edges. Such graphs, first systematically studied by Thomason [20, 21] and Chung, Graham, and Wilson [2], share several properties with genuine random graphs of the same edge density. For example, it was shown that if $\varrho=\varrho(d,\ell)$ is sufficiently small, then any (ϱ,d)-quasirandom graph G is ℓ-universal, meaning that G contains approximately the same number of copies of any ℓ-vertex graph F as the random graph of the same density.

Theorem 1. *For every graph F, every $d>0$ and every $\gamma>0$, there exist $\varrho>0$ and n_0 so that any (ϱ,d)-quasirandom graph G on $n\geq n_0$ vertices contains $(1\pm\gamma)d^{e_F}n^{v_F}$ labeled copies of F.*

As usual, in the result above we write e_F for the number of edges in F and we write v_F for the number of vertices in F. In this note, we address the extent to which Theorem 1 can be generalized to hypergraphs.

Definition 2. *A k-uniform hypergraph $H=(V,E)$ is (ϱ,d)-**quasirandom** if for any subset $U\subseteq V$ of size $|U|\geq\varrho|V|$, we have $e_H(U)=(d\pm\varrho)\binom{|U|}{k}$.*

It is known that Theorem 1 does not generally extend to k-uniform hypergraphs, for $k\geq 3$. Indeed, let F_0 be the 3-uniform hypergraph consisting of two triples intersecting in two vertices, and consider the following two (ϱ,d)-quasirandom n-vertex hypergraphs H_1 and H_2. Let $H_1=G^{(3)}(n,1/8)$ be the random 3-uniform hypergraph on n vertices whose triples appear independently with probability 1/8. Let $H_2=K_3(G(n,1/2))$ be the 3-uniform hypergraph whose triples correspond to triangles of the random graph $G(n,1/2)$ on n vertices, where the edges of $G(n,1/2)$

Date: February 22, 2008.

The first author was partially supported by FAPESP and CNPq through a Temático-ProNEx project (Proc. FAPESP 2003/09925-5) and by CNPq (Proc. 308509/2007-2, 485671/2007-7 and 486124/2007-0).

The second author was supported by NSF grant DMS 0639839.

The third author was supported by NSF grant DMS 0300529.

appear independently with probability 1/2. It is easy to check that, w.h.p., both H_1 and H_2 are $(\varrho, 1/8)$-quasirandom for any $\varrho > 0$. However, w.h.p., H_1 contains $(1 \pm o(1))n^4/64$ copies of F_0, while H_2 contains $(1 \pm o(1))n^4/32$ such copies, approximately twice as many.

The hypergraph F_0, while very elementary, has one property which causes the extension of Theorem 1 to fail: it contains two vertices belonging to more than one edge. We will show that removing this "obstacle" allows an extension of Theorem 1.

Definition 3. *We say a k-uniform hypergraph F is* ***linear*** *if $|e \cap f| \leq 1$ for all distinct edges e and f of F. We denote by $\mathcal{L}^{(k)}$ the family of all k-uniform, linear hypergraphs and set $\mathcal{L}^{(k)}_\ell = \{F \in \mathcal{L}^{(k)} \colon v_F \leq \ell\}$.*

Theorem 4. *For every integer $k \geq 2$, $d > 0$ and $\gamma > 0$, and every $F \in \mathcal{L}^{(k)}$, there exist $\varrho > 0$ and n_0 so that any (ϱ, d)-quasirandom k-uniform hypergraph $H = (V, E)$ on $n \geq n_0$ vertices contains $(1 \pm \gamma) d^{e_F} n^{v_F}$ labeled copies of F.*

We will also consider some other related results that extend known graph results to hypergraphs in a similar way to how Theorem 4 extends Theorem 1.

Definition 5. *A k-uniform hypergraph $H = (V, E)$ is (ϱ, d)-**dense** if for any subset $U \subseteq V$ of size $|U| \geq \varrho |V|$, we have $e_H(U) \geq d\binom{|U|}{k}$.*

For graphs, a simple induction on $\ell \geq 2$ shows that every (ϱ, d)-dense graph on sufficiently many vertices contains a copy of K_ℓ, as long as $\varrho \leq d^{\ell-2}$. However, the analogous statement for $k \geq 3$ fails. Indeed, let T_n be a tournament on n vertices chosen uniformly at random, and let $H = H(T_n)$ be the 3-uniform hypergraph whose triples correspond to directed triangles of T_n. Then, w.h.p., H is (ϱ, d)-dense for any $\varrho > 0$ and $0 < d < 1/4$. (In fact, H is $(\varrho, 1/4)$-quasirandom.) However, since every tournament on four vertices contains at most two directed triangles, H is $K_4^{(3)}$-free. (In fact, H does not even contain three triples on any four vertices.) In this note, we prove that, on the other hand, a (ϱ, d)-dense hypergraph H will contain (many) copies of linear hypergraphs of fixed size.

Definition 6. *For integers $\ell \geq k$ and $\xi > 0$, we say a k-uniform hypergraph $H = (V, E)$ is $(\xi, \mathcal{L}^{(k)}_\ell)$-**universal** if the number of copies of any $F \in \mathcal{L}^{(k)}_\ell$ is at least $\xi |V|^\ell$.*

Theorem 7. *For all integers $\ell \geq k \geq 2$ and every $d > 0$, there exist $\varrho = \varrho(\ell, k, d) > 0$, $\xi = \xi(\ell, k, d) > 0$, and $n_0 = n_0(\ell, k, d)$ so that every (ϱ, d)-dense k-uniform hypergraph $H = (V, E)$ on $n \geq n_0$ vertices is $(\xi, \mathcal{L}^{(k)}_\ell)$-universal.*

We shall also prove an easy corollary of Theorem 7 (upcoming Corollary 8), which roughly asserts the following. Suppose $H = (V, E)$ is a 'non-universal' hypergraph of density d. We prove that V may be partitioned into nearly equal-sized classes $V_1, \dots, V_t$ so that the number of edges of H crossing at least two such classes is slightly larger than it would be expected if $V = V_1 \dot\cup \dots \dot\cup V_t$ were a random partition. More precisely, for $t \in \mathbb{N}$, let $\tau_t(H)$ be the *maximal t-cut-density* of H, defined by

$$\tau_t(H) = \max\{\hat{d}_H(U_1, \dots, U_t) \colon U_1 \dot\cup \dots \dot\cup U_t = V \text{ and } |U_1| \leq \dots \leq |U_t| \leq |U_1| + 1\},$$

where

$$\hat{d}_H(U_1, \dots, U_t) = \frac{|E(H) \setminus \bigcup_{i=1}^t \binom{U_i}{k}|}{\binom{|V|}{k} - \sum_{i=1}^t \binom{|U_i|}{k}}.$$

Corollary 8. *For all integers $\ell \geq k \geq 2$ and every $d > 0$, there exist $t \in \mathbb{N}$, $\beta = \beta(\ell, k, d), \xi = \xi(\ell, k, d) > 0$ and $n_0 = n_0(\ell, k, d)$ so that every k-uniform hypergraph $H = (V, E)$ on $n \geq n_0$ vertices and $e_H \geq d\binom{n}{k}$ edges satisfies the following. If H is not $(\xi, \mathcal{L}_\ell^{(k)})$-universal, then $\tau_t(H) \geq d + \beta$.*

Corollary 8 is somewhat related to a result from [11] and its strengthening due to Nikiforov [10].

2. Tools

A key tool we use in this paper is the so-called *weak hypergraph regularity lemma.* This result is a straightforward extension of Szemerédi's regularity lemma [18] for graphs. Let $H = (V, E)$ be a k-uniform hypergraph and let $W_1, \ldots, W_k$ be mutually disjoint non-empty subsets of V. We denote by $d_H(W_1, \ldots, W_k) = d(W_1, \ldots, W_k)$ the *density* of the k-partite induced subhypergraph $H[W_1, \ldots, W_k]$ of H, defined by

$$d_H(W_1, \ldots, W_k) = \frac{e_H(W_1, \ldots, W_k)}{|W_1| \cdot \ldots \cdot |W_k|}\,.$$

We say the k-tuple $(V_1, \ldots, V_k)$ of mutually disjoint subsets $V_1, \ldots, V_k \subseteq V$ is (ε, d)*-regular*, for positive constants ε and d, if

$$|d_H(W_1, \ldots, W_k) - d| \leq \varepsilon$$

for all k-tuples of subsets $W_1 \subseteq V_1, \ldots, W_k \subseteq V_k$ satisfying $|W_1| \cdot \ldots \cdot |W_k| \geq \varepsilon |V_1| \cdot \ldots \cdot |V_k|$. Note, in particular, that if $(V_1, \ldots, V_k)$ is (ε, d)-regular, then

$$\big| H[W_1, \ldots, W_k] - d|W_1| \cdot \ldots \cdot |W_k| \big| \leq \varepsilon |V_1| \cdot \ldots \cdot |V_k| \tag{1}$$

holds for *any* $W_1 \subseteq V_1, \ldots, W_k \subseteq V_k$. We say the k-tuple $(V_1, \ldots, V_k)$ is ε*-regular* if it is (ε, d)-regular for some $d \geq 0$. The weak regularity lemma then states the following.

Theorem 9. *For all integers $k \geq 2$ and $t_0 \geq 1$, and every $\varepsilon > 0$, there exist $T_0 = T_0(k, t_0, \varepsilon)$ and $n_0 = n_0(k, t_0, \varepsilon)$ so that for every k-uniform hypergraph $H = (V, E)$ on $n \geq n_0$ vertices, there exists a partition $V = V_0 \dot\cup V_1 \dot\cup \ldots \dot\cup V_t$ so that the following hold:*

- (*i*) $t_0 \leq t \leq T_0$,
- (*ii*) $|V_0| \leq \varepsilon n$ *and* $|V_1| = \cdots = |V_t|$, *and*
- (*iii*) *for all but at most* $\varepsilon\binom{t}{k}$ *sets* $\{i_1, \ldots, i_k\} \subseteq [t]$, *the k-tuple* $(V_{i_1}, \ldots, V_{i_k})$ *is* ε*-regular.*

The proof of Theorem 9 follows the lines of the original proof of Szemerédi [18] (for details see e.g. [1, 3, 17]).

A key feature of Szemerédi's regularity lemma is the so-called *counting lemma.* This lemma provides good estimates on the number of subgraphs of a fixed isomorphism type in an appropriate collection of ε-regular pairs. To be precise, let F be a graph (hypergraph) on the vertex set $[\ell]$ and let G be an ℓ-partite graph (hypergraph) with vertex partition $V(G) = V_1 \dot\cup \ldots \dot\cup V_\ell$. A copy F_0 of F in G, on the vertices $v_1 \in V_1, \ldots, v_\ell \in V_\ell$, is said to be *partite-isomorphic* to F if $i \mapsto v_i$ defines a homomorphism. The counting lemma for graphs asserts that if (V_i, V_j) is (ε, d_{ij})-regular, where $d_{ij}^\ell \gg \varepsilon > 0$ whenever $\{i, j\} \in E(F)$, then the number of labeled partite-isomorphic copies F_0 of F in G is within the interval $(1 \pm \gamma) \prod_{\{i,j\} \in E(F)} d_{ij} \prod_{i \in [\ell]} |V_i|$, where $\gamma \to 0$ as $\varepsilon \to 0$. It is known that this fact does not extend to k-uniform hypergraphs ($k \geq 3$), and that stronger regularity lemmas are needed in that case (see, e.g., [5, 9, 12, 13, 19]). However, weak regularity is sufficient for estimating the number of linear subhypergraphs in an appropriately ε-regular environment.

Lemma 10 (Counting lemma for linear hypergraphs)**.** *For all integers $\ell \geq k \geq 2$ and every $\gamma, d_0 > 0$, there exist $\varepsilon = \varepsilon(\ell, k, \gamma, d_0) > 0$ and $m_0 = m_0(\ell, k, \gamma, d_0)$ so that the following holds.*

Let $S = ([\ell], F) \in \mathcal{L}_\ell^{(k)}$ and let $H = (V_1 \dot\cup \ldots \dot\cup V_\ell, E)$ be an ℓ-partite, k-uniform hypergraph where $|V_1|, \ldots, |V_\ell| \geq m_0$. Suppose, moreover, that for all edges $f \in F$, the k-tuple $(V_i)_{i \in f}$ is (ε, d_f)-regular, where $d_f \geq d_0$. Then the number of partite-isomorphic copies of S in H is within the interval

$$(1 \pm \gamma) \prod_{f \in F} d_f \prod_{i \in [\ell]} |V_i| \,.$$

Proof. Let integers $\ell \geq k \geq 2$ and $\gamma, d_0 > 0$ be fixed. We shall prove, by induction on $|F|$, the number of edges of S, that $\varepsilon = \gamma(d_0/2)^{|F|}$ will suffice to count copies of S (with 'precision' γ), provided m_0 is large enough. (In this way, $\varepsilon = \gamma(d_0/2)^{\binom{\ell}{2}}$ works for all $S \in \mathcal{L}_\ell^{(k)}$.) If $|F| = 0$ or $|F| = 1$, the result is trivial. It is also easy to see that the result holds whenever S consists of pairwise disjoint edges, since then the number of partite-isomorphic copies of S in H is within

$$\prod_{f \in F} (d_f \pm \varepsilon) \prod_{i \in [\ell]} |V_i| = (1 \pm (\varepsilon/d_0))^{|F|} \prod_{f \in F} d_f \prod_{i \in [\ell]} |V_i| = (1 \pm \gamma) \prod_{f \in F} d_f \prod_{i \in [\ell]} |V_i|.$$

Now, generally, take m_0 large enough so that we can apply the induction assumption on $|F| - 1$ edges with precision $\gamma/2$ and d_0 (and note that $\varepsilon = \gamma(d_0/2)^{|F|} < (\gamma/2)(d_0/2)^{|F|-1}$). All copies of various subhypergraphs disussed below are tacitly assumed to be partite-isomorphic.

Let $S = ([\ell], F) \in \mathcal{L}_\ell^{(k)}$ have $|F| \geq 2$ edges and let $H = (V, E)$ be a k-uniform hypergraph satisfying the assumptions of Lemma 10. Fix an edge $e \in F$ and set $S_- = ([\ell], F \setminus \{e\})$ to be the hypergraph obtained from S by removing the edge e. Moreover, for a copy T_- of S_- in H, we denote by e_{T_-} the unique k-tuple of vertices which together with T_- forms a copy of S in H. Furthermore, let $1_E \colon \binom{V}{k} \to \{0,1\}$ be the indicator function of the edge set E of H. In this notation, a copy T_- of S_- in H extends to a copy of S if, and only if, $1_E(e_{T_-}) = 1$. Consequently, summing over all copies T_- of S_- in H, we can count the number $\#\{S \subseteq H\}$ of copies of S in H by

$$\begin{aligned} \#\{S \subseteq H\} &= \sum_{T_- \subseteq H} 1_E(e_{T_-}) = \sum_{T_- \subseteq H} (d_e + 1_E(e_{T_-}) - d_e) \\ &= d_e \times \#\{S_- \subseteq H\} + \sum_{T_- \subseteq H} (1_E(e_{T_-}) - d_e) \\ &= (1 \pm \tfrac{\gamma}{2}) \prod_{f \in F} d_f \prod_{i \in [l]} |V_i| + \sum_{T_- \subseteq H} (1_E(e_{T_-}) - d_e) \,, \end{aligned} \tag{2}$$

where we used the induction assumption for S_- for the last estimate.

It is left to bound the error term $\sum_{T_- \subseteq H} 1_E(e_{T_-}) - d_e$ in (2). For that, we will appeal to the regularity of $(V_i)_{i \in e}$. Let $S_* = S[[\ell] \setminus e]$ be the induced subhypergraph of S obtained by removing the vertices of e and all edges of S intersecting e. For a copy T_* of S_* in H, let $\mathrm{ext}(T_*)$ be the set of k-tuples $K \in \prod_{i\in e} V_i$ such that $V(T_*)\dot\cup K$ spans a copy of T_- in H. Since S is a linear hypergraph, we have $|f \cap e| \le 1$ for every edge f of S_-. Hence, for every $i \in e$, there exists a subset $W_i^{T_*} \subseteq V_i$ such that

$$\mathrm{ext}(T_*) = \prod_{i\in e} W_i^{T_*}\,.$$

Indeed, for every $i \in e$, the set $W_i^{T_*}$ consists of those vertices $v \in V_i$ with the property that $V(T_*)\dot\cup\{v\}$ spans a copy of S induced on $V(S_*)\dot\cup\{i\}$ in H. With this notation, we can bound the error term in (2) as follows:

$$\begin{aligned}\Big|\sum_{T_-\subseteq H} 1_E(e_{T_-}) - d_e\Big| &\le \sum_{T_*\subseteq H}\Big|\sum_{K\in\mathrm{ext}(T_*)} 1_E(K) - d_e\Big|\\ &= \sum_{T_*\subseteq H}\Big|\sum\Big\{1_E(K) - d_e\colon\ K \in \prod_{i\in e} W_i^{T_*}\Big\}\Big| \le \sum_{T_*\subseteq H}\varepsilon\prod_{i\in e}|V_i|\,,\end{aligned}$$

where the ε-regularity was used for the last estimate. Indeed, for a fixed copy $T_* \subseteq H$, we have

$$\Big|\sum\Big\{1_E(K) - d_e\colon\ K\in\prod_{i\in e} W_i^{T_*}\Big\}\Big| = \Big|\big|H\cap\prod_{i\in e} W_i^{T_*}\big| - d_e\prod_{i\in e}\big|W_i^{T_*}\big|\Big|,$$

so that we may appeal to (1). Now, because of the choice of ε we have

$$\Big|\sum_{T_-\subseteq H} 1_E(e_{T_-}) - d_e\Big| \le \varepsilon\sum_{T_*\subseteq H}\prod_{i\in e}|V_i| \le \varepsilon\prod_{i\in[\ell]}|V_i| \le \frac{\gamma}{2}\prod_{f\in F} d_f\prod_{i\in[\ell]}|V_i|\,,$$

and Lemma 10 follows from (2). □

3. Quasirandom hypergraphs

In this section, we prove Theorem 4 according to the following outline. We first observe that a (ϱ, d)-quasirandom (k-uniform) hypergraph H is (ε, d)-regular w.r.t. any disjoint family $U_1, \dots, U_k \subset V(H)$ of large and equal-sized sets. As such, any partition $U_1\dot\cup\dots\dot\cup U_\ell$ within $V(H)$ of $\ell \ge k$ large equal-sized sets will satisfy the hypothesis of the counting lemma (Lemma 10), and will therefore contain the "right" number of copies of any hypergraph $F \in \mathcal{L}_\ell^{(k)}$. Applying this argument to a partition chosen at random then yields the "right" number of copies of F in H.

Proof of Theorem 4. Let $k \ge 2$, $d, \gamma > 0$ and $F \in \mathcal{L}^{(k)}$ on the vertex set $\{1, \dots, \ell\}$ be given. We set

$$\varepsilon = \varepsilon(\ell, k, \gamma/2, d) \quad\text{and}\quad \varrho = \frac{\varepsilon^2}{\ell(2k)^k} \tag{3}$$

and let $n \ge m_0(\ell, k, \gamma/2, d)/\varrho$ be sufficiently large, where the constants $\varepsilon(\ell, k, \gamma/2, d)$ and $m_0(\ell, k, \gamma/2, d)$ are given by Lemma 10. Let H be a (ϱ, d)-quasirandom k-uniform hypergraph on n vertices.

Following the outline (above), let $U_i \subset V$, $1 \le i \le k$, be mutually disjoint sets of size $|U_i| = m \ge \varrho n/\varepsilon$. We claim that $(U_1, \dots, U_k)$ is (ε, d)-regular w.r.t. H. Indeed, let $V_i \subseteq U_i$, $1 \le i \le k$, be given so that $|V_1| \cdot \ldots \cdot |V_k| \ge \varepsilon m^k$. (Note, in particular, that this implies $|V_i| \geq \varepsilon m \geq \varrho n$ for all $1 \leq i \leq k$.) To show that $|H[V_1, \dots, V_k]| = (d \pm \varepsilon)|V_1| \cdot \dots \cdot |V_k|$, we observe, from inclusion-exclusion, that

$$|H[V_1, \dots, V_k]| = \sum_{I \subseteq [k]} (-1)^{|I|} \Big| H\Big[\bigcup_{j \in [k] \setminus I} V_j \Big] \Big|.$$

The (ϱ, d)-quasi-randomness of H (together with $|V_i| \geq \varrho n$ for all $1 \leq i \leq k$) implies

$$\begin{aligned}
|H[V_1, \dots, V_k]| &= \sum_{I \subseteq [k]} (-1)^{|I|} (d \pm \varrho) \binom{\big|\bigcup_{j \in [k] \setminus I} V_j\big|}{k} \\
&= d \sum_{I \subseteq [k]} (-1)^{|I|} \binom{\big|\bigcup_{j \in [k] \setminus I} V_j\big|}{k} \pm \varrho \sum_{I \subseteq [k]} \binom{\big|\bigcup_{j \in [k] \setminus I} V_j\big|}{k} \\
&= d \sum_{I \subseteq [k]} (-1)^{|I|} \binom{\big|\bigcup_{j \in [k] \setminus I} V_j\big|}{k} \pm \varrho (2k)^k m^k \\
&= d |V_1| \cdot \dots \cdot |V_k| \pm \varrho (2k)^k m^k \\
&= \big(d \pm \varrho (2k)^k / \varepsilon\big) |V_1| \cdot \dots \cdot |V_k| \\
&= (d \pm \varepsilon) |V_1| \cdot \dots \cdot |V_k|,
\end{aligned}$$

where the first term in the line above was obtained by inclusion-exclusion.

To finish the proof of Theorem 4, consider an ℓ-tuple of mutually disjoint sets $U_1, \dots, U_\ell$ with $|U_1| = \dots = |U_\ell| = m$, where m is a fixed integer satisfying $n/\ell \geq m \geq \varrho n / \varepsilon$. Then every k-tuple $I \in \binom{[\ell]}{k}$ satisfies that $(U_i)_{i \in I}$ is (ε, d)-regular (as shown above), and so by the choice of ε in (3), we can apply the counting lemma for linear hypergraphs (Lemma 10) to $U_1 \dot\cup \dots \dot\cup U_\ell$. Consequently, $H[U_1, \dots, U_\ell]$ contains $(1 \pm \gamma/2) d^{e_F} m^\ell$ partite-isomorphic copies of F (recall $V(F) = [\ell]$). Now, on the one hand, we note that there are $\binom{n}{m}\binom{n-m}{m} \dots \binom{n-(\ell-1)m}{m}$ choices for the partition $U_1 \dot\cup \dots \dot\cup U_\ell$. On the other hand, for each ℓ-tuple of vertices $(u_1, \dots, u_\ell)$ in $V(H)$, there are $\binom{n-\ell}{m-1}\binom{n-m-(\ell-1)}{m-1} \dots \binom{n-(\ell-1)m-1}{m-1}$ such partitions $U_1 \dot\cup \dots \dot\cup U_\ell$ for which $(u_1, \dots, u_\ell) \in U_1 \times \dots \times U_\ell$. Consequently, the number of labeled copies of F in H is given by

$$\begin{aligned}
&(1 \pm \gamma/2) d^{e_F} m^\ell \frac{\binom{n}{m}\binom{n-m}{m} \dots \binom{n-(\ell-1)m}{m}}{\binom{n-\ell}{m-1}\binom{n-m-(\ell-1)}{m-1} \dots \binom{n-(\ell-1)m-1}{m-1}} \\
&\qquad\qquad = (1 \pm \gamma/2) d^{e_F} \frac{n!}{(n-\ell)!} = (1 \pm \gamma) d^{e_F} n^{v_F},
\end{aligned}$$

where for the last step we use that n is sufficiently large. □

4. Universal hypergraphs

In this section, we prove Theorem 7. The proof relies on the weak hypergraph regularity lemma, which allows us to locate a sufficiently dense and ε-regular ℓ-partite subhypergraph in any (ϱ, d)-dense hypergraph. The $(\xi, \mathcal{L}_\ell^{(k)})$-universality then follows from Lemma 10.

Proof of Theorem 7. Let integers $\ell \geq k \geq 2$ and $d > 0$ be given. To define the promised constants ϱ and ξ, we first consider a few auxiliary constants. Set $d_0 = d/(4k!)$ and $q = \lceil 1/d_0 \rceil$ and let $s = r_k(q, \ell)$ be the (k-uniform) Ramsey number for q and ℓ, i.e., s is the smallest integer s.t. any 2-coloring of $E(K_s^{(k)})$ yields a copy of $K_q^{(k)}$ in the first color, or a copy of $K_\ell^{(k)}$ in the second color. Set $\varepsilon = \min\left\{1/(2\binom{s}{k}),\ \varepsilon(\ell, k, 1/2, d_0)\right\}$, where $\varepsilon(\ell, k, 1/2, d_0)$ is given by Lemma 10 applied with ℓ, k, $\gamma = 1/2$, and d_0. Moreover, let $T_0 = T_0(k, s, \varepsilon)$ be given by Theorem 9 applied with k, $t_0 = s$, and ε. We now define the promised constants as

$$\varrho = \frac{q}{T_0} \qquad \text{and} \qquad \xi = \frac{d_0^{\binom{\ell}{2}}}{2T_0^\ell},$$

and let n_0 be sufficiently large.

Let $H = (V, E)$ be a (ϱ, d)-dense k-uniform hypergraph. The weak hypergraph regularity lemma yields a partition $V_0 \dot\cup V_1 \dot\cup \ldots \dot\cup V_t$, $s \le t \le T_0$ (s and T_0 defined above) which satisfies properties (ii) and (iii) of Theorem 9 (with ε defined above). We consider the following auxiliary, so-called *reduced hypergraph*, $R = ([t], E_R)$, where $e \in \binom{[t]}{k}$ is an edge in E_R if, and only if, $(V_i)_{i\in e}$ is an ε-regular k-tuple. Hence,

$$|E_R| \ge (1-\varepsilon)\binom{t}{k} > (1 - 1/\tbinom{s}{k})\binom{t}{k} \ge \mathrm{ex}(t, K_s^{(k)}),$$

where $\mathrm{ex}(t, K_s^{(k)})$ is the Turán number for $K_s^{(k)}$, i.e., the largest number of k-tuples among all $K_s^{(k)}$-free k-uniform hypergraphs on t vertices (the inequality we used above is well-known). Consequently, R contains a copy of $K_s^{(k)}$, and we denote this copy by $R_s \subseteq R$. Now, we 2-color the edges of R_s according to the density of the corresponding k-tuple. More precisely, we color the edge $e = \{i_1, \ldots, i_k\}$ "sparse" if $d(V_{i_1}, \ldots, V_{i_k}) \le d_0$, and we color it "dense" otherwise. We now argue that R_s does not contain a "sparse" copy of $K_q^{(k)}$.

Indeed, suppose R_s does contain a "sparse" clique $K_q^{(k)}$. Let $i_1, \ldots, i_q$ be the vertices of this clique, and set $U = \dot\bigcup_{j=1}^q V_{i_j}$. Since $i_1, \ldots, i_q$ spanned a "sparse" clique in R_s, the number of edges $e_H(U)$ can be bounded from above by

$$\begin{aligned} e_H(U) &\le d_0\binom{q}{k}\left(\frac{n}{t}\right)^k + q\binom{n/t}{2}\binom{qn/t}{k-2} \\ &< \left(d_0 + \frac{1}{q}\right) q^k \left(\frac{n}{t}\right)^k \le \frac{d(qn/t)^k}{2k!} < d\binom{|U|}{k}, \end{aligned} \tag{4}$$

where we used the choice of d_0 and q and the fact that n is sufficiently large. Clearly, (4) violates the (ϱ, d)-denseness of H, and so R_s contains no "sparse" clique $K_q^{(s)}$.

By the choice of $s = r_k(q, \ell)$, R_s must contain a "dense" clique $K_\ell^{(s)}$. Let $i_1, \ldots, i_\ell$ be the vertex set of that clique. From the preparation above, $H[V_{i_1}, \ldots, V_{i_s}]$ satisfies the hypothesis of the counting lemma for linear hypergraphs (Lemma 10), and therefore, $H \supseteq H[V_{i_1}, \ldots, V_{i_s}]$ contains at least

$$\frac{d_0^{e(S)}}{2}\left(\frac{n}{t}\right)^\ell \ge \frac{d_0^{\binom{\ell}{2}}}{2T_0^\ell} n^\ell = \xi n^\ell$$

copies of any $S \in \mathcal{L}_\ell^{(k)}$, making H $(\xi, \mathcal{L}_\ell^{(k)})$-universal. □

5. Non-universal hypergraphs

In this section, we deduce Corollary 8 from Theorem 7, according to the following outline. Since the given hypergraph H is not universal (for linear hypergraphs), Theorem 7 implies that there must be a subset $U \subseteq V$, of linear size, containing only "few" edges. We apply this observation repeatedly, obtaining a partition $V_1 \dot\cup \dots \dot\cup V_t$ of nearly the entire vertex set, where $H[V_i]$ is "sparse" for every $i \in [t]$. This, however, implies that the number of edges of H intersecting at least two classes from the partition must be slightly larger than expected. Finally, this "extra" density will "survive" when we distribute the remaining vertices of H into $V_1, \dots, V_t$.

Proof of Corollary 8. Let integers $\ell \geq k \geq 2$ and $d > 0$ be fixed. To define the promised constants t, β and ξ, we first consider a few auxiliary constants. Set $c = d/4$. Theorem 7 yields constants $\varrho' = \varrho'(\ell, k, c)$, $\xi' = \xi'(\ell, k, c)$, and $n_0' = n_0'(\ell, k, c)$. Set

$$\varsigma = \min\left\{(\varrho')^2, \tfrac{c^2}{16k^2}\right\}. \tag{5}$$

We now define the promised constants as

$$t = \left\lceil \frac{1-\sqrt{\varsigma}}{\varsigma} \right\rceil, \qquad \beta = \frac{d}{4t^{k-1}} \qquad \text{and} \qquad \xi = \xi' \varsigma^{\ell/2}$$

and let $n_0 \geq \max\{n_0'/\sqrt{\varsigma}, t/\varsigma, 2kt\}$ be sufficiently large.

Note that it suffices to prove Corollary 8 for hypergraphs H for which n is a multiple of t. Indeed, otherwise we could first remove constantly many ($x = n \pmod t$) vertices from H. For the resulting hypergraph H', we would obtain $\tau_t(H') \geq d + \beta - o(1)$, and so distributing the removed x vertices appropriately into the corresponding cut of H' implies $\tau_t(H) \geq d + \beta - o(1)$, where $o(1)$ tends to 0 as $n \to \infty$.

So, let $H = (V, E)$ be a k-uniform hypergraph on $n = mt \geq n_0$ vertices (for some $m \in \mathbb{N}$) with at least $d\binom{n}{k}$ edges which is not $(\xi, \mathcal{L}_\ell^{(k)})$-universal. Because of the choice of ξ, we infer from Theorem 7 that no subset $W \subseteq V$ with $|W| \geq \sqrt{\varsigma} n$ is $(\sqrt{\varsigma}, c)$-dense. In other words, every such W contains a subset $W' \subseteq W$, $|W'| \geq \sqrt{\varsigma}|W| \geq \varsigma n$ such that $e_H(W') \leq c\binom{|W'|}{k}$. In fact, a simple averaging argument shows that there must be such a set W' with $|W'| = \lfloor \varsigma n \rfloor$. Repeatedly selecting disjoint such W' yields a vertex partition $V = V_0 \dot\cup V_1 \dot\cup \dots \dot\cup V_t$ such that for all $i \in [t]$,

$$|V_i| = \lfloor \varsigma n \rfloor \quad \text{and} \quad e_H(V_i) \leq c\binom{\varsigma n}{k}, \quad \text{and} \quad |V_0| \leq (\sqrt{\varsigma} + \varsigma)n\,.$$

Indeed such a partition exists, since $(t-1)\lfloor \varsigma n \rfloor < (1 - \sqrt{\varsigma})n$ (owing to the choice of t) and $t\lfloor \varsigma n \rfloor \geq t\varsigma n - t \geq (1-\sqrt{\varsigma})n - \varsigma n$ (owing to the choices of t and n_0).

We now redistribute the vertices of V_0 among the classes $V_1, \dots, V_t$ and obtain a partition $U_1 \dot\cup \dots \dot\cup U_t = V$ such that, for each $i \in [t]$, $|U_i| = m = n/t$ and

$$e_H(U_i) \leq c\binom{\varsigma n}{k} + \frac{|V_0|}{t}\binom{m}{k-1} \leq c\binom{m}{k} + (\sqrt{\varsigma} + \varsigma)m\binom{m}{k-1}.$$

Because of (5), we have $(\sqrt{\varsigma} + \varsigma)k \leq c/2$, and so

$$e_H(U_i) \leq \left(c + (\sqrt{\varsigma} + \varsigma)k\frac{m}{m-k+1}\right)\binom{m}{k} \leq 2c\binom{m}{k},$$

where we also used that $m = n/t \geq 2k$. Consequently, the number of edges which are not completely contained in any one of the sets U_i is at least $d\binom{n}{k} - 2ct\binom{m}{k}$, and so

$$\tau_t(H) \geq \frac{\left|E(H) \setminus \bigcup_{i=1}^{t}\binom{U_i}{k}\right|}{\binom{n}{k} - t\binom{m}{k}} \geq \frac{d\binom{n}{k} - 2ct\binom{m}{k}}{\binom{n}{k} - t\binom{m}{k}} \geq d + \beta\,, \tag{6}$$

where we used the choice of $c = d/4$ and $\beta = d/(4t^{k-1})$ and the fact that n is sufficiently large for the last inequality. □

6. Concluding remarks

Subgraphs of locally dense graphs. The following question seems interesting already for graphs. Recall from Theorem 1 that a (ϱ, d)-quasirandom n-vertex graph H contains $(1 \pm o(1))d^{e_F}n^{v_F}$ labeled copies of any fixed graph F. It is conceivable that replacing (ϱ, d)-quasirandomness by (ϱ, d)-denseness would not decrease this number. We believe the following question has an affirmative answer.

Question 1. Is it true that for any $\gamma, d > 0$ and any graph F, there exist $\varrho > 0$ and n_0 so that any (ϱ, d)-dense graph H on $n \geq n_0$ vertices contains at least $(1-\gamma)d^{e_F}n^{v_F}$ labeled copies of F?

One may check that the answer to Question 1 is positive when F is a clique or more generally, a complete ℓ-partite graph for some fixed ℓ. Sidorenko [15, 16] made a related conjecture stating that any graph G with at least $d\binom{n}{2}$ edges contains at least $(1-o(1))d^{e_F}n^{v_F}$ labeled copies of any given bipartite graph F. Sidorenko's conjecture is known to be true for even cycles, complete bipartite graphs and was recently proved for a certain family of graphs including Boolean cubes [7]. Since our assumption in Question 1 is stronger than that made in Sidorenko's conjecture, the positive answer to Sidorenko's conjecture would also validate Question 1 for all bipartite graphs. To our knowledge, the smallest non-bipartite graph for which Question 1 is open is the 5-cycle.

Regularity and partial Steiner systems. In this note, we established that a fairly weak concept of regularity provides a counting lemma for linear hypergraphs. In order to extend this result to partial Steiner (r, k)-systems (k-uniform hypergraphs in which every r-set is covered at most once), a stronger concept of regularity will be needed. For example, when $r = 3 \leq k$, one will need a concept of regularity for k-uniform hypergraphs H which relates the edges of H to certain subgraphs of $K^{(2)}_{|V(H)|}$ (rather than to subsets of $V(H)$). Such concepts of regularity for $k = 3$ were considered in [4, 6]. For arbitrary $r \leq k$, one will need that H is regular w.r.t. certain subhypergraphs $G^{(r)}$ of $K^{(r)}_{|V(H)|}$, where $G^{(r)}$ has to be regular w.r.t. certain subhypergraphs $G^{(r-1)}$ of $K^{(r-1)}_{|V(H)|}$, and so on. This stronger concept of regularity is related to the hypergraph regularity lemmas from [5, 14, 19].

Remark on Theorem 4. Note that the parameter ϱ in the concept of (ϱ, d)-quasirandomness plays two roles. On the one hand, it "governs the locality", i.e., the size of the subsets to which the condition of uniform edge distribution applies. On the other hand, it "governs the precision" of that condition. The following result shows that, in fact, one can (formally) relax the condition on the locality, if the precision remains high enough.

Theorem 11. *For all integers $\ell \geq k \geq 2$, $\gamma, d > 0$, $1/k > \varepsilon > 0$ and every $F \in \mathcal{L}^{(k)}$, there exist $\delta > 0$ and n_0 so that any k-uniform hypergraph $H = (V, E)$ on $n \geq n_0$ vertices with the property that $e_H(U) = (d \pm \delta)\binom{|U|}{k}$ for every $U \subseteq V$ with $|U| \geq \varepsilon|V|$ contains $(1 \pm \gamma)d^{e_F} n^{v_F}$ labeled copies of F.*

Theorem 11 can be proved in a similar way to Theorem 4, and so we omit the details. The main idea, however, is to show first that a hypergraph satisfying the assumptions of Theorem 11 is, in fact, (ϱ, d)-quasirandom for some $\varrho = \varrho(\delta)$ with $\varrho(\delta) \to 0$ as $\delta \to 0$.

Non-universality and large cuts. For graphs, Corollary 8 has the consequence that if one selects, uniformly at random, a set $I \in \binom{[t]}{t/2}$ (say, w.l.o.g., that t is even), then the set $U = \bigcup_{i \in I} V_i$ induces a cut larger than $(d + \beta)(n/2)^2 = (d + \beta - o(1))(1/2)\binom{n}{2}$, for some small $\beta > 0$ independent of n (see [8, 10] for related results). For $k \geq 3$, Corollary 8 does not seem to yield immediately a similar result, and the following question remains open.

Question 2. Is it true that for all integers $\ell \geq k \geq 3$ and $d, \xi > 0$, there exist $\beta > 0$ and n_0 so that if $H = (V, E)$ is a k-uniform hypergraph on $n \geq n_0$ vertices and $d\binom{n}{k}$ edges which is not $(\xi, \mathcal{L}_\ell^{(k)})$-universal, then there exists a set $U \subseteq V$ of size $\lfloor n/2 \rfloor$ such that

$$\left|\left\{e \in E\colon\ 1 \leq |e \cap U| \leq k-1\right\}\right| \geq (d + \beta)\left(1 - \frac{1}{2^{k-1}}\right)\binom{n}{k}\,?$$

References

1. F. R. K. Chung, *Regularity lemmas for hypergraphs and quasi-randomness*, Random Structures Algorithms **2** (1991), no. 2, 241–252. 2
2. F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1989), no. 4, 345–362. 1
3. P. Frankl and V. Rödl, *The uniformity lemma for hypergraphs*, Graphs Combin. **8** (1992), no. 4, 309–312. 2
4. ______, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164. 6
5. W. T. Gowers, *Hypergraph regularity and Szemerédi's theorem*, Ann. of Math. (2), to appear. 2, 6
6. ______, *Quasirandomness, counting and regularity for 3-uniform hypergraphs*, Combin. Probab. Comput. **15** (2006), no. 1-2, 143–184. 6
7. H. Hatami, *Graph norms and Sidorenko's conjecture*, manuscript. 6
8. Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*, Recent advances in algorithms and combinatorics, CMS Books Math./Ouvrages Math. SMC, vol. 11, Springer, New York, 2003, pp. 289–351. 6
9. B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 113–179. 2
10. V. Nikiforov, *Edge distribution of graphs with few copies of a given graph*, Combin. Probab. Comput. **15** (2006), no. 6, 895–902. 1, 6
11. V. Rödl, *On universality of graphs with uniformly distributed edges*, Discrete Math. **59** (1986), no. 1-2, 125–134. 1
12. V. Rödl and M. Schacht, *Regular partitions of hypergraphs: Counting Lemmas*, Combin. Probab. Comput. **16** (2007), no. 6, 887–901. 2

13. ———, *Regular partitions of hypergraphs: Regularity Lemmas*, Combin. Probab. Comput. **16** (2007), no. 6, 833–885. 2
14. V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures Algorithms **25** (2004), no. 1, 1–42. 6
15. A. F. Sidorenko, *Inequalities for functionals generated by bipartite graphs*, Diskret. Mat. **3** (1991), no. 3, 50–65. 6
16. ———, *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), no. 2, 201–204. 6
17. A. Steger, *Die Kleitman–Rothschild Methode*, Ph.D. thesis, Forschungsinstitut für Diskrete Mathematik, Rheinische Friedrichs–Wilhelms–Universität Bonn, March 1990. 2
18. E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401. 2, 2
19. T. Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A **113** (2006), no. 7, 1257–1280. 2, 6
20. A. Thomason, *Pseudorandom graphs*, Random graphs '85 (Poznań, 1985), North-Holland Math. Stud., vol. 144, North-Holland, Amsterdam, 1987, pp. 307–331. 1
21. ———, *Random graphs, strongly regular graphs and pseudorandom graphs*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–195. 1

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, Brazil
E-mail address: `yoshi@ime.usp.br`

Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620, USA.
E-mail address: `bnagle@math.usf.edu`

Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
E-mail address: `rodl@mathcs.emory.edu`

Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany
E-mail address: `schacht@informatik.hu-berlin.de`