Uniform edge distribution in hypergraphs is hereditary

Dhruv Mubayi [*] Vojtěch Rödl [†]

Submitted: Jul 11, 2003; Accepted: Aug 16, 2004 ; Published: Aug 31, 2004

Abstract

Let $\alpha \in (0,1), l \geq 2$ and let $\mathcal{H}_n$ be an l-graph on n vertices. $\mathcal{H}_n$ is (α, ξ)-uniform if every ξn vertices of $\mathcal{H}_n$ span $(\alpha \pm \xi)\binom{\xi n}{l}$ edges. Our main result is the following. For all $\widetilde{\delta}$, there exist δ, r, n_0 such that, if $n > n_0$ and $\mathcal{H}_n^{(l)}$ is (α, δ)-uniform, then all but $\exp\{-r^{1/l}/20\}\binom{n}{r}$ r-sets of vertices induce a subhypergraph that is $(\alpha, \widetilde{\delta})$-uniform.

We also present the following application. Let $\mathcal{F}$ be a fixed l-graph, and $c > 0$. Then there is an n_0 and r' such that: If $\mathcal{H}$ is an n vertex l-graph ($n > n_0$) such that the deletion of any cn^l edges of $\mathcal{H}$ leaves an l-graph that admits no homomorphism into $\mathcal{F}$, then there exists $\mathcal{H}' \subset \mathcal{H}$ on r' vertices, that also admits no homomorphism into $\mathcal{F}$. This extends a recent result of Alon and Shapira, who proved it when $\mathcal{F}$ is a complete graph.

1 Introduction

Throughout this paper we write $x = y \pm z$ for the statement $x \in [y - z, y + z]$, and we omit ceiling and floor symbols. We write $\mathcal{H}_n^{(l)}$ to denote an n vertex l-graph.

Definition 1.1. *$\mathcal{H}_n^{(l)}$ is (α, ξ)-uniform if every ξn vertices of $\mathcal{H}_n^{(l)}$ span $(\alpha \pm \xi)\binom{\xi n}{l}$ edges. $\mathcal{H}_n^{(l)}$ is (α, ξ)-dense if every ξn vertices of $\mathcal{H}_n^{(l)}$ span at least $\alpha\binom{\xi n}{l}$ edges.*

Our main result, Theorem 1.2 below, says that in an l-graph whose edges are uniformly distributed, this property is inherited to almost all sets of constant size.

[*]Department of Mathematics, Statistics, and Computer Science, University of Illinois, 851 S. Morgan Street, Chicago, IL 60607-7045; research supported in part by the National Science Foundation under grants DMS-9970325 and DMS-0400812

[†]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA; research supported in part by the National Science Foundation under grants DMS-0071261 and DMS-0300529

2000 Mathematics Subject Classification: 05C15, 05C35, 05C65, 05D05

Keywords: *Ramsey Turán problems, extremal hypergraph theory*

Theorem 1.2. *(Sections 6–8) Let* $\alpha \in (0,1), l \geq 2$ *be fixed. For all* $\widetilde{\delta}$*, there exist* δ, r, n_0 *such that, if* $n > n_0$ *and* $\mathcal{H}_n^{(l)}$ *is* (α, δ)*-uniform, then all but* $\exp\{-r^{1/l}/20\}\binom{n}{r}$ r*-sets of vertices induce a subhypergraph that is* $(\alpha, \widetilde{\delta})$*-uniform.*

Remark 1.3. *By bounding some inequalities more carefully in the proof of Theorem 1.2, we can improve the factor* $\exp\{-r^{1/l}/20\}$ *to* $\exp\{-cr\}$*, where* c *is an absolute constant.*

Remark 1.4. *Note that the minimum* r *with the property of Theorem 1.2 must satisfy* $r \to \infty$ *as* $\widetilde{\delta} \to 0$*. Otherwise, we could choose* $\widetilde{\delta}$ *such that* $\widetilde{\delta} r < l$ *and* $\alpha > \widetilde{\delta}$*. With this choice, no* r*-set* R *can be* $(\alpha, \widetilde{\delta})$*-uniform, since every subset of* R *of size* $\widetilde{\delta} r < l$ *has density* $0 < \alpha - \widetilde{\delta}$*.*

Remark 1.5. *We will in fact prove Theorem 1.2 with the following quantification:*

$$\forall \widetilde{\delta} \quad \exists r_0 \quad \forall r > r_0 \quad \exists \delta, n_0 \quad \forall n > n_0$$

Remark 1.6. *Note also that if* $\delta' < \delta$*, and Theorem 1.2 holds for* δ*, then it also holds for* δ'*, since an* (α, δ')*-uniform hypergraph is also* (α, δ)*-uniform. We may therefore take* δ *to be as small as we need in relation to* r*, in particular, we assume that* $\delta \ll 1/r^r$*.*

Remark 1.7. *After proving Theorem 1.2, we found out that Alon, de la Vega, Kannan, and Karpinski [4] independently and simultaneously proved a similar theorem (Theorem 9 in the extended version of [4]) using completely different techniques. It may be possible to use their ideas to prove Theorem 1.2 with* $\exp\{-r^{1/l}/20\}$ *replaced by 1/40 .*

Remark 1.8. *We recently learned that Theorem 1.2 was also obtained for l=3 by Nagle and Czygrinow [6] using an alternative proof. Those authors infer Theorem 1.2 by first establishing an analogue in a stronger environment provided by an application of the Frankl-Rodl hypergraph regularity lemma (cf. [14]).*

Theorem 1.2 has already had applications to Ramsey-Turán problems for hypergraphs. In particular, it is used in [19] to construct infinitely many l-graphs whose Ramsey-Turán density is strictly positive and less than the Turán density. In this paper we present another application.

Erdős asked the following question, which generalizes a theorem of Bollobás, Erdős, Simonovits, and Szemerédi [5]: Given positive integers c and k, do there exist positive integers $f_k(c)$ and n_0 such that the following holds: If G is an n vertex graph with $n > n_0$ with the property that the chromatic number of G cannot be made less than k by the omission of at most cn^2 edges, then G contains a k-chromatic subgraph with at most $f_k(c)$ vertices.

It is well-known that chromatic number is a global property, namely, that a graph can have large chromatic number without having a small subgraph with large chromatic number. Another motivation for Erdős' question is: to what extent is chromatic number global in the sense described above? Duke and the second author [8] answered Erdős' question. Recently, a new approach was developed by Goldreich, Goldwasser, and Ron

[15] which gives better bounds on $f_k(c)$. Their motivation came from computer science, to find efficient algorithms to test whether a graph has a given property. These results were further improved by Alon and Krivelevich [1].

As a relatively straightforward application of Theorem 1.2, we generalize these results to hypergraphs. Moreover, instead of speaking just about chromatic number, we observe that the same proof gives a slightly stronger statement about homomorphisms. Recall that the l-graph $\mathcal{H}$ has a homomorphism into the l-graph $\mathcal{F}$ if there is a function $f : V(\mathcal{H}) \to V(\mathcal{F})$ such that if $\{u_1, \ldots, u_l\}$ is an edge of $\mathcal{H}$, then $\{f(u_1), \ldots, f(u_l)\}$ is an edge of $\mathcal{F}$.

Theorem 1.9. *(Section 3) Let $\mathcal{F}$ be a fixed l-graph, and $c > 0$. Then there is an n_0 and r' such that: If $\mathcal{H}$ is an n vertex l-graph ($n > n_0$) such that the deletion of any cn^l edges of $\mathcal{H}$ leaves an l-graph that admits no homomorphism into $\mathcal{F}$, then there exists $\mathcal{H}' \subset \mathcal{H}$ on r' vertices, that also admits no homomorphism into $\mathcal{F}$.*

2 The Regularity Lemma

In this section we describe our main tool needed to prove Theorem 1.2, the Szemerédi Regularity Lemma. In an l-graph $\mathcal{H}$, let $X_1, \ldots, X_l$ be pairwise disjoint sets of vertices. Write $e(X_1, \ldots, X_l)$ for the number of edges with exactly one point in each X_i. The *density* of the l-tuple $X_1, \ldots, X_l$ is

$$d(X_1, \ldots, X_l) = \frac{e(X_1, \ldots, X_l)}{|X_1| \ldots |X_l|}.$$

Given $\epsilon > 0$, the l-tuple $X_1, \ldots, X_l$ is ϵ-regular with density $\alpha = d(X_1, \ldots, X_l)$ if for every choice of $Y_i \subseteq X_i$ with

$$|Y_i| \geq \epsilon |X_i| \text{ for each } i \ ,$$

we have

$$d(Y_1, \ldots, Y_l) = \alpha \pm \epsilon.$$

We also extend the definition of (α, ϵ)-uniform to a partite situation.

Definition 2.1. *The l-partite $\mathcal{H} = \mathcal{H}(V_1, \ldots, V_l)$ is (α, ϵ)-uniform if for every choice of $V_i' \subset V_i$, with $|V_i'| = \epsilon |V_i|$, the density $d(V_1', \ldots, V_l') = \alpha \pm \epsilon$.*

Remark 2.2. *For an l-tuple $L = X_1, \ldots, X_l$ there is a subtle distinction between the following two statements*

1) L is ϵ-regular with density α

2) L is (α, ϵ)-uniform.

It is clear that 1) implies 2), but if 2) holds, then the best we can say regarding 1) is that L is 2ϵ-regular with density $\alpha' = \alpha \pm \epsilon$. Throughout this paper we use 1) in contexts where the density of L is not central and 2) where it is.

Consider a partition $\mathcal{P} = V_0 \cup V_1 \cup \cdots \cup V_t$ of $V(\mathcal{H})$. We say that $\mathcal{P}$ is an ϵ-regular partition if

1) $|V_0| < \epsilon |V|$,

2) $|V_1| = \cdots = |V_t|$,

3) the l-tuple $V_{i_1}, \ldots, V_{i_l}$ is ϵ-regular for all but ϵt^l choices of l of the V_i's.

With these notions we can state the celebrated Szemerédi Regularity Lemma. Below we state a version for l-graphs, $l \geq 2$. Its proof is essentially the same as for the case $l = 2$.

Lemma 2.3. *(Regularity Lemma, see [18]) Let $l \geq 2$ be fixed. For every $\epsilon > 0$ and every integer $t_0 \geq 1$ there exist integers T, n_0 such that every l-graph of order at least n_0 admits an ϵ-regular partition $V_0, \ldots, V_t$ with $t_0 \leq t \leq T$.*

3 Proof of Theorem 1.9

In this section, we sketch the proof of Theorem 1.9. Our mail tool is Theorem 3.1 below. Later in Section 6.2, we will see that Theorem 3.1 is a special case of Theorem 6.2, which is the main result proved in this paper. Moreover, we will also show that Theorem 6.2 is essentially equivalent to Theorem 1.2.

Theorem 3.1. *Fix $0 < \alpha < 1$ and $l \geq 2$. Let $G = G(A_1, \ldots, A_l)$ be an l-partite l-graph with $|A_i| = n$ for all i. For all $\widetilde{\delta}$, there exist δ, r, n_0 such that if $n > n_0$, the following holds: Suppose that G is (α, δ)-uniform. Then all but $\exp\{-r^{1/l}/10\}\binom{n}{r}^l$ choices of l-tuples $(\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_l)$ of r-sets $\widetilde{A}_i \in [A_i]^r, i \in [l]$ are $(\alpha, \widetilde{\delta})$-uniform.*

Proof of Theorem 1.9 (Sketch): Let $c, \mathcal{F}$ be given, and let $\mathcal{H}$ be an l-graph on $n > n_0$ vertices (n_0 comes from Lemma 2.3 and Theorem 3.1) such that after deleting any cn^l edges, the resulting l-graph admits no homomorphism into $\mathcal{F}$. Let $\epsilon \ll c$, and $t_0 \gg \max\{|V(\mathcal{F})|, 1/c\}$. Let $V = V_0 \cup \ldots \cup V_t$ be an ϵ-regular partition of $\mathcal{H}$, where $|V_0| \leq \epsilon |V| = \epsilon n$. We now apply Theorem 3.1 to each ϵ-regular l-tuple that has density at least $c/2$. In our setup, ϵ plays the role of the parameter δ in Theorem 3.1. By Theorem 3.1, if the l-tuple $\{V_{i_1}, \ldots, V_{i_l}\}$ is ϵ-regular with density $\alpha \geq c/2$, then for all but $\exp\{-r^{1/l}/10\}\binom{m}{r}$ choices of l-tuples $(\tilde{V}_{i_1}, \ldots, \tilde{V}_{i_l})$ of r-sets $\tilde{V}_{i_j} \subset [V_{i_j}]^r$, the l-tuple $(\tilde{V}_{i_1}, \ldots, \tilde{V}_{i_l})$ is $(\alpha, \widetilde{\delta})$-uniform. Consequently, if we choose an r-set $\tilde{V}_i$ from each V_i randomly, then the expected number of l-tuples of chosen r sets violating the conclusion of Theorem 3.1 is at most $\exp\{-r^{1/l}/10\}\binom{t}{l}$. Hence there exists a choice of r-sets $V_i' \subset V_i$ for $i \in [t]$, such that all but $2\exp\{-r^{1/l}/10\}\binom{t}{l}$ of the l-tuples of these r-sets satisfy the conclusion of Theorem 3.1. Delete from $\mathcal{H}$ all edges that

1a) have nonempty intersection with V_0,

1b) lie within at most $l-1$ distinct V_is,

1c) contain one vertex from each of $V_{i_i}, \ldots, V_{i_l}$, where $V_{i_1}, \ldots, V_{i_l}$ is an ϵ-irregular l-tuple, or is ϵ-regular with density less than $c/2$,

2) lie within an l-tuple $V_{i_1}, \ldots, V_{i_l}$ for which $V'_{i_1}, \ldots, V'_{i_l}$ violates the conclusion of Theorem 3.1.

It is easy to see that the number of edges deleted from $\mathcal{H}$ in 1a)–1c) is at most $(c/2)n^l$. The number of edges deleted in 2) is at most $2\exp\{-r^{1/l}/10\}\binom{t}{l}(n/t)^l < (c/2)n^l$, since $c \gg 1/r$. Therefore we have altogether deleted at most cn^l edges, and the resulting l-graph admits no homomorphism into $\mathcal{F}$.

Define the reduced l-graph $\mathcal{H}^*$ of $\mathcal{H}$ with vertex set $\{V_1, \ldots, V_t\}$, where $\{V_{i_1}, \ldots, V_{i_l}\}$ forms an edge if the corresponding l-tuple in $\mathcal{H}$ is ϵ-regular with density at least $c/2$. Then it is easy to see that $\mathcal{H}^*$ also admits no homomorphism into $\mathcal{F}$. If $\{V_{i_1}, \ldots, V_{i_l}\}$ is an edge of $\mathcal{H}^*$, then $V'_{i_1}, \ldots, V'_{i_l}$ satisfies the following property: for every choice of $C_{i_1} \subset V'_{i_1}, \ldots, C_{i_l} \subset V'_{i_l}$, where $|C_{i_j}| \geq |V'_i|/|V(\mathcal{F})| \geq \widetilde{\delta} r$ for all j, there is an edge with one element in each of $C_{i_1}, \ldots, C_{i_l}$. From this it is easy to show that the subhypergraph $\mathcal{H}'$ of $\mathcal{H}$ with vertex set $V'_1 \cup \ldots \cup V'_t$ admits no homomorphism into $\mathcal{F}$ (see Duke-Rödl [8] for the details in the graph case). Moreover, the number of vertices in $\mathcal{H}'$ is $tr = r'$ which is independent of n. □

4 Hypergeometric Lemmas

In this section we state several generalizations of the hypergeometric inequalities that are proved in the Appendix. We begin with the following two statements which follow easily from large deviation results of the hypergeometric distribution (see [17]).

Lemma 4.1. *Let $|X| = x$, and $|Y| = y$ with $Y \subseteq X$. Let $R \in [X]^r$ be a random r-set chosen independently from the family of all r-sets of X. Let $r > r_0$ and $3/2 > \sigma \geq 1/\log r$. Then with probability greater than $1 - e^{-\sqrt{r}}$, the following holds:*
a) $|R \cap Y| = \frac{ry}{x}(1 \pm \sigma)$ if $y \geq x/r^{1/4}$,
b) $|R \cap Y| < r^{19/20}$ if $y \leq x/r^{1/10}$.

Proof. See Appendix. □

An easy consequence of Lemma 4.1 is the following:

Proposition 4.2. *For all T, there exist r_0, n_0, such that if $r > r_0, n > n_0$, the following holds: Let $B = B_1 \cup \ldots \cup B_T$, where the B_i's are pairwise disjoint and each of size n/T. Then all but $Te^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $R \in [B]^r$ satisfy $|R \cap B_i| = (r/T)(1 \pm 1/\log r)$ for each $i \in [T]$.*

Proof. It suffices to show that for a randomly chosen r-set R, the probability that $|R \cap B_i| \neq (r/T)(1 \pm 1/\log r)$ is less than $e^{-\sqrt{r}}$. This follows from Lemma 4.1 part a) since r_0 is sufficiently large that $T < r^{1/4}$ and hence $n/T > n/r^{1/4}$. □

We also need the following fractional extension of Proposition 4.2.

Lemma 4.3. *For all c, ξ, there exist r_0, n_0 such that if $r > r_0, n > n_0$, the following holds: If $\sum_{i \in [n]} x_i = cn$, where $0 \leq x_i \leq 1$ for each i, then all but $e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $R \in [n]^r$ satisfy $\sum_{i \in R} x_i = (c \pm \xi)r$.*

Proof. See Appendix. □

The following is an extension of the hypergeometric inequalities to edge sets of weighted graphs. Given a set S in a weighted graph, the weight of S is the sum of the weights of all the edges induced by S. We write $deg_G(x)$ for the weighted degree of vertex x.

Lemma 4.4. *Fix $k \geq 2$. For every β, there exists r_0, n_0 such that if $r > r_0, n > n_0$, the following holds. Let G be an n vertex weighted graph with total weight at most βn^k, maximum weighted degree at most n^{k-1}, and maximum edge weight at most n^{k-2}. Then all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets of vertices induce a subgraph with weight at most $2\beta n^{k-2}r^2$.*

Proof. See Appendix. □

Proposition 4.5. *Let $G = (A, B)$ be a bipartite graph with n^2/t edges and $|A| = |B| = n$. Let $\mathcal{B} \subset E(G)$ be a set of size at most $\epsilon' n^2/t$. Let $\mathcal{B}' \subset [E(G)]^2$ be a set size at most $\epsilon' n^4/t^2$. Then the following holds: For all ϵ', t, there exist r_0, n_0, such that if $r > r_0, n > n_0$, then for all but $6e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$,*

$$|(\widetilde{A} \times \widetilde{B}) \cap E(G)| = \frac{r^2}{t}(1 \pm 1/10)$$

and in the graph spanned by $\widetilde{A} \times \widetilde{B}$,
1) at most $2\epsilon' r^2/t$ edges lie in $\mathcal{B}$, and
2) at most $2\epsilon' r^4/t^2$ pairs of edges lie in $\mathcal{B}'$.

Proof. See Appendix. □

5 Notation for Theorem 1.2

In this section we generalize notation from Section 3 to weighted hypergraphs. Let $l \geq 2$, and let $A_1, \ldots, A_l$ be pairwise disjoint sets. Let $w : A_1 \times \ldots \times A_l \to [0, 1]$. Think of w as a *weight* function on the edges of the complete l-partite l-graph $\mathcal{H}$ with parts $A_1, \ldots, A_l$. In the case $\mathcal{H}$ is not complete, the weight function w assigns zero to all edges outside $\mathcal{H}$.

The weight of the l-tuple $L = A_1, \ldots, A_l$ under w is

$$w(L) = \sum\{w(a_1, \ldots, a_l); (a_1, \ldots, a_l) \in A_1 \times \ldots \times A_l\}.$$

The *density* of L is

$$d(L) = \frac{w(L)}{|A'_1| \cdots |A'_l|}.$$

For $A'_i \subset A_i$, $i \in [l]$, define the density

$$d(A'_1, \ldots, A'_l) = \frac{\sum\{w(a_1, \ldots, a_l); (a_1, \ldots, a_l) \in A'_1 \times \cdots \times A'_l\}}{|A'_1| \cdots |A'_l|}.$$

We define w to be ϵ-regular of density $\alpha = d(L)$ if for all choices of $A'_i \subset A_i$, with $|A'_i| \geq \epsilon n$ for each $i \in [l]$,

$$d(A'_1, \ldots, A'_l) = \alpha \pm \epsilon.$$

As in Section 3, w is (α, ϵ)-uniform if for all choices of $A'_i \subset A_i$, with $|A'_i| \geq \epsilon n$ for each $i \in [l]$, we have

$$d(A'_1, \ldots, A'_l) = \alpha \pm \epsilon.$$

Given functions $\gamma_i : A_i \to [0, 1]$, $i \in [l]$, and $|\gamma_i| = \sum_{a \in A_i} \gamma_i(a)$, the fractional density

$$d(\gamma_1, \ldots, \gamma_l) = \frac{\sum\{w(a_1 \ldots a_l) \prod_i \gamma_i(a_i); (a_1, \ldots, a_l) \in A_1 \times \cdots \times A_l\}}{|\gamma_1| \ldots |\gamma_l|}.$$

Remark 5.1. *All these definitions extend naturally to l-uniform hypergraphs that are not necessarily l-partite.*

6 Proof of Theorem 1.2

6.1 The Structure

In order to prove Theorem 1.2, we will prove the following stronger weighted version of it.

Theorem 6.1. *Fix $0 < \alpha < 1$ and $l \geq 2$. Let $\mathcal{H}$ be a weighted l-graph with vertex set X of size n and weight function w. For all $\widetilde{\delta}$, there exist δ, r, n_0 such that, if $n > n_0$ the following holds: Suppose that w is (α, δ)-uniform. Then for all but* $\exp\{-r^{1/l}/20\}\binom{n}{r}$ *r-sets of vertices $\widetilde{X} \in [X]^r$, the restriction of w to $\widetilde{X}$ is $(\alpha, \widetilde{\delta})$-uniform.*

For $l \geq 2$, let $WT(l)$ be the statement of Theorem 6.1 for l-graphs, and $UWT(l)$ be the statement of Theorem 1.2 for l-graphs. By restricting to 0-1 weight functions, we see that

$$WT(l) \Longrightarrow UWT(l).$$

We will prove $WT(l)$ by induction on l, with the base case $l = 2$ in Section 6.4. Therefore our next step is to prove

$$WT(l) \Longrightarrow WT(l+1) \tag{1}$$

for each $l \geq 2$.
We will only prove

$$WT(2) \Longrightarrow UWT(3). \tag{2}$$

Extending (2) to $WT(2) \Longrightarrow WT(3)$ and further extending this to (1) is simply a matter of adding notation. In the interest of clarity we decided not to write the proof for this more general statement. We feel that our proof of (2) will convince the reader of (1), and Section 6.5.6 describes the minor notational changes needed to rigorously prove (1).

6.2 An equivalent l-partite formulation

In proving Theorem 6.1, it is more convenient to work with an l-partite l-graph. Indeed, the theorem that we prove is:

Theorem 6.2. *Fix $0 < \alpha < 1$ and $l \geq 2$. Let $G = G(A_1, \ldots, A_l)$ be an l-partite l-graph with $|A_i| = n$ for all i and weight function w. For all $\widetilde{\delta}$, there exist δ, r, n_0 such that if $n > n_0$, the following holds: Suppose that w is (α, δ)-uniform. Then for all but $\exp\{-r^{1/l}/10\}\binom{n}{r}^l$ choices of l-tuples $(\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_l)$ of r-sets $\widetilde{A}_i \in [A_i]^r, i \in [l]$, the restriction of w to $\cup_i \widetilde{A}_i$ is $(\alpha, \widetilde{\delta})$-uniform.*

Remark 6.3. *Note that Remarks 1.4, 1.5 and 1.6 after Theorem 1.2 also apply to Theorems 6.1 and 6.2.*

We need some additional notation. An *equipartition* P of a set S is a partition of S into sets of equal size. The equipartition P *contains* an edge e of S if e has nonempty intersection with each of the parts of P.

Proof that Theorem 6.2 $\Longrightarrow$ Theorem 6.1: As mentioned earlier, we only consider the unweighted case and $l = 3$. Alternatively, the weights are all 0 or 1. We describe the structure of the proof, postponing the details to the Appendix.

To prove that Theorem 6.2 $\Longrightarrow$ Theorem 6.1, we will show that given a hypergraph satisfying the hypothesis of Theorem 6.1, every equipartition of its vertex set yields a 3-partite hypergraph satisfying the hypothesis of Theorem 6.2. More precisely,

$$\text{hypothesis of Theorem 6.1 with } \delta, n_0 \Longrightarrow \text{hypothesis of Theorem 6.2 with } 14\delta, n_0/3 \quad (3)$$

and similarly

$$\text{conclusion of Theorem 6.2 with } \widetilde{\delta}/2, r/3 \Longrightarrow \text{conclusion of Theorem 6.1 with } \widetilde{\delta}, r. \quad (4)$$

In fact we prove (4) in contrapositive form, namely

$$\text{conclusion of Theorem 6.1 with } \widetilde{\delta}, r \text{ fails} \Longrightarrow \text{conclusion of Theorem 6.2 with } \widetilde{\delta}/2, r/3 \text{ fails.} \quad (5)$$

Proof of (3): See Appendix. □

Proof of (5): See Appendix. □

6.3 Degree Conditions for Regularity: the weighted case

In this section we extend degree conditions for ϵ-regularity of graphs to weighted graphs (c.f. [2] Lemmas 3.1 and 3.2).

First we describe the results for the unweighted case. Let $G = (X, Y)$ be a bipartite graph. Consider the following two statements where $0 \leq \epsilon, \epsilon', \alpha \leq 1$:

(i) $UNI(\epsilon)$: G is (α, ϵ)-uniform

(ii) $DPC(\epsilon')$:
a) All but $\epsilon'|X|$ vertices of X have degree $|Y|(\alpha \pm \epsilon')$
b) All but $\epsilon'|X|^2$ pairs of vertices in X have a common neighborhood of size $|Y|(\alpha^2 \pm \epsilon')$

The next Proposition is almost identical to Lemma 3.1 and Lemma 3.2 in [2]. We refer the reader to [2] or [7] for a proof.

Proposition 6.4. *$UNI(\epsilon)$ and $DPC(\epsilon')$ are equivalent in the following sense*
1) $\forall \epsilon, \exists \epsilon', n_0, \forall \alpha$, if $\min\{|X|, |Y|\} > n_0$, then $DPC(\epsilon') \Rightarrow UNI(\epsilon)$.
2) $\forall \epsilon', \exists \epsilon, n_0, \forall \alpha$, if $\min\{|X|, |Y|\} > n_0$, then $UNI(\epsilon) \Rightarrow DPC(\epsilon')$.

To extend this result to the weighted case, the following Proposition is useful.

Proposition 6.5. *Let $\mathcal{H}(A_1, \ldots, A_l)$ be an l-partite l-graph, w be a weight function on $\mathcal{H}$ of density α, and $\gamma_i : A_i \to \{0, 1\}$. Suppose that $\delta > 0$ and $|\gamma_i|$ is an integer for each i. Then the following are equivalent:*
1) w is (α, δ)-uniform
2) for all choices of $\gamma_i : A_i \to [0, 1]$ with $|\gamma_i| > \delta n$,

$$d(\gamma_1, \ldots, \gamma_l) = \alpha \pm \delta.$$

Proof. See Appendix. □

Now we extend Proposition 6.4 to the weighted case. We need the following set up: Let $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$ and $w : A \times B \to [0, 1]$. Let $w_{i,j} = w(a_i, b_j)$, and let $\vec{w_i} = (w_{i,1}, \ldots, w_{i,n})$. Set $\alpha = (1/n^2) \sum_{i,j} w_{i,j}$. Consider the following two statements:

(i) $UNI(\delta)$: w is (α, δ)-uniform

(ii) $DPC(\delta')$: All but $\delta' n^2$ pairs i, i' have dot product $\vec{w_i} \cdot \vec{w_{i'}} = (\alpha \pm \delta')^2 n$.

Proposition 6.6. *$UNI(\delta)$ and $DPC(\delta')$ are equivalent in the following sense:*
1) $\forall \delta, \exists \delta', n_0, \forall \alpha, \forall n > n_0, DPC(\delta') \Rightarrow UNI(\delta)$
2) $\forall \delta', \exists \delta, n_0 \forall \alpha, \forall n > n_0, UNI(\delta) \Rightarrow DPC(\delta')$

Proof. See Appendix. □

6.4 Proof of Theorem 6.2 for graphs

Theorem 6.7. *Fix $0 < \alpha < 1$. Let $G = G(A, B)$ be a bipartite graph with $|A| = |B| = n$ and weight function w. For all $\widetilde{\delta}$, there exist δ, r, n_0 such that if $n > n_0$, the following holds: Suppose that w is (α, δ)-uniform. Then for all but $exp\{-\sqrt{r}/10\}\binom{n}{r}^2$ choices of r-sets $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$, the restriction of w to $\widetilde{A} \cup \widetilde{B}$ is $(\alpha, \widetilde{\delta})$-uniform.*

Proof. Our constants will satisfy the hierarchy $\widetilde{\delta} \gg \delta'' \geq \delta' \gg \delta$. The choice of these constants, however will ensure that $\widetilde{\delta} \to 0$ as $\delta \to 0$, thus allowing any $\widetilde{\delta} > 0$ as input for the Theorem.

Since w is (α, δ)-uniform, G satisfies $DPC(\delta')$ for some δ' by Proposition 6.6 Part 2. Define an auxiliary graph G_A with vertex set A and edges consisting of pairs i, i' forming the exceptional set in $DPC(\delta')$. The definition of $DPC(\delta')$ implies that G_A has at most $\delta' n^2$ edges. Now apply Lemma 4.4 with $\beta = \delta'$ and $k = 2$. We obtain that all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $\widetilde{A} \in [A]^r$ induce at most $2\delta' r^2$ edges from G_A. Fix once such $\widetilde{A}$.

We will show that for most of the r-sets $\widetilde{B} \in [B]^r$, the pair $\widetilde{A}, \widetilde{B}$ satisfies $DPC(\delta'')$ where $\delta'' = 3\delta$. This will imply by Proposition 6.6 Part 1 that w restricted to $\widetilde{A} \cup \widetilde{B}$ is $(\alpha, \widetilde{\delta})$-uniform as required.

Fix a pair $\{i, i'\} \in [\widetilde{A}]^2$ that is not an edge of G_A. For each $j \in B$, define $x_j = w_{i,j} w_{i',j}$. By the choice of i, i', we have

$$\sum_{j \in B} x_j = \sum_{j \in B} w_{i,j} w_{i',j} = \vec{w_i} \cdot \vec{w_{i'}} = (\alpha \pm \delta')^2 n.$$

By Lemma 4.3 with $c = (\alpha \pm \delta')^2$ and $\xi = \delta'$, all but $e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $\widetilde{B}$ satisfy

$$\sum_{j \in \widetilde{B}} w_{i,j} w_{i',j} = \sum_{j \in \widetilde{B}} x_j = (\alpha \pm 3\delta')^2 r. \tag{6}$$

Therefore all but $\binom{r}{2} e^{-\sqrt{r}/2}\binom{n}{r} < e^{-\sqrt{r}/3}\binom{n}{r}$ of the r-sets $\widetilde{B} \in [B]^r$ satisfy (6) for all choices $\{i, i'\} \in [\widetilde{A}]^2 - E(G_A)$. Consequently, the number of choices of $\widetilde{A}, \widetilde{B}$ that do not satisfy $DPC(\delta'')$ is at most

$$2e^{-\sqrt{r}/2}\binom{n}{r}^2 + e^{-\sqrt{r}/3}\binom{n}{r}^2 < e^{-\sqrt{r}/10}\binom{n}{r}^2.$$

For a pair $\widetilde{A}, \widetilde{B}$ that satisfies $DPC(\delta'')$, Proposition 6.6 Part 1 implies that the restriction of w to $\widetilde{A} \cup \widetilde{B}$ is $(\alpha, \widetilde{\delta})$-uniform. □

6.5 The Induction step: Extension to hypergraphs

In this section we prove the $l = 3$ and unweighted case of Theorem 6.2.

Let $\alpha, \widetilde{\delta} > 0$ be given and let $\mathcal{H} = \mathcal{H}(A, B, C)$ be a 3-partite triple system with parts A, B, C each of size n. We will show that there exist δ, r, n_0 such that, if $n > n_0$, and $\mathcal{H}$ is (α, δ)-uniform, then for all but $e^{-r^{1/3}/10}\binom{n}{r}^3$ choices of triples of r-sets $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r, \widetilde{C} \in [C]^r$, the hypergraph $\widetilde{\mathcal{H}} = \mathcal{H}(\widetilde{A}, \widetilde{B}, \widetilde{C})$ is $(\alpha, \widetilde{\delta})$-uniform. Throughout the proof, our constants implicitly satisfy the following hierarchy:

$$\alpha \gg \widetilde{\delta} \gg \widetilde{\delta}^{(2)} \gg \epsilon'' \gg \epsilon' \gg \epsilon \ggg 1/t \gg 1/r \gg \delta \gg 1/n. \tag{7}$$

Remark 6.8. *In the notation above, by $x \gg y$ we mean that $x > (100y)^{100}$ holds. The relation $\epsilon \ggg 1/t$ between ϵ and t is given by the Regularity Lemma (Lemma 2.3 applied with $l = 2$). Since t is guaranteed to exist only for t being a very fast growing function of ϵ (see [16]), we will assume that $\epsilon > 1/\log t > 1/\log r$.*

6.5.1 Outline of Proof

In this subsection we give the outline of the proof. Let $\alpha, \widetilde{\delta}$ be given.

Define $\Gamma = (A \times B, C)$ to be the bipartite graph with edge set

$$E(\Gamma) = \{((a,b),c) : (a,b,c) \in E(\mathcal{H})\}.$$

Observe that Γ has the same density as $\mathcal{H}$, namely α.

Step 1: Apply the Szemerédi Regularity Lemma to Γ with parameter $\epsilon > 0$ chosen appropriately (see Remark 6.9 after Step 4). This gives a partition

$$A \times B = G_1 \cup \cdots \cup G_t, \qquad |G_i| = n^2/t, \quad i = 1, \ldots, t,$$

$$C = C_1 \cup \cdots \cup C_t, \qquad |C_j| = n/t, \quad j = 1, \ldots, t,$$

so that all but ϵt^2 pairs G_i, C_j are ϵ-regular.

For each $j = 1, \ldots, t$, consider the weighted graph $H_j(A, B)$ with weight w_j, where $w_j(a,b)$ is the number of $c \in C_j$ for which $\{a,b,c\}$ is an edge in $\mathcal{H}$.

Step 2: Fix $\widetilde{\delta}^{(2)} \ll \widetilde{\delta}$ and apply the induction hypothesis $WT(2)$, or equivalently, Theorem 6.7 to $H_j(A,B)$ to conclude that for most choices of r-sets $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$, w_j restricted to $\widetilde{A} \cup \widetilde{B}$ is $(\alpha, \widetilde{\delta}^{(2)})$-uniform. Since $WT(2)$ is a statement which remains true with a smaller choice of δ and a larger choice of r, we can choose δ and r such that (7) holds.

Let $\Gamma_{i,j}$ be the subgraph of Γ induced by $G_i \cup C_j$. Let $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B}) \subset \Gamma_{i,j}$ be the subgraph induced by $[G_i \cap (\widetilde{A} \times \widetilde{B})] \cup C_j$.

Step 3: Choose ϵ' appropriately (see Remark 6.9 after Step 4) and show that for most $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$, $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ satisfies $DPC(3\epsilon')$ for most i, j.

Now fix $\widetilde{A}, \widetilde{B}$ satisfying the conditions of both Steps 2 and 3. Let $\alpha_{i,j}$ be the density of $\Gamma_{i,j}$. Let $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C}) \subset \Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ be the subgraph consisting only of edges touching $\widetilde{C}$.

Step 4: Choose $\epsilon'' \ll \widetilde{\delta}^{(2)}$ and show that for most $\widetilde{C} \in [C]^r$, the graph $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C})$ is $(\alpha_{i,j}, \epsilon'')$-uniform.

Remark 6.9. *We fixed $\epsilon'' \ll \widetilde{\delta}^{(2)}$ above. The proof of Step 4 relies on Proposition 6.4 part 1 which yields $\epsilon' = \epsilon'(\epsilon'')$ considered in Step 3. Furthermore, the proof of Step 3 is based on Proposition 6.4 part 2, which yields $\epsilon = \epsilon(\epsilon')$ used in Step 1.*

Now fix $\widetilde{C}$ satisfying the conditions of Step 4.

Step 5: Show that $\mathcal{H}(\widetilde{A}, \widetilde{B}, \widetilde{C})$ is $(\alpha, \widetilde{\delta})$-uniform.

In the following subsections, we provide the details of Steps 2–5.

6.5.2 Step 2

Claim 6.10. *For all but $e^{-\sqrt{r}/10}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$ and all $j \in [t]$*

(*) $d(A', B', C_j) = \alpha \pm \widetilde{\delta}^{(2)}$ *whenever* $A' \subset \widetilde{A}, B' \subset \widetilde{B}$ *with* $|A'| > \widetilde{\delta}^{(2)}|\widetilde{A}|, |B'| > \widetilde{\delta}^{(2)}|\widetilde{B}|$.

Proof. For each $j \in [t]$, consider the weighted bipartite graph $H_j(A, B)$ and weights

$$w_j(a, b) = |\{c \in C_j; (a, b, c) \in \mathcal{H}\}|.$$

Because $\mathcal{H}$ is (α, δ)-uniform and $|C_j| = n/t \gg n/r \gg \delta n$, we have $d(A', B', C_j) = \alpha \pm \delta$ whenever $A' \subset A, B' \subset B$ with $|A'| \geq \delta|A|, |B'| \geq \delta|B|$. Consequently, w_j is (α, δ)-uniform. By the induction assumption, Theorem 6.7, applied to the weighted graph $H_j(A, B)$, we conclude that for all but $e^{-\sqrt{r}/10}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$, the weight function w_j restricted to $\widetilde{A} \cup \widetilde{B}$ is $(\alpha, \widetilde{\delta}^{(2)})$-uniform. Translating this statement to $\mathcal{H}$, yields the Claim. □

6.5.3 Step 3

Let $\Gamma_{i,j}$ be the subgraph of Γ induced by $G_i \cup C_j$. Let $\alpha_{i,j}$ be the density of $\Gamma_{i,j}$. Let $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ be the subgraph of $\Gamma_{i,j}$ induced by $[G_i \cap (\widetilde{A} \times \widetilde{B})] \cup C_j$. For $(a, b) \in A \times B$, let

$$\Gamma_j(a, b) = \{c \in C_j : (a, b, c) \in E(\mathcal{H})\}.$$

In this Subsection we prove the following Claim.

Claim 6.11. *For all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$,*

$$|(\widetilde{A} \times \widetilde{B}) \cap G_i| = \frac{r^2}{t}(1 \pm 1/10). \tag{8}$$

For all ϵ, there exists an ϵ' such that the following holds: If $\Gamma_{i,j}$ is ϵ-regular, then the following are also satisfied for all but $4e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$:

$$\textit{All but } 2\epsilon'\tfrac{r^2}{t} \textit{ elements } (a, b) \in (\widetilde{A} \times \widetilde{B}) \cap G_i \textit{ have } |\Gamma_j(a, b)| = (\alpha_{i,j} \pm \epsilon')|C_j| \tag{9}$$

$$\textit{All but } 2\epsilon'\tfrac{r^4}{t^2} \textit{ elements } \{(a, b), (a', b')\} \in [(\widetilde{A} \times \widetilde{B}) \cap G_i]^2 \textit{ have } |\Gamma_j(a, b) \cap \Gamma_j(a', b')| = (\alpha_{i,j}^2 \pm \epsilon')|C_j|. \tag{10}$$

Furthermore, $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ has density $\alpha_{i,j} \pm 5\epsilon'$.

Proof. We apply the proof of Proposition 4.5 to the bipartite graph with bipartition $A \cup B$ whose edges consist of all elements of G_i. This immediately gives (8) for all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices of $\widetilde{A}, \widetilde{B}$.

We now focus on proving (9) and (10). Since $\Gamma_{i,j}$ is ϵ-regular of density $\alpha_{i,j}$, it is $(\alpha_{i,j}, \epsilon)$-uniform. Proposition 6.4 Part 2 implies that $\Gamma_{i,j}$ satisfies $DPC(\epsilon')$ for some ϵ'. In other words,

a) All but $\epsilon'|G_i|$ elements $(a,b) \in G_i$ have $|\Gamma_j(a,b)| = (\alpha_{i,j} \pm \epsilon')|C_j|$, and
b) All but $\epsilon'|G_i|^2$ elements $\{(a,b),(a',b')\} \in [G_i]^2$ have $|\Gamma_j(a,b) \cap \Gamma_j(a',b')| = (\alpha_{i,j}^2 \pm \epsilon')|C_j|$.

Let $\mathcal{B}_{i,j}$ and $\mathcal{B}'_{i,j}$ denote the exceptions in a) or b). More precisely, let $\mathcal{B}_{i,j}$ denote the set of elements $(a,b) \in [G_i]$ for which $|\Gamma_j(a,b)| \neq (\alpha_{i,j} \pm \epsilon')n/t$. Let $\mathcal{B}'_{i,j}$ denote the set of elements $\{(a,b),(a',b')\} \in [G_i]^2$ for which $|\Gamma_j(a,b) \cap \Gamma_j(a',b')| \neq (\alpha_{i,j}^2 \pm \epsilon')n/t$.

By a) and b) above, $|\mathcal{B}_{i,j}| < \epsilon'|G_i| = \epsilon' \frac{n^2}{t}$ and $|\mathcal{B}'_{i,j}| < \epsilon'|G_i|^2 = \epsilon' \frac{n^4}{t^2}$. We now apply the proof of Proposition 4.5 to the bipartite graph with bipartition $A \cup B$ whose edges consist of all elements of G_i. This gives (9) and (10) for all but $4e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices of $\widetilde{A}, \widetilde{B}$.

For all other choices $\widetilde{A}, \widetilde{B}$, the number of edges in $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ is at most

$$2\epsilon'(r^2/t)|C_j| + |(\widetilde{A} \times \widetilde{B}) \cap G_i|(\alpha_{i,j} + \epsilon')|C_j|.$$

Since $|(\widetilde{A} \times \widetilde{B}) \cap G_i| > \frac{9r^2}{10t} > \frac{2r^2}{3t}$, this is at most

$$3\epsilon'|(\widetilde{A} \times \widetilde{B}) \cap G_i||C_j| + (\alpha_{i,j} + \epsilon')|(\widetilde{A} \times \widetilde{B}) \cap G_i||C_j| < (\alpha_{i,j} + 5\epsilon')|(\widetilde{A} \times \widetilde{B}) \cap G_i||C_j|.$$

On the other hand, the number of edges in $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ is at least

$$(|(\widetilde{A} \times \widetilde{B}) \cap G_i| - 2\epsilon' r^2/t)(\alpha_{i,j} - \epsilon')|C_j|.$$

Since $|(\widetilde{A} \times \widetilde{B}) \cap G_i| > \frac{2r^2}{3t}$, this is at least

$$(1 - 3\epsilon')|(\widetilde{A} \times \widetilde{B}) \cap G_i|(\alpha_{i,j} - \epsilon')|C_j| > (\alpha_{i,j} - 5\epsilon')|(\widetilde{A} \times \widetilde{B}) \cap G_i||C_j|.$$

Therefore the density of $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B})$ is $\alpha_{i,j} \pm 5\epsilon'$. □

Definition 6.12. *The set C_j is good if for all but $\sqrt{\epsilon}t$ of the i's, the graph $\Gamma_{i,j}$ is ϵ-regular.*

Since the number of ϵ-irregular pairs G_i, C_j is at most ϵt^2, we infer that at most $\sqrt{\epsilon}t$ of the C_j's are not good.

We now fix $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$ satisfying (*) in Claim 6.10, and each of (8), (9), (10) for every i,j for which $\Gamma_{i,j}$ is ϵ-regular. The number of exceptions to (*) is at most $e^{-\sqrt{r}/10}\binom{n}{r}^2$, and the number of exceptions to (8), (9), (10) over all i,j is at most $(t^2 e^{-\sqrt{r}/3})\binom{n}{r}^2$. Therefore the total number of exceptions is at most $2e^{-\sqrt{r}/10}\binom{n}{r}^2$, and there are at least $(1 - 2e^{-\sqrt{r}/10})\binom{n}{r}^2$ pairs $\widetilde{A}, \widetilde{B}$ remaining. For the rest of the proof, fix one such pair $\widetilde{A}, \widetilde{B}$.

6.5.4 Step 4

Given an r-set $\widetilde{C} \in [C]^r$, we let $\widetilde{C}_j = C_j \cap \widetilde{C}$.

Claim 6.13. *For all but* $(t+tr^2)e^{-\sqrt{r}}\binom{n}{r}$ *of the* r*-sets* $\widetilde{C}\in[C]^r$*, and all* $j\in[t]$

1) $|\widetilde{C}_j|=\frac{r}{n}|C_j|(1\pm 1/\log r)=\frac{r}{t}(1\pm 1/\log r)$,

and for all $(a,b)\in\widetilde{A}\times\widetilde{B}$,

2) $|\Gamma_j(a,b)\cap\widetilde{C}_j|=|\Gamma_j(a,b)|\frac{|\widetilde{C}_j|}{|C_j|}(1\pm 1/\log r)$ *if* $|\Gamma_j(a,b)|\geq n/r^{1/4}$*, and*

3) $|\Gamma_j(a,b)\cap\widetilde{C}_j|<r^{19/20}$ *if* $|\Gamma_j(a,b)|\leq n/r^{1/10}$.

Proof. By Proposition 4.2 all but $te^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $\widetilde{C}\in[C]^r$ satisfy 1). Similarly, by Lemma 4.1, for a fixed $(a,b)\in\widetilde{A}\times\widetilde{B}$, all but $te^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $\widetilde{C}\in[C]^r$ satisfy 2) or 3). Since $|\widetilde{A}\times\widetilde{B}|=r^2$, the total number of exceptions to 2) or 3) is at most $tr^2e^{-\sqrt{r}}\binom{n}{r}$. Therefore the total number of exceptions to 1), 2) or 3) is at most $(t+tr^2)e^{-\sqrt{r}}\binom{n}{r}$. $\square$

Claim 6.14. *Let* $\Gamma_{i,j}$ *be* ϵ*-regular. Suppose that* $\alpha_{i,j}\geq\sqrt{2\epsilon'}$*. Then all but* $(r^2+r^4)e^{-\sqrt{r}}\binom{n}{r}$ *of the* r*-sets* $\widetilde{C}\in[C]^r$ *satisfy*

1) $|\Gamma_j(a,b)\cap\widetilde{C}_j|=(\alpha_{i,j}\pm 2\epsilon')|\widetilde{C}_j|$ *for all but* $2\epsilon' r^2/t$ *elements* $(a,b)\in G_i\cap(\widetilde{A}\cap\widetilde{B})$*, and*

2) $|\Gamma_j(a,b)\cap\Gamma_j(a',b')\cap\widetilde{C}_j|=(\alpha_{i,j}^2\pm 2\epsilon')|\widetilde{C}_j|$ *for all but* $2\epsilon' r^4/t^2$ *elements* $\{(a,b),(a',b')\}\in[G_i\cap(\widetilde{A}\times\widetilde{B})]^2$.

Proof. By (9) and (10),

1') $|\Gamma_j(a,b)|=(\alpha_{i,j}\pm\epsilon')|C_j|$ for all but $2\epsilon' r^2/t$ of the $(a,b)\in(\widetilde{A}\times\widetilde{B})\cap G_i$
and

2') $|\Gamma_j(a,b)\cap\Gamma_j(a',b')|=(\alpha_{i,j}^2\pm\epsilon')|C_j|$ for all but $2\epsilon' r^4/t^2$ elements $\{(a,b),(a',b')\}\in[(\widetilde{A}\times\widetilde{B})\cap G_i]^2$.

Since $\alpha_{i,j}\geq\sqrt{2\epsilon'}$, $\alpha_{i,j}-\epsilon'>\alpha_{i,j}/2$. Also, $\epsilon'\ggg 1/r$ (and due to Remark 6.8 we further have $\epsilon'>2/\log r$) hence 1') yields

$$|\Gamma_j(a,b)|\geq(\alpha_{i,j}-\epsilon')|C_j|\geq\alpha_{i,j}|C_j|/2>\epsilon' n/t>n/t^2>n/r^{1/4}.$$

Consequently Lemma 4.1 implies that for all but $r^2e^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $\widetilde{C}\in[C]^r$ and all (a,b) that are not exceptions in 1'),

$$|\Gamma_j(a,b)\cap\widetilde{C}_j|=|\Gamma_j(a,b)|\frac{|\widetilde{C}_j|}{|C_j|}(1\pm 1/\log r)=(\alpha_{i,j}\pm\epsilon')|\widetilde{C}_j|(1\pm 1/\log r)=(\alpha_{i,j}\pm 2\epsilon')|\widetilde{C}_j|.$$

Therefore 1) holds for all these sets $\widetilde{C}$.

Similarly, $\alpha_{i,j}^2-\epsilon'>\alpha_{i,j}^2/2$, and therefore

$$(\alpha_{i,j}^2-\epsilon')|C_j|\geq\alpha_{i,j}^2|C_j|/2>\epsilon' n/t>n/r^{1/4}$$

Now we apply 2') and Lemma 4.1 to conclude 2).

The total number of exceptions to 1) and 2) is at most $(r^2+r^4)e^{-\sqrt{r}}\binom{n}{r}$ as needed. □

For the rest of the proof, we fix a $\widetilde{C}$ that satisfies the conditions of Claim 6.13, and Claim 6.14 for every pair i,j for which $\Gamma_{i,j}$ is ϵ-regular. The number of exceptions is at most

$$(t+tr^2+r^2t^2+r^4t^2)e^{-\sqrt{r}}\binom{n}{r} < e^{-\sqrt{r}/3}\binom{n}{r}.$$

Let $\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C}) \subset \Gamma_{i,j}(\widetilde{A}\times\widetilde{B})$ be the subgraph consisting only of edges touching $\widetilde{C}$.

Claim 6.15. *Suppose that $\Gamma_{i,j}$ is ϵ-regular. Then*

$$\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C}) \text{ is } (\alpha_{i,j},\epsilon'')\text{–uniform.} \tag{11}$$

Proof. We distinguish between two cases.
Case 1: $\alpha_{i,j} > \sqrt{2\epsilon'}$. We first observe that $\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C})$ satisfies $DPC(3\epsilon')$. This follows immediately from Claim 6.14 (see (9) and (10)) and the inequalities $2\epsilon' r^2/t < 3\epsilon'|G_i\cap(\widetilde{A}\times\widetilde{B})|$ and $2\epsilon' r^4/t^2 < 3\epsilon'|G_i\cap(\widetilde{A}\times\widetilde{B})|^2$ which bound the number of exceptions in each case. By Claim 6.11, $\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C})$ has density $\alpha_{i,j}\pm 5\epsilon'$. Now Proposition 6.4 part 1 implies that (11) holds.

For the other case we use the following trivial fact:
Fact: Suppose that a bipartite graph G has density less than $(\epsilon'')^3$, and $\tilde{\alpha} < (\epsilon'')$. Then G is $(\tilde{\alpha},\epsilon'')$-uniform.

Case 2: $\alpha_{i,j} \le \sqrt{2\epsilon'}$. We show that the number of edges in $\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C})$ is less than $(\epsilon'')^3|G_i\cap(\widetilde{A}\times\widetilde{B})||\widetilde{C}_j|$. Since $\alpha_{i,j}\le\sqrt{2\epsilon'}<\epsilon''$ (see Remark 6.8), the Fact implies that (11) holds. Let

$$S_1 = \{(a,b)\in G_i\cap(\widetilde{A}\times\widetilde{B}) : |\Gamma_j(a,b)| = (\alpha_{i,j}\pm\epsilon')|C_j| \text{ and } |\Gamma_j(a,b)|\ge n/r^{1/10}\},$$
$$S_2 = \{(a,b)\in G_i\cap(\widetilde{A}\times\widetilde{B}) : |\Gamma_j(a,b)| = (\alpha_{i,j}\pm\epsilon')|C_j| \text{ and } |\Gamma_j(a,b)| < n/r^{1/10}\},$$
$$S_3 = \{(a,b)\in G_i\cap(\widetilde{A}\times\widetilde{B}) : |\Gamma_j(a,b)| \ne (\alpha_{i,j}\pm\epsilon')|C_j|\}.$$

Then the number of edges in $\Gamma_{i,j}(\widetilde{A}\times\widetilde{B},\widetilde{C})$ is

$$\sum_{(a,b)\in S_1}|\Gamma_j(a,b)\cap\widetilde{C}| + \sum_{(a,b)\in S_2}|\Gamma_j(a,b)\cap\widetilde{C}| + \sum_{(a,b)\in S_3}|\Gamma_j(a,b)\cap\widetilde{C}|. \tag{12}$$

By Claim 6.13 part 2, and $\alpha_{i,j}\le\sqrt{2\epsilon'}$,

$$\sum_{(a,b)\in S_1}|\Gamma_j(a,b)\cap\widetilde{C}| \le |G_i\cap(\widetilde{A}\times\widetilde{B})|(\alpha_{i,j}+\epsilon')|\widetilde{C}_j|(1+1/\log r) < \frac{(\epsilon'')^3}{3}|G_i\cap(\widetilde{A}\times\widetilde{B})||\widetilde{C}_j|.$$

By Claim 6.13 parts 1 and 3, and $r\gg t$,

$$\sum_{(a,b)\in S_2}|\Gamma_j(a,b)\cap\widetilde{C}| \le |G_i\cap(\widetilde{A}\times\widetilde{B})|r^{19/20} < \frac{(\epsilon'')^3}{3}|G_i\cap(\widetilde{A}\times\widetilde{B})||\widetilde{C}_j|.$$

Since $\Gamma_{i,j}$ is ϵ-regular, (8) and (9) yield $|S_3| \le 2\epsilon' r^2/t < 3\epsilon' |G_i \cap (\widetilde{A} \times \widetilde{B})|$. Therefore

$$\sum_{(a,b)\in S_3} |\Gamma_j(a,b) \cap \widetilde{C}| \le 3\epsilon' |G_i \cap (\widetilde{A} \times \widetilde{B})||\widetilde{C}_j| < \frac{(\epsilon'')^3}{3} |G_i \cap (\widetilde{A} \times \widetilde{B})||\widetilde{C}_j|.$$

We conclude that (12) is bounded above by $(\epsilon'')^3 |G_i \cap (\widetilde{A} \times \widetilde{B})||\widetilde{C}_j|$. This completes the proof. □

6.5.5 Step 5

In the remainder of the proof, we show that $\widetilde{A}, \widetilde{B}, \widetilde{C}$ satisfies the conclusion of the Theorem. In other words, we will prove that $\mathcal{H}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \widetilde{\mathcal{C}})$ is $(\alpha, \widetilde{\delta})$-uniform. Note that the number of triples $\widetilde{A}, \widetilde{B}, \widetilde{C}$ that are exceptions is at most

$$[2\exp\{-\sqrt{r}/3\} + \exp\{-\sqrt{r}/10\}] \binom{n}{r}^3 < \exp\{-r^{1/3}/10\} \binom{n}{r}^3.$$

Pick $A' \subset \widetilde{A}, B' \subset \widetilde{B}, C' \subset \widetilde{C}$, with $|A'| > \widetilde{\delta}|\widetilde{A}|, |B'| > \widetilde{\delta}|\widetilde{B}|, |C'| \ge \widetilde{\delta}|\widetilde{C}|$. To complete the proof, we must show that

Lemma 6.16.

$$d(A', B', C') = \alpha \pm \widetilde{\delta} \tag{13}$$

Proof. Our first goal is to prove

Claim 6.17. *Let A', B' be as above, and $\widetilde{C}$ satisfy the conditions in Claim 6.13. Then*

$$d(A', B', \widetilde{C}_j) = \alpha \pm 2\widetilde{\delta}^{(2)}. \tag{14}$$

Proof. Let

$$S_1 = \{(a,b) \in A' \times B' : |\Gamma_j(a,b)| \ge n/r^{1/4}\}$$

and

$$S_2 = \{(a,b) \in A' \times B' : |\Gamma_j(a,b)| < n/r^{1/4}\}.$$

Then

$$d(A', B', \widetilde{C}_j) = \frac{e(A', B', \widetilde{C}_j)}{|A'||B'||\widetilde{C}_j|} = \frac{\sum_{(a,b)\in S_1} |\Gamma_j(a,b) \cap \widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|} + \frac{\sum_{(a,b)\in S_2} |\Gamma_j(a,b) \cap \widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|}.$$

By Definition of S_2, Claim 6.13 parts 1 and 3, and $\widetilde{\delta} \gg \epsilon \gg 1/t \gg 1/r$,

$$\frac{\sum_{(a,b)\in S_2} |\Gamma_j(a,b) \cap \widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|} < \frac{r^2 r^{19/20}}{(\widetilde{\delta})^2 r^2 (r/2t)} < \epsilon.$$

Since $\widetilde{\delta}^{(2)} \gg \epsilon$, to finish the proof it suffices to show that

$$\frac{\sum_{(a,b)\in S_1} |\Gamma_j(a,b)\cap \widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|} = \alpha \pm \frac{3}{2}\widetilde{\delta}^{(2)}. \tag{15}$$

We estimate $\sum_{(a,b)\in S_1} |\Gamma_j(a,b)\cap\widetilde{C}_j|$ by estimating $e(A',B',C_j)$, which we in turn estimate by using **(*)**:

$$e(A',B',C_j) = d(A',B',C_j)|A'||B'||C_j| = (\alpha \pm \widetilde{\delta}^{(2)})|A'||B'||C_j|. \tag{16}$$

On the other hand,

$$e(A',B',C_j) = \sum_{(a,b)\in S_1} \Gamma_j(a,b) + \sum_{(a,b)\in S_2} \Gamma_j(a,b),$$

while $\sum_{(a,b)\in S_2} \Gamma_j(a,b) < r^2 n/r^{1/4}$. Consequently,

$$\sum_{(a,b)\in S_1} \Gamma_j(a,b) = (\alpha \pm \widetilde{\delta}^{(2)})|A'||B'||C_j| \pm r^{7/4}n,$$

and therefore

$$\sum_{(a,b)\in S_1} \Gamma_j(a,b)\frac{|\widetilde{C}_j|}{|C_j|} = (\alpha \pm \widetilde{\delta}^{(2)})|A'||B'||\widetilde{C}_j| \pm 2r^{11/4}.$$

Now by the Definition of S_1 and Claim 6.13 part 2,

$$\frac{\sum_{(a,b)\in S_1} |\Gamma_j(a,b)\cap \widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|} = \frac{\sum_{(a,b)\in S_1} |\Gamma_j(a,b)|\frac{|\widetilde{C}_j|}{|C_j|}(1 \pm \frac{1}{\log r})}{|A'||B'||\widetilde{C}_j|} = \left[\alpha \pm \left(\widetilde{\delta}^{(2)} + \epsilon\right)\right](1 \pm \frac{1}{\log r}).$$

This implies (15) since $\widetilde{\delta}^{(2)} \gg \epsilon \gg 1/r$. $\square$

We now focus on C'. Let $C'_j = C' \cap C_j$. Then

$$d(A',B',C') = \sum_{j=1}^{t} \frac{|C'_j|}{|C'|}\, d(A',B',C'_j). \tag{17}$$

Break the sum in (17) into three parts depending first on whether C_j is good or not, and then on whether $|C'_j|$ is big or not:

$$J_1 = \{j : C_j \text{ is good and } |C'_j| > \sqrt{\epsilon''}r/t\},$$

$$J_2 = \{j : C_j \text{ is good and } |C'_j| \le \sqrt{\epsilon''}r/t\},$$
$$J_3 = \{j : C_j \text{ is not good}\}.$$

For $i = 1, 2, 3$, define

$$\sum^{(i)} = \sum_{j \in J_i} \frac{|C'_j|}{|C'|}\ d(A', B', C'_j).$$

Next we show that the contribution of $\sum^{(2)}$ and $\sum^{(3)}$ is negligible. By definition

$$\sum^{(2)} \leq \sum_{|C'_j| \leq \sqrt{\epsilon''} r/t} \frac{|C'_j|}{|C'|} < \frac{t\ \sqrt{\epsilon''} r/t}{\widetilde{\delta}|\widetilde{C}|} = \frac{\sqrt{\epsilon''} r}{\widetilde{\delta} r} < \frac{(\epsilon'')^{1/3}}{2}. \tag{18}$$

Since $C'_j \subseteq \widetilde{C}_j$, Claim 6.13 part 1 implies that $|C'_j| \leq 2r/t$. Moreover, at most $\sqrt{\epsilon} t$ of the C_j's are not good. Consequently,

$$\sum^{(3)} \leq \sum_{j \in J_3} \frac{|C'_j|}{|C'|} < \frac{\sqrt{\epsilon}\ t\ \ \max_j |C'_j|}{\widetilde{\delta}|\widetilde{C}|} < \frac{\sqrt{\epsilon}\ t\ 2r/t}{\widetilde{\delta} r} = \frac{2\sqrt{\epsilon}}{\widetilde{\delta}} < \frac{(\epsilon'')^{1/3}}{2}. \tag{19}$$

We have shown that

$$\sum^{(1)} \leq d(A', B', C') \leq \sum^{(1)}\ +\ (\epsilon'')^{1/3}. \tag{20}$$

To complete the proof of Lemma 6.16, it suffices to prove the following Claim.

Claim 6.18. *Let $j \in J_1$. Then*

$$d(A', B', C'_j) = \alpha \pm \widetilde{\delta}/2. \tag{21}$$

Before we prove Claim 6.18, we show how to complete the proof of Lemma 6.16. Indeed, if we assume Claim 6.18, then (17) and (20) together with $\epsilon \ll \widetilde{\delta}$ yield

$$d(A', B', C') \leq \left[\sum_{j \in J_1} \frac{|C'_j|}{|C'|}(\alpha + \widetilde{\delta}/2)\right] + (\epsilon'')^{1/3} < \alpha + \widetilde{\delta}/2 + (\epsilon'')^{1/3} < \alpha + \widetilde{\delta}.$$

On the other hand,

$$d(A', B', C') \geq \sum^{(1)} \geq \sum_{j \in J_1} \left\{\frac{|C'_j|}{|C'|}(\alpha - \widetilde{\delta}/2)\right\}$$

$$= (\alpha - \widetilde{\delta}/2) \left[\sum_{j=1}^{t} \frac{|C'_j|}{|C'|} - \sum_{j \in J_2 \cup J_3} \frac{|C'_j|}{|C'|}\right] \geq (\alpha - \widetilde{\delta}/2)(1 - (\epsilon'')^{1/3}) \geq (\alpha - \widetilde{\delta}),$$

where the second last inequality comes from the fact that

$$\sum_{j \in J_2} \frac{|C'_j|}{|C'|} \leq \sum_{|C'_j| \leq \sqrt{\epsilon''} r/t} \frac{|C'_j|}{|C'|} < \frac{(\epsilon'')^{1/3}}{2} \qquad \text{and} \qquad \sum_{j \in J_3} \frac{|C'_j|}{|C'|} < \frac{(\epsilon'')^{1/3}}{2}$$

proved in (18) and (19).

Proof of Claim 6.18. Let $G'_i = G_i \cap (A' \times B')$, and let $e(G'_i, C'_j)$ be the number of edges between G'_i and C'_j in $\Gamma_{i,j}$. Then

$$e(A', B', C'_j) = \sum_{i=1}^{t} e(G'_i, C'_j). \tag{22}$$

Break the sum in (22) into three parts by the following sets:

$$I_1^{(j)} = \{i : |G'_i| > \sqrt{\epsilon''} r^2/t \text{ and } \Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C}) \text{ is } (\alpha_{i,j}, \epsilon'')\text{–uniform}\},$$

$$I_2^{(j)} = \{i : |G'_i| > \sqrt{\epsilon''} r^2/t \text{ and } \Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C}) \text{ is not } (\alpha_{i,j}, \epsilon'')\text{–uniform}\},$$

$$I_3^{(j)} = \{i : |G'_i| \le \sqrt{\epsilon''} r^2/t\}.$$

Then

$$\sum_{i=1}^{t} e(G'_i, C'_j) = \sum_{i \in I_1^{(j)}} e(G'_i, C'_j) + \sum_{i \in I_2^{(j)}} e(G'_i, C'_j) + \sum_{i \in I_3^{(j)}} e(G'_i, C'_j).$$

We now bound the contribution from the last two sums. If $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C})$ is not $(\alpha_{i,j}, \epsilon'')$-uniform, then by Claim 6.15, $\Gamma_{i,j}$ is not ϵ-regular. Since $j \in J_1$, C_j is good and therefore $\Gamma_{i,j}$ is not ϵ-regular for at most $\sqrt{\epsilon} t$ of the G_i's. Also, $G'_i = G_i \cap (A' \times B') \subseteq G_i \cap (\widetilde{A} \times \widetilde{B})$, hence by (8),

$$|G'_i| \le |G_i \cap (\widetilde{A} \times \widetilde{B})| < 2r^2/t. \tag{23}$$

Consequently,

$$\sum_{i \in I_2^{(j)}} e(G'_i, C'_j) \le \sqrt{\epsilon} t |G'_i||C'_j| < \sqrt{\epsilon} t (2r^2/t)|C'_j| = 2\sqrt{\epsilon} r^2 |C'_j|.$$

Clearly

$$\sum_{i \in I_3^{(j)}} e(G'_i, C'_j) \le |C'_j| \sum_{|G'_i| \le \sqrt{\epsilon''} r^2/t} |G'_i| \le |C'_j| \sqrt{\epsilon''} r^2.$$

Now we estimate $\sum_{i \in I_1^{(j)}} e(G'_i, C'_j)$. For an $i \in I_1^{(j)}$ and $j \in J_1$ we have

(a) $\Gamma_{i,j}(\widetilde{A} \times \widetilde{B}, \widetilde{C})$ is $(\alpha_{i,j}, \epsilon'')$-uniform
(b) $|G'_i| > \sqrt{\epsilon''} r^2/t > \epsilon'' |G_i \cap (\widetilde{A} \times \widetilde{B})|$
(c) $|C'_j| > \sqrt{\epsilon''} r/t > (\sqrt{\epsilon''}/2)|\widetilde{C}_j| > \epsilon'' |\widetilde{C}_j|$.

Therefore

$$\sum_{i \in I_1^{(j)}} e(G'_i, C'_j) = \sum_{i \in I_1^{(j)}} |G'_i||C'_j|(\alpha_{i,j} \pm \epsilon''). \tag{24}$$

Now

$$d(A', B', C'_j) = \frac{\sum_{i\in I_1^{(j)}} e(G'_i, C'_j) + \sum_{i\in I_2^{(j)}} e(G'_i, C'_j) + \sum_{i\in I_3^{(j)}} e(G'_i, C'_j)}{|A'||B'||C'_j|}. \tag{25}$$

Since $|A'| > \widetilde{\delta}|\widetilde{A}|, |B'| > \widetilde{\delta}|\widetilde{B}|$,

$$\frac{\sum_{i\in I_2^{(j)}} e(G'_i, C'_j)}{|A'||B'||C'_j|} < \frac{2\sqrt{\epsilon}r^2|C'_j|}{|A'||B'||C'_j|} = \frac{2\sqrt{\epsilon}r^2}{|A'||B'|} < \frac{2\sqrt{\epsilon}}{(\widetilde{\delta})^2} < \frac{\widetilde{\delta}}{16}, \tag{26}$$

and

$$\frac{\sum_{i\in I_3^{(j)}} e(G'_i, C'_j)}{|A'||B'||C'_j|} < \frac{|C'_j|\sqrt{\epsilon''}r^2}{|A'||B'||C'_j|} = \frac{\sqrt{\epsilon''}r^2}{|A'||B'|} < \frac{\sqrt{\epsilon''}}{(\widetilde{\delta})^2} < \frac{\widetilde{\delta}}{16}. \tag{27}$$

Consequently in view of (25), to finish the proof of the Claim it suffices to prove the following Subclaim

Subclaim 6.19. *For every $j \in J_1$,*

$$\frac{\sum_{i\in I_1^{(j)}} e(G'_i, C'_j)}{|A'||B'||C'_j|} = \alpha \pm \frac{\widetilde{\delta}}{4}. \tag{28}$$

Proof. Using (24) for $j \in J_1$ we have

$$\frac{\sum_{i\in I_1^{(j)}} e(G'_i, C'_j)}{|A'||B'||C'_j|} = \frac{\sum_{i\in I_1^{(j)}} |G'_i|(\alpha_{i,j} \pm \epsilon'')}{|A'||B'|}. \tag{29}$$

Now we use Claim 6.17 (Equation (14)) as well as (a) and (b) above:

$$\alpha \pm 2\widetilde{\delta}^{(2)} = d(A', B', \widetilde{C}_j) = \frac{\sum_{i=1}^{t} e(G'_i, \widetilde{C}_j)}{|A'||B'||\widetilde{C}_j|} = \frac{\sum_{i\in I_1^{(j)}} |G'_i||\widetilde{C}_j|(\alpha_{i,j} \pm \epsilon'')}{|A'||B'||\widetilde{C}_j|} + \frac{\sum_{i\notin I_1^{(j)}} e(G'_i, \widetilde{C}_j)}{|A'||B'||\widetilde{C}_j|}. \tag{30}$$

By arguments similar to those used for (26) and (27),

$$\frac{\sum_{i\notin I_1^{(j)}} e(G'_i, \widetilde{C}_j)}{|A'||B'||\widetilde{C}_j|} < \frac{\sum_{i\notin I_1^{(j)}} |G'_i||\widetilde{C}_j|}{|A'||B'||\widetilde{C}_j|} < \frac{\widetilde{\delta}}{8}. \tag{31}$$

Consequently, by (31) and (30)

$$\frac{\sum_{i\in I_1^{(j)}} |G'_i|(\alpha_{i,j} \pm \epsilon'')}{|A'||B'|} = \frac{\sum_{i\in I_1^{(j)}} |G'_i||\widetilde{C}_j|(\alpha_{i,j} \pm \epsilon'')}{|A'||B'||\widetilde{C}_j|} = \alpha \pm \left(2\widetilde{\delta}^{(2)} + \frac{\widetilde{\delta}}{8}\right) = \alpha \pm \frac{\widetilde{\delta}}{4}.$$

Equation (28) now follows from (29). This completes the proof of Subclaim 6.19, Claim 6.18, and the Theorem. □

6.5.6 Comments on the weighted case

In this subsection we briefly describe the relatively minor modifications needed to prove $WT(l) \Longrightarrow WT(l+1)$ for $l > 2$. We explain only the implication $WT(l) \Longrightarrow UWT(l+1)$, since extending the result to the weighted case is essentially identical to our proof for the unweighted conclusion. We only need to add notation that not only counts edges, but counts edges based on their weight.

Our focus is therefore on $WT(l) \Longrightarrow UWT(l+1)$ for $l > 2$. Our strategy is the same with slight modifications to Steps 1–3.

Let $\alpha, \widetilde{\delta} > 0$ be given and let $\mathcal{H} = \mathcal{H}(A_1, \ldots, A_{l+1})$ be an $(l+1)$-partite $(l+1)$-graph with parts $A_1, \ldots, A_{l+1}$ each of size n. We will show that there exist δ, r, n_0 such that, if $n > n_0$, and $\mathcal{H}$ is (α, δ)-uniform, then for all but $e^{-r^{1/(l+1)}/10}\binom{n}{r}^l$ choices of $(l+1)$-tuples of r-sets $\widetilde{A}_i \in [A_i]^r, i \in [l+1]$, the hypergraph $\widetilde{\mathcal{H}} = \mathcal{H}(\widetilde{A}_1, \ldots, \widetilde{A}_{l+1})$ is $(\alpha, \widetilde{\delta})$-uniform. Throughout the proof, our constants implicitly satisfy the following hierarchy:

$$\alpha \gg \widetilde{\delta} \gg \widetilde{\delta}^{(l)} \gg \epsilon'' \gg \epsilon' \gg \epsilon \gg 1/t \gg 1/r \gg \delta \gg 1/n.$$

Define $\Gamma = (A_1 \times \cdots \times A_l, A_{l+1})$ to be the bipartite graph with edge set

$$E(\Gamma) = \{(a_1, \ldots, a_l), a_{l+1}) : (a_1, \ldots, a_{l+1}) \in E(\mathcal{H})\}.$$

Observe that Γ has the same density as $\mathcal{H}$, namely α.

Step 1: Apply the Szemerédi Regularity Lemma with parameter $\epsilon > 0$ to Γ. This gives a partition

$$A_1 \times \cdots \times A_l = G_1 \cup \cdots \cup G_t, \qquad |G_i| = n^2/t, \quad i = 1, \ldots, t,$$
$$A_{l+1} = C_1 \cup \cdots \cup C_t, \qquad |C_j| = n/t, \quad j = 1, \ldots, t,$$

so that all but ϵt^2 pairs G_i, C_j are ϵ-regular.

For each $j = 1, \ldots, t$, consider the weighted l-graph $H_j(A_1, \ldots, A_l)$ with weight w_j, where $w_j(a_1, \ldots, a_l)$ is the number of $a_{l+1} \in A_{l+1}$ for which $\{a_1, \ldots, a_l, a_{l+1}\}$ is an edge in $\mathcal{H}$.

Step 2: Apply the induction hypothesis $WT(l)$ to $H_j(A_1, \ldots, A_l)$ to conclude that for most choices of r-sets $\widetilde{A}_i \in [A_i]^r, i \in [l]$, w_j restricted to $\widetilde{A}_1 \cup \cdots \cup \widetilde{A}_l$ is $(\alpha, \widetilde{\delta}^{(l)})$-uniform.

Let $\Gamma_{i,j}$ be the subgraph of Γ induced by $G_i \cup C_j$. Let $\Gamma_{i,j}(\widetilde{A}_1 \times \cdots \times \widetilde{A}_l) \subset \Gamma_{i,j}$ be the subgraph induced by $[G_i \cap (\widetilde{A}_1 \times \cdots \times \widetilde{A}_l)] \cup C_j$.

Step 3: Show that for most $\widetilde{A}_i \in [A_i]^r, i \in [l]$, $\Gamma_{i,j}(\widetilde{A}_i \times \cdots \times \widetilde{A}_l)$ satisfies $DPC(3\epsilon')$ for most i, j.

Fix $\widetilde{A}_1, \ldots, \widetilde{A}_l$ satisfying Steps 2) and 3).

Step 4: Show that for most $\widetilde{A}_{l+1} \in [A_{l+1}]^r$, $\Gamma_{i,j}(\widetilde{A}_1 \times \cdots \times \widetilde{A}_l, \widetilde{A}_{l+1})$ is $(\alpha_{i,j}, \epsilon'')$-uniform.

Now fix $\widetilde{A}_{l+1}$ satisfying the conditions of Step 4.

Step 5: Show that $\mathcal{H}(\widetilde{A}_1, \cdots, \widetilde{A}_l, \widetilde{A}_{l+1})$ is $(\alpha, \widetilde{\delta})$-uniform.

7 Appendix

7.1 Proofs from Section 4

Lemma 4.1: Let $|X| = x$, and $|Y| = y$ with $Y \subseteq X$. Let $R \in [X]^r$ be a random r-set chosen independently from the family of all r-sets of X. Let $r > r_0$ and $3/2 > \sigma \geq 1/\log r$. Then with probability greater than $1 - e^{-\sqrt{r}}$, the following holds:

a) $|R \cap Y| = \frac{ry}{x}(1 \pm \sigma)$ if $y \geq x/r^{1/4}$,

b) $|R \cap Y| < r^{19/20}$ if $y \leq x/r^{1/10}$.

Proof. The following two inequalities apply to a binomially distributed random variable $X = Bi(N, p)$ with mean μ (see [17] pages 27-28).

1) $Pr(|X - \mu| \geq \epsilon\mu) \leq 2\exp\{-\frac{\epsilon^2}{3}\mu\}$ for $0 < \epsilon \leq 3/2$, and

2) $Pr(X \geq a) \leq \exp\{-a\}$ for $a \geq 7\mu$. Here we use the fact that these estimates apply to the hypergeometric distribution (see [17] pages 29-30). In the range $y \geq x/r^{1/4}$ we use 1): here $\mu = ry/x$, $0 < \epsilon = \sigma \leq 3/2$ and thus due to our assumption on r ($r > r_0, \sigma > 1/\log r$) we infer that

$$Pr(||R \cap Y| - ry/x| > \sigma ry/x) < 2\exp\{-\frac{\sigma^2}{3}ry/x\} < \exp\{-\sqrt{r}\}.$$

In the range $y \leq x/r^{1/10}$, we use 2) with $a = r^{19/20}$. Notice that $a = r^{19/20} > 7ry/x = 7\mu$, so

$$Pr(|R \cap Y| > r^{19/20}) \leq \exp\{-r^{19/20}\} < \exp\{-\sqrt{r}\}.$$

This completes the proof. $\square$

Lemma 4.3: For all c, ξ, there exist r_0, n_0 such that if $r > r_0, n > n_0$, the following holds: If $\sum_{i\in[n]} x_i = cn$, where $0 \leq x_i \leq 1$ for each i, then all but $e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $R \in [n]^r$ satisfy $\sum_{i\in R} x_i = (c \pm \xi)r$.

Proof. The idea is to partition the elements x_i into intervals and apply the hypergeometric inequalities to each interval that has a large proportion of the x_i's. By choosing intervals of appropriate size, we ensure that the major part of the contribution of $\sum_i x_i$ comes from x_i's in these intervals.

Let $\epsilon = \xi/10$. Define $I_0 = \{i : x_i < \epsilon\}$ and for $1 \leq j \leq \frac{1-\epsilon}{\epsilon^2}$, let $I_j = \{i : l_{j-1} \leq x_i < l_j\}$, where $l_j = \epsilon + j\epsilon^2$. Let $J_1 = \{j : |I_j| > \epsilon^3 n\}$ and $J_2 = \{j : |I_j| \leq \epsilon^3 n\}$. We now apply Lemma 4.1 with $y = |I_j|$.

If $j \in J_1$, then since $r_0 = r_0(\xi)$ we have $y > \epsilon^3 n > n/r^{1/4}$. Therefore all but $e^{-\sqrt{r}}$ of the r-sets $R \in [n]^r$ have $|R \cap I_j| = \frac{|I_j|}{n}r(1 \pm \epsilon)$.

If $j \in J_2$, then all but $e^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $R \in [n]^r$ have either

i) $|R \cap I_j| = \frac{|I_j|}{n}r(1 \pm \epsilon) < 2\epsilon^3 r$ or

ii) $|R \cap I_j| < r^{19/20} < 2\epsilon^3 r$.

We conclude that for all but $\frac{1}{\epsilon^2}e^{-\sqrt{r}}\binom{n}{r}$ of the r-sets $R \in [n]^r$,

iii) $|R \cap I_j| = \frac{|I_j|}{n}r(1 \pm \epsilon)$ for all $j \in J_1$ and

iv) $|R \cap I_j| < 2\epsilon^3 r$ for all $j \in J_2$.
Since $\frac{1}{\epsilon^2}e^{-\sqrt{r}} < e^{-\sqrt{r}/2}$ it suffices to show that the conclusion of the Lemma holds for these r-sets. Now fix an R satisfying iii) and iv) and note that $|l_j|/|l_{j-1}| < 1+\epsilon$. Then by iv),

$$\sum_{i \in R} x_i = \sum_{j \in J_1} \sum_{i \in I_j \cap R} x_i + \sum_{j \in J_2} \sum_{i \in I_j \cap R} x_i < \left(\sum_{j \in J_1} \sum_{i \in I_j \cap R} x_i \right) + \frac{1}{\epsilon^2} 2\epsilon^3 r = \left(\sum_{j \in J_1} \sum_{i \in I_j \cap R} x_i \right) + 2\epsilon r.$$

Using iii) now gives

$$\sum_{j \in J_1} \sum_{i \in I_j \cap R} x_i < \sum_{j \in J_1} |I_j \cap R| l_j < \frac{r(1+\epsilon)}{n} \sum_{j \in J_1} |I_j| l_j < \frac{r(1+\epsilon)^2}{n} \sum_{j \in J_1} |I_j| l_{j-1} \le \frac{r(1+\epsilon)^2}{n} \sum_{i=1}^{n} x_i.$$

Consequently,

$$\sum_{i \in R} x_i < \frac{r(1+\epsilon)^2}{n} \sum_{i=1}^{n} x_i + 2\epsilon r = cr(1+\epsilon)^2 + 2\epsilon r < cr(1+3\epsilon) + 2\epsilon r < r(c+5\epsilon) \le r(c+\xi).$$

On the other hand, iii) and iv) together give

$$\begin{aligned} \sum_{i \in R} x_i &\ge \sum_{j \in J_1} \sum_{i \in I_j \cap R} x_i \ge \sum_{j \in J_1} |I_j \cap R| l_{j-1} \ge \frac{\sum_{j \in J_1} |I_j \cap R| l_j}{1+\epsilon} \ge \frac{r(1-\epsilon)}{n(1+\epsilon)} \sum_{j \in J_1} |I_j| l_j \\ &\ge \frac{r(1-\epsilon)}{n(1+\epsilon)} \left(\sum_{i=1}^{n} x_i - \sum_{j \in J_2} |I_j| l_j \right) \ge \frac{r(1-\epsilon)}{n(1+\epsilon)} (cn - \epsilon n) = \frac{r(1-\epsilon)(c-\epsilon)}{(1+\epsilon)} \\ &> r(c-3\epsilon) > r(c-\xi). \end{aligned}$$

We have shown that for all but $e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $R \in [n]^r$, the sum $\sum_{i \in R} x_i = (c \pm \xi) r$ as required. $\square$

The following lemma is used in the proof of Lemma 4.4. Its proof is very similar to that of Lemma 4.3 so we omit it.

Lemma 7.1. *Given $r > r_0$, there exists $n(r)$ such that for $n > n(r)$ the following holds: Suppose that $\sum_{i \in [n]} x_i < n/r^{1/10}$ with $0 \le x_i \le 1$. Then for all but $e^{-\sqrt{r}/2}\binom{n}{r}$ $R \in [n]^r$, we have $\sum_{i \in R} x_i < r^{19/20}$.*

Given a weighted graph G, we write $deg_G(x)$ for the weighted degree of x, so we sum over the weights of all edges incident with x.

Lemma 4.4: Fix $k \ge 2$. For every β, there exists r, n_0 such that if $n > n_0$ then the following holds. Let G be an n vertex weighted graph with total weight at most βn^k, maximum weighted degree at most n^{k-1}, and maximum edge weight at most n^{k-2}. Then all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets of vertices induce a subgraph with weight at most $2\beta n^{k-2} r^2$.

Proof. We may assume by adding weight to edges if needed that G has weight exactly βn^k. We may also divide all weights by n^{k-2} and assume that G has total weight βn^2, maximum weighted degree n, and maximum edge weight 1. Let $V(G) = [n]$, and for each $i \in [n]$, let $x_i = deg_G(i)/n$. Then

$$\sum_{i=1}^{n} x_i = \frac{2\beta n^2}{n} = 2\beta n, \qquad \text{and} \qquad 0 \le x_i = \frac{deg_G(i)}{n} \le 1$$

so by Lemma 4.3, all but $e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets $R \in [n]^r$ have

$$\sum_{i\in R} deg_G(i) = n \sum_{i\in R} x_i = (2\beta \pm \beta/10)nr.$$

From now we restrict only to these r-sets.

Given a vertex i, an r-set R containing i is i-bad if either

i) $deg_R(i) > (1 + \frac{\delta}{10})deg_G(i)\frac{r-1}{n-1}$ and $deg_G(i) > n/r^{1/4}$, or

ii) $deg_R(i) > r^{19/20}$ and $deg_G(i) \le n/r^{1/10}$

If i) holds, then by Lemma 4.3 at most $e^{-\sqrt{r}/2}\binom{n-1}{r-1}$ of the r-sets containing i are i-bad. If ii) holds, then since the edge weight is at most 1, Lemma 7.1 implies that at most $e^{-\sqrt{r}/2}\binom{n-1}{r-1}$ of the r-sets containing i are i-bad. Hence for any vertex i, at most $e^{-\sqrt{r}/2}\binom{n-1}{r-1}$ of the r-sets containing i are i-bad. Therefore the number of r-sets R that contain an i for which R is i-bad is at most $ne^{-\sqrt{r}/2}\binom{n-1}{r-1} < e^{-\sqrt{r}/2}\binom{n}{r}$. For an R that contains no i for which it is i-bad, we have

$$\sum_{i\in R} deg_R(i) < r \cdot r^{19/20} + \sum_{i\in R} 2deg_G(i)\frac{r-1}{n-1} < r^{39/20} + 2\frac{r-1}{n-1}\sum_{i\in R} deg_G(i) < 2\beta r^2.$$

Altogether we have considered all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the r-sets. The weight of one of the sets R we have considered is $\sum_{i\in R} deg_R(i)/2 < 2\beta r^2$ as needed. $\square$

7.2 Proofs from Section 6.2

Recall that an *equipartition* P of a set S is a partition of S into sets of equal size. The equipartition P *contains* an edge e of S if e has nonempty intersection with each of the parts of P.

Let $p(x) = \binom{3x}{x}\binom{2x}{x}$. Note that $p(t)/6$ is the number of equipartitions of a $3t$ element set into three subsets. Then for $a < b$, an easy computation using $\binom{n}{k} = n!/[k!(n-k)!]$ yields

$$\frac{p(a)p(b-a)}{p(b)} = \frac{\binom{b}{a}^3}{\binom{3b}{3a}}. \tag{32}$$

The following Lemma deals with 3-graphs, but can easily be extended to l-graphs.

Lemma 7.2. *Let $\alpha, \widetilde{\delta} > 0$, and let S be a set of size $3s$ with density at least $\alpha + \widetilde{\delta}$. Then the number of equipartitions $P = S_1 \cup S_2 \cup S_3$ of S containing at least $(\alpha + \widetilde{\delta}/2)s^3$ triples of S is at least*

$$\frac{p(s)}{6}\left(1 - O\left(\frac{1}{\widetilde{\delta}^2 s}\right)\right) = \frac{1}{6}\binom{3s}{s}\binom{2s}{s}\left(1 - O\left(\frac{1}{\widetilde{\delta}^2 s}\right)\right). \tag{33}$$

Proof. We consider the space of all equipartitions of S. For each edge e in S, let X_e be the indicator random variable for e belonging to some equipartition P. Then

$$Pr(X_e = 1) = \frac{6p(s-1)}{p(s)}.$$

We now estimate the expected number $E(X) = E(\sum_e X_e)$ of edges contained in an equipartition. By linearity of expectation and (32) we have

$$\begin{aligned} E(X) &\geq (\alpha + \widetilde{\delta})\binom{3s}{3} Pr(X_e = 1) \\ &= 6(\alpha + \widetilde{\delta})\binom{3s}{3}\frac{p(s-1)p(1)}{p(s)p(1)} = 6(\alpha + \widetilde{\delta})\binom{3s}{3}\frac{\binom{s}{1}^3}{6\binom{3s}{3}} = (\alpha + \widetilde{\delta})s^3. \end{aligned}$$

To show that most equipartitions contain close to the expected number of edges, we use Chebyschev's inequality which states that for any positive t, we have $Pr(|X - E(X)| > t\sigma) \leq 1/t^2$, where σ is the standard deviation. We must therefore compute the variance σ^2. Now

$$\sigma^2 = E(X^2) - [E(X)]^2 = E((\sum_e X_e)^2) - [E(X)]^2 = E(X) + \sum_{e \neq f} E(X_e X_f) - [E(X)]^2.$$

The term $\sum_{e \neq f} E(X_e X_f)$ can be computed by considering those pairs e, f that are disjoint and those that intersect. Since $|e \cap f| \neq \emptyset$ implies that $|e \cup f| \leq 5$, we have

$$\sum_{\substack{e,f \\ |e \cap f| \neq \emptyset}} E(X_e)E(X_f) < s^5.$$

On the other hand,

$$\sum_{\substack{e,f \\ |e \cap f| = \emptyset}} E(X_e)E(X_f) = \frac{6^2 p(s-2)}{p(s)} \sum_{\substack{e,f \\ |e \cap f| = \emptyset}} 1 = (\alpha + \widetilde{\delta})^2 s^6 + O(s^5) = [E(X)]^2 + O(s^5).$$

We have shown that the standard deviation σ of X is $O(s^{5/2})$. Therefore

$$Pr(X < (\alpha + \widetilde{\delta}/2)s^3) < Pr(|X - E(X)| > (\widetilde{\delta}/2)s^3) < O\left(\frac{1}{\widetilde{\delta}^2 s}\right),$$

and consequently

$$Pr(X \geq (\alpha + \widetilde{\delta}/2)s^3) \geq 1 - \frac{c}{\widetilde{\delta}^2 s}$$

for some absolute constant c. Since the total number of equipartitions of S is $(1/6)\binom{3s}{s}\binom{2s}{s}$, the number of equipartitions which contain at least $(\alpha + \widetilde{\delta}/2)s^3$ triples of S is given by (33) as needed. $\square$

Proof of Theorem 6.1 from Theorem 6.2: Recall that we only consider the unweighted case and $l = 3$. Alternatively, the weights are all 0 or 1.

Proof of (3): Given a δ-regular $\mathcal{H}$ of density α, we show that for all equipartitions $A = A_1 \cup A_2 \cup A_3$ of $V(\mathcal{H})$, the resulting 3-partite 3-graph $\mathcal{H}(A_1, A_2, A_3)$ is 14δ-regular of density α.

For $i = 1, 2, 3$, pick $A_i' \subset A_i$, each of size at least $v/3 > 14\delta(n/3)$. We may assume that each has size exactly $v/3$ since a simple averaging argument implies the result in the case of strict inequality. We will prove that the number of edges in $\mathcal{H}(A_1', A_2', A_3')$ is $(1 \pm 14\delta)\alpha v^3$.

The set $A_1' \cup A_2' \cup A_3'$ contains $(1 \pm \delta)\alpha v^3$ triples since $\mathcal{H}$ is δ-regular and $v > \delta n$. Similarly, for each i, j, the set $A_i' \cup A_j'$ contains $(1 \pm \delta)\alpha\binom{2v}{3}$ triples, and A_i' contains $(1 \pm \delta)\alpha\binom{v}{3}$ triples. Therefore the number of triples in $A_1' \cup A_2' \cup A_3'$ with at least one element in each of A_1', A_2', and A_3' is by inclusion-exclusion

$$(1 \pm \delta)\alpha\binom{3v}{3} - 3(1 \pm \delta)\alpha\binom{2v}{3} + 3(1 \pm \delta)\alpha\binom{v}{3} = (1 \pm 14\delta)\alpha v^3. \qquad \square$$

Proof of (5): We will assume that there exist $\exp\{-r^{1/l}/20\}\binom{n}{r}$ r-sets $\widetilde{X} \in [X]^r$ (here $l = 3$), which do not satisfy property $(\alpha, \widetilde{\delta})$. Our goal is to prove that there exists an equipartition $A = A_1 \cup A_2 \cup A_3$ of X and $\exp\{-r^{1/l}/20\}\binom{n/3}{r/3}^3$ triples of $r/3$-sets $\widetilde{A}_i \in [A_i]^{r/3}, i \in [3]$, which do not satisfy property $(\alpha, \widetilde{\delta}/2)$.

Let $\mathcal{R}$ be the family of r-sets $\widetilde{X} \in [X]^r$, not satisfying property $(\alpha, \widetilde{\delta})$. This means that for each $\widetilde{X} \in \mathcal{R}$, there is a $Y = Y_{\widetilde{X}} \subset X$, $|\widetilde{Y}| = \widetilde{\delta}r$ that has density outside $\alpha \pm \widetilde{\delta}$. Let

$$\mathcal{B} = \{(\widetilde{X}, Y_{\widetilde{X}}) : \widetilde{X} \in \mathcal{R}\}$$

and

$$\mathcal{B}' = \{(\widetilde{X}, Y_{\widetilde{X}}) \in \mathcal{B} : d(Y_{\widetilde{X}}) > \alpha + \widetilde{\delta}\}.$$

We may assume without loss of generality that $|\mathcal{B}'| \geq |\mathcal{B}|/2 \geq \exp\{-r^{1/l}/20\}\binom{n}{r}/2$.

For an equipartition $P = A_1 \cup A_2 \cup A_3$ and $R \subset X$, let $P_R = R_1 \cup R_2 \cup R_3$ be the restriction of P to R. We prove that such a P exists that satisfies the following two properties for at least $\exp\{-r^{1/l}/10\}\binom{n/3}{r/3}^3$ elements $(\widetilde{X}, Y) \in \mathcal{B}'$:

1) P_Y is an equipartition of Y and the density $d(Y_1, Y_2, Y_3) > (\alpha + \widetilde{\delta}/2)$.

2) $P_{\widetilde{X}}$ is an equipartition of X.

This will complete the proof.

Set $n' = n/3$, $r' = r/3$ and $s' = |Y|/3 = \widetilde{\delta} r/3$. By Lemma 7.2 the number of equipartitions of Y that contain at least $(\alpha + \widetilde{\delta}/2)(s')^3$ triples is at least $(1/6)p(s')\left(1 - O\left(\frac{1}{\widetilde{\delta}^2 s'}\right)\right)$. The number of ways to extend each of these equipartitions of Y to an equipartition of $\widetilde{X}$ is $p(r' - s')$. The number of ways to extend each of these equipartitions of $\widetilde{X}$ to an equipartition of X is $p(n' - r')$.

Now form a bipartite graph with parts $\mathcal{B}', \mathcal{P}$, where $\mathcal{P}$ consists of all equipartitions of X, and join $(\widetilde{X}, Y) \in \mathcal{B}'$ to $P \in \mathcal{P}$ if 1) and 2) are satisfied. We have just argued that every vertex of $\mathcal{B}'$ has degree at least

$$d = \frac{1}{6} p(s')p(r' - s')p(n' - r') \left(1 - O\left(\frac{1}{\widetilde{\delta}^2 s'}\right)\right) \geq \frac{1}{12} p(s')p(r' - s')p(n' - r'),$$

where the inequality holds since r is sufficiently large. Hence there is a $P \in \mathcal{P}$ with degree at least $|\mathcal{B}'|d/|\mathcal{P}| \geq |\mathcal{B}|d/2p(n')$. Now we lower bound this and use (32) to simplify:

$$\begin{aligned}
\frac{|\mathcal{B}|d}{2p(n')} &\geq \frac{1}{24} \exp\{-r^{1/l}/20\} \binom{n}{r} \frac{p(s')p(r' - s')p(n' - r')}{p(n')} \\
&= \frac{1}{24} \exp\{-r^{1/l}/20\} \binom{3n'}{3r'} \frac{p(s')p(r' - s')}{p(r')} \frac{p(r')p(n' - r')}{p(n')} \\
&= \frac{1}{24} \exp\{-r^{1/l}/20\} \binom{3n'}{3r'} \frac{\binom{r'}{s'}^3}{\binom{3r'}{3s'}} \frac{\binom{n'}{r'}^3}{\binom{3n'}{3r'}} \\
&\geq \exp\{-r^{1/l}/10\} \binom{n'}{r'}^3 = \exp\{-r^{1/l}/10\} \binom{n/3}{r/3}^3.
\end{aligned}$$

Consequently there is an equipartition of X satisfying 1) and 2) for at least $\exp\{-r^{1/l}/10\}\binom{n/3}{r/3}^3$ elements $(\widetilde{X}, Y) \in \mathcal{B}'$. This concludes the proof. $\square$

7.3 Proofs from Section 6.3

Proposition 6.5: Let $\mathcal{H}(A_1, \ldots, A_l)$ be an l-partite l-graph, w be a weight function on $\mathcal{H}$ of density α, and $\gamma_i : A_i \to \{0, 1\}$. Suppose that $\delta > 0$ and $|\gamma_i|$ is an integer for each i. Then the following are equivalent:

1) w is (α, δ)-uniform

2) for all choices of $\gamma_i : A_i \to [0, 1]$ with $|\gamma_i| > \delta n$,

$$|d(\gamma_1, \ldots, \gamma_l) = \alpha \pm \delta.$$

Proof. Clearly 2) implies 1) since we may consider γ_i that take only 0-1 values. We now prove that 1) implies 2). Suppose that w is (α, δ)-uniform and $|d(\gamma_1, \ldots, \gamma_l) - \alpha| > \delta$. Assume by symmetry that $d(\gamma_1, \ldots, \gamma_l) > \alpha + \delta$. We will transform γ_i to β_i such that for each i

i) $|\beta_i| = |\gamma_i|$,

ii) β_i is a 0-1 function,

iii) $d(\beta_1,\ldots,\beta_l) \geq d(\gamma_1,\ldots,\gamma_l) > \alpha+\delta$.

Now letting $A'_i = \{z \in A_i : \beta_i(z) = 1\}$ we have that $|A'_i| = |\beta_i| \geq \delta n$ and

$$d(A'_1,\ldots,A'_l) = d(\beta_1,\ldots,\beta_l) > \alpha+\delta.$$

This contradicts 1).

We now show how to create β_i for which i), ii), iii) hold. For $x \in A_1$, let

$$d(x) = \sum\{w(x,a_2,\ldots a_l)\prod_{i=2}^{l}\gamma_i(a_i) : (x,a_2,\ldots,a_l) \in A_1 \times \ldots A_l\}.$$

Suppose that there are $x, x' \in A_1$ with neither $\gamma_1(x)$ nor $\gamma_1(x')$ being an integer. We will transform γ_1 to γ'_1 such that

i') $|\gamma'_1| = |\gamma_1|$,

ii') at least one of $\gamma'_1(x)$ or $\gamma'_1(x')$ is an integer,

iii') $d(\gamma'_1,\gamma_2,\ldots,\gamma_l) \geq d(\gamma_1,\ldots,\gamma_l)$.

Assume without loss of generality that $d(x) \leq d(x')$. Let $\epsilon = \min\{1-\gamma_1(x'),\gamma_1(x)\}$. Set $\gamma'_1(x) = \gamma_1(x)-\epsilon$, $\gamma'_1(x') = \gamma_1(x')+\epsilon$, and $\gamma'_1(z) = \gamma_1(z)$ for $z \notin \{x,x'\}$. Clearly i') holds and by the choice of ϵ, either $\gamma'_1(x) = 0$ or $\gamma'_1(x') = 1$ so ii') holds. Since $d(x) \leq d(x')$,

$$\gamma'_1(x)d(x)+\gamma'_1(x')d(x') = \gamma_1(x)d(x)+\gamma_1(x')d(x')+\epsilon(d(x')-d(x)) \geq \gamma_1(x)d(x)+\gamma_1(x')d(x').$$

Consequently,

$$d(\gamma'_1,\gamma_2,\ldots,\gamma_l) = \frac{\sum_{z\in A_1}\gamma'_1(z)d(z)}{|\gamma'_1||\gamma_2|\cdots|\gamma_l|} \geq \frac{\sum_{z\in A_1}\gamma_1(z)d(z)}{|\gamma_1||\gamma_2|\cdots|\gamma_l|} = d(\gamma_1,\gamma_2,\ldots,\gamma_l),$$

which proves iii').

We repeat such a transformation to γ'_1 until the resulting function takes on only integer values. Since $|\gamma_1|$ is an integer, the resulting function cannot have only one non integer value, and this process therefore terminates to form β_1. Repeating this for each i yields i), ii), and iii) as required. $\square$

Recall the following set up: Let $A = \{a_1,\ldots,a_n\}$, $B = \{b_1,\ldots,b_n\}$ and $w : A \times B \to [0,1]$. Let $w_{i,j} = w(a_i,b_j)$, and let $\vec{w_i} = (w_{i,1},\ldots,w_{i,n})$. Set $\alpha = (1/n^2)\sum_{i,j} w_{i,j}$. Consider the following two statements:

$UNI(\delta)$: w is (α,δ)-uniform

$DPC(\delta')$: The dot product $\vec{w_i}\cdot\vec{w_{i'}} = (\alpha\pm\delta')^2 n$ for all but $\delta' n^2$ pairs i, i'.

Proposition 6.6: The statements $UNI(\delta)$ and $DPC(\delta')$ are equivalent in the following sense:

1) $\forall\delta, \exists\delta', n_0$, such that $\forall n > n_0, DPC(\delta') \Rightarrow UNI(\delta)$

2) $\forall\delta', \exists\delta, n_0$, such that $\forall n > n_0, UNI(\delta) \Rightarrow DPC(\delta')$

Proof. The proofs of both 1) and 2) are similar to Lemmas 3.1 and 3.2 in [2]. Here we only prove 2), since 1) follows by extending the ideas of [2] in a similar fashion.

Suppose that $UNI(\delta)$ holds and let $A' \subset A$ be the set of vertices a_i for which

$$d(a_i) = \sum_{j=1}^{n} w_{i,j} > (\alpha + \delta)n.$$

Then

$$d(A', B) = \frac{\sum_{a_i \in A'} \sum_{b_j \in B} w_{i,j}}{|A'||B|} = \frac{\sum_{a_i \in A'} d(a_i)}{|A'||B|} > \alpha + \delta.$$

Since w is (α, δ)-uniform we conclude that $|A'| < \delta n$. Arguing similarly for vertices with $d(a_i) < (\alpha - \delta)n$, we obtain that for at least $(1 - 2\delta)n$ vertices $a_i \in A$, we have $d(a_i) = (\alpha \pm \delta)n$. We now show that for each such a_i, the dot product $\vec{w_i} \cdot \vec{w_{i'}} = (\alpha \pm \delta)^2 n$ for all but at most $2\delta n$ of the $a_{i'}$s.

Fixing i, let $A'' \subset A$ be the set of those $a_{i'}$ for which $\vec{w_i} \cdot \vec{w_{i'}} > (\alpha + \delta)^2 n$. Define $\gamma_1(x) = 1$ for $x \in A''$ and 0 for $x \notin A''$, and define $\gamma_2(b_j) = w_{i,j}$ for all $b_j \in B$. Then $|\gamma_1| = |A''|$, and $|\gamma_2| = d(a_i) = (\alpha \pm \delta)n$. Consequently,

$$d(\gamma_1, \gamma_2) = \frac{\sum_{a_{i'} \in A''} \sum_{b_j \in B} w_{i',j} \gamma_1(a_{i'}) \gamma_2(b_j)}{|\gamma_1||\gamma_2|} = \frac{\vec{w_i} \cdot \vec{w_{i'}}}{|\gamma_2|} > \frac{(\alpha + \delta)^2 n}{(\alpha + \delta)n} = \alpha + \delta.$$

If $|\gamma_1| \geq \delta n$, then by Proposition 6.5, this contradicts (α, δ)-uniformity of w. Hence $|A''| = |\gamma_1| < \delta n$. Arguing similarly for those $a_{i'}$ for which $\vec{w_i} \cdot \vec{w_{i'}} < (\alpha - \delta)^2 n$, we conclude that $\vec{w_i} \cdot \vec{w_{i'}} = (\alpha \pm \delta)^2 n$ for all but at most $2\delta n$ of the $a_{i'}$s. Therefore the number of pairs i, i' for which $\vec{w_i} \cdot \vec{w_{i'}} \neq (\alpha \pm \delta)^2 n$ is at most $(2\delta n)n + (n - 2\delta n)2\delta n < 4\delta^2 n$. We conclude that for any δ', there is a $\delta = \delta(\delta', \alpha)$ such that $UNI(\delta) \Rightarrow DPC(\delta')$. □

7.4 Proofs from Section 6.5

Proposition 4.5: Let $G = (A, B)$ be a bipartite graph with n^2/t edges and $|A| = |B| = n$. Let $\mathcal{B} \subset E(G)$ be a set of size at most $\epsilon' n^2/t$. Let $\mathcal{B}' \subset [E(G)]^2$ be a set size at most $\epsilon' n^4/t^2$. Then the following holds: For all ϵ', t, there exist r, n_0, such that if $n > n_0$, then for all but $6e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$,

$$|(\widetilde{A} \times \widetilde{B}) \cap E(G)| = \frac{r^2}{t}(1 \pm 1/10) \tag{34}$$

and in the graph spanned by $\widetilde{A} \times \widetilde{B}$,

1) at most $2\epsilon' r^2/t$ edges lie in $\mathcal{B}$, and

2) at most $2\epsilon' r^4/t^2$ pairs of edges lie in $\mathcal{B}'$.

Proof. We first show that for all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$, (34) holds. For $i \in A$, let $x_i = deg(i)/n$. Then $\sum_{i\in A} x_i = n/t$ so by Lemma 4.3 all but $e^{-\sqrt{r}/2}\binom{n}{r}$ choices $\widetilde{A} \in [A]^r$ have

$$\sum_{i\in\widetilde{A}} deg(i) = n\sum_{i\in\widetilde{A}} x_i = (1/t \pm 1/10)rn.$$

Fix such an $\widetilde{A}$, and for each $j \in B$, define $y_j = deg_{\widetilde{A}}(j)/r$. Then $\sum_{j\in B} y_j = \sum_{i\in\widetilde{A}} deg(i)/r = (1/t \pm 1/10)n$. By Lemma 4.3, all but $e^{-\sqrt{r}/2}\binom{n}{r}$ choices $\widetilde{B} \in [B]^r$ have

$$e(\widetilde{A}, \widetilde{B}) = \sum_{j\in\widetilde{B}} deg_{\widetilde{A}}(j) = r\sum_{j\in\widetilde{B}} y_j = (1/t \pm 1/10)r^2.$$

The total number of exceptions $\widetilde{A}, \widetilde{B}$ is at most $e^{-\sqrt{r}/2}\binom{n}{r}^2 + e^{-\sqrt{r}/2}\binom{n}{r}^2 = 2e^{-\sqrt{r}/2}\binom{n}{r}^2$ as claimed.

Applying this argument to the graph with edge set $\mathcal{B}$, we obtain that 1) holds for all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A} \in [A]^r, \widetilde{B} \in [B]^r$.

We now prove that 2) holds for all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ exceptions. Define a weighted graph G_A with vertex set A (possibly with loops), where the weight $w(uv)$ of edge uv is the number of $\{e, e'\} \in \mathcal{B}'$ with u an endpoint of e and v an endpoint of e'. Then the sum of the weights of all edges in G_A is $|\mathcal{B}'| \leq \epsilon' n^4/t^2$, the maximum weighted degree is at most n^3, and the maximum edge weight is at most n^2. By Lemma 4.4, setting $\beta = \epsilon'/t^2$ and $k = 4$, all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the $\widetilde{A} \in [A]^r$ have

$$\sum_{u,v\in\widetilde{A}} w(uv) < 2\epsilon' \frac{n^4}{t^2}\frac{r^2}{n^2} = 2\epsilon' \frac{n^2 r^2}{t^2}.$$

Now define a weighted graph G_B with vertex set B (possibly with loops), where the weight $w(uv)$ of edge uv is the number of $\{e, e'\} \in \mathcal{B}'$ with $e = au$ and $e' = a'v$ where $a, a' \in A$. Then the sum of the weights of all edges in G_B is $\sum_{u,v\in\widetilde{A}} w(uv) < 2\epsilon' \frac{n^2r^2}{t^2}$, the maximum weighted degree is at most $r^2 n$, and the maximum edge weight is at most r^2. Therefore by Lemma 4.4 (scaling down by a factor of r^2, setting $\beta = 2\epsilon'/t^2$, and $k = 2$), all but $2e^{-\sqrt{r}/2}\binom{n}{r}$ of the $\widetilde{B} \in [B]^r$ have

$$\sum_{u,v\in\widetilde{B}} w(uv) < 2\epsilon' \frac{n^2 r^2}{t^2}\frac{r^2}{n^2} = 2\epsilon' \frac{r^4}{t^2}.$$

Therefore the number of elements of $\mathcal{B}'$ in the graph spanned by $\widetilde{A} \times \widetilde{B}$ is at most $2\epsilon r^4/t^2$, and this holds for all but $2e^{-\sqrt{r}/2}\binom{n}{r}^2$ choices $\widetilde{A}, \widetilde{B}$.

We have argued that altogether the number of $\widetilde{A}, \widetilde{B}$ not satisfying (34), 1) or 2), is at most $6e^{-\sqrt{r}/2}\binom{n}{r}^2$, thereby completing the proof. $\square$

8 Acknowledgments

Part of this work was done while the first author was a postdoctoral researcher in the Theory Group at Microsoft Research. The authors thank Mathias Schacht for useful comments, and Jozef Skokan for his careful reading of an earlier draft of this paper.

References

[1] N. Alon, M. Krivelevich, Testing k-colorability. SIAM J. Discrete Math. 15 (2002), no. 2, 211–227.

[2] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, R. Yuster, The algorithmic aspects of the regularity lemma. J. Algorithms 16 (1994), no. 1, 80–109.

[3] N. Alon, A. Shapira, Testing satisfiability, Proc. of the 13^{th} Annual ACM-SIAM SODA, ACM Press (2002), 645-654. Also: J. Algorithms, in press.

[4] N. Alon, W. F. de la Vega, R. Kannan and M. Karpinski, Random Sampling and Approximation of MAX-CSP Problems, Proc. of the 34 ACM STOC, ACM Press (2002) 232–239.

[5] B. Bollobás, P. Erdős, M. Simonovits, E. Szemerédi, Extremal graphs without large forbidden subgraphs. Advances in graph theory (Cambridge Combinatorial Conf., Trinity Coll., Cambridge, 1977). Ann. Discrete Math. 3 (1978), 29–41.

[6] A. Czygrinow, B. Nagle, On the hereditary nature of weakly-regular triple systems, in preparation.

[7] R. A. Duke, H. Lefmann, V. Rödl, A fast approximation algorithm for computing the frequencies of subgraphs in a given graph. SIAM J. Comput. 24 (1995), no. 3, 598–620.

[8] R. A. Duke, V. Rödl, On graphs with small subgraphs of large chromatic number. Graphs Combin. 1 (1985), no. 1, 91–96

[9] P. Erdős, M. Simonovits, A limit theorem in graph theory. Studia Sci. Math. Hungar 1 1966 51–57.

[10] P. Erdős, V. T. Sós, On Ramsey-Turán type theorems for hypergraphs. Combinatorica 2 (1982), no. 3, 289–295.

[11] P. Erdős, A. H. Stone, On the structure of linear graphs. Bull. Amer. Math. Soc. 52, (1946). 1087–1091.

[12] P. Frankl, Z. Füredi, A new generalization of the Erdős-Ko-Rado theorem. Combinatorica 3 (1983), no. 3-4, 341–349.

[13] P. Frankl, V. Rödl, Some Ramsey-Turán type results for hypergraphs. Combinatorica 8 (1988), no. 4, 323–332.

[14] P. Frankl, V. Rödl, Extremal problems on set systems, Rand. Struct. Alg. 20 (2002), pp 131-164.

[15] O. Goldreich, S. Goldwasser, D. Ron, Property testing and its connection to learning and approximation. J. ACM 45 (1998), no. 4, 653–750.

[16] W. T. Gowers, Lower bounds of tower type for Szemeré di's uniformity lemma. Geom. Funct. Anal. 7 (1997), no. 2, 322–337.

[17] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley (2000)

[18] J. Komlós, M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, in: Combinatorics, Paul Erdos is Eighty, (D. Miklos et. al, eds.), Bolyai Society Mathematical Studies, Vol.2, 1996, pp. 295–352.

[19] D. Mubayi, V. Rödl, Supersaturation for Ramsey-Turán problems, submitted.

[20] A. F. Sidorenko, On Ramsey-Turán numbers for 3-graphs. J. Graph Theory 16 (1992), no. 1, 73–78.

[21] M. Simonovits, Extremal graph problems, degenerate extremal problems, and supersaturated graphs. Progress in graph theory (Waterloo, Ont., 1982), 419–437, Academic Press, Toronto, ON, 1984.

[22] M. Simonovits, V. T. Sós, Ramsey-Turán theory, Combinatorics, graph theory, algorithms and applications. Discrete Math. 229 (2001), no. 1-3, 293–340.