Ryan R. Martina, Jozef Skokan

Asymptotic multipartite version of the Alon–Yuster theorem

Article (Accepted version) (Refereed)

Original citation:
Martin, Ryan R. and Skokan, Jozef (2017) *Asymptotic multipartite version of the Alon–Yuster theorem.* Journal of Combinatorial Theory, Series B. ISSN 0095-8956

DOI: 10.1016/j.jctb.2017.05.004

Asymptotic multipartite version of the Alon-Yuster theorem

Ryan R. Martin[a,1], Jozef Skokan[b]

[a] *Department of Mathematics, Iowa State University, Ames, Iowa 50011*
[b] *Department of Mathematics, London School of Economics, London, WC2A 2AE, UK and Department of Mathematics, University of Illinois, 1409 W. Green Street, Urbana, IL 61801*

Abstract

In this paper, we prove the asymptotic multipartite version of the Alon-Yuster theorem, which is a generalization of the Hajnal-Szemerédi theorem: If $k \geq 3$ is an integer, H is a k-colorable graph and $\gamma > 0$ is fixed, then, for every sufficiently large n, where $|V(H)|$ divides n, and for every balanced k-partite graph G on kn vertices with each of its corresponding $\binom{k}{2}$ bipartite subgraphs having minimum degree at least $(k-1)n/k + \gamma n$, G has a subgraph consisting of $kn/|V(H)|$ vertex-disjoint copies of H.

The proof uses the Regularity method together with linear programming.

Keywords: tiling, Hajnal-Szemerédi, Alon-Yuster, multipartite, regularity, linear programming
2010 AMS Subject Classification: 05C35, 05C70

1. Introduction

1.1. Motivation

One of the celebrated results of extremal graph theory is the theorem of Hajnal and Szemerédi on tiling simple graphs with vertex-disjoint copies of a given complete graph K_k on k vertices. Let G be a simple graph with vertex-set $V(G)$ and edge-set $E(G)$. We denote by $\deg_G(v)$, or simply $\deg(v)$, the degree of a vertex $v \in V(G)$ and we denote by $\delta(G)$ the minimum degree of the graph G. For a graph H such that $|V(H)|$ divides $|V(G)|$, we say that G has a *perfect H-tiling* (also a *perfect H-factor* or *perfect H-packing*) if there is a subgraph of G that consists of $|V(G)|/|V(H)|$ vertex-disjoint copies of H.

The theorem of Hajnal and Szemerédi can be then stated in the following way:

Email addresses: `rymartin@iastate.edu` (Ryan R. Martin), `j.skokan@lse.ac.uk` (Jozef Skokan)

[1] This author's research partially supported by NSF grant DMS-0901008, NSA grant H98230-13-1-0226 and by an Iowa State University Faculty Professional Development grant.

Theorem 1 (Hajnal, Szemerédi [10]). *If G is a graph on n vertices, $k \mid n$, and $\delta(G) \geq (k-1)n/k$, then G has a perfect K_k-tiling.*

The case of $k = 3$ was first proven by Corrádi and Hajnal [5] before the general case. The original proof in [10] was relatively long and intricate. A shorter proof was provided later by Kierstead and Kostochka [16]. Kierstead, Kostochka, Mydlarz and Szemerédi [17] improved this proof and gave a fast algorithm for finding K_k-tilings in n-vertex graphs with minimum degree at least $(k-1)n/k$.

The question of finding a minimum-degree condition for the existence of a perfect H-tiling in the case when H is not a clique and n obeys some divisibility conditions was first considered by Alon and Yuster [1]:

Theorem 2 (Alon, Yuster [1]). *Let H be an h-vertex graph with chromatic number k and let $\gamma > 0$. If n is large enough, $h \mid n$ and G is a graph on n vertices with $\delta(G) \geq (k-1)n/k + \gamma n$, then G has a perfect H-tiling.*

Komlós, Sárközy and Szemerédi [20] removed the γn term from the minimum degree condition and replaced it with a constant that depends only on H.

Kühn and Osthus [23] determined that $(1 - 1/\chi^*(H))\,n + C$ was the necessary minimum degree to guarantee an H-tiling in an n-vertex graph for n sufficiently large, and they also showed that this was best possible up to the additive constant. The constant $C = C(H)$ depends only on H and χ^* is an invariant related to the so-called critical chromatic number of H, which was introduced by Komlós [18].

1.2. Background

In this paper, we consider the multipartite variant of Theorem 2. Before we can state the problem, we need a few definitions.

Given a graph G, the *blow-up of G by m*, denoted by $G(m)$, is the graph obtained by replacing each vertex $v \in V(G)$ with a set U_v of m vertices and replacing every edge $\{v_1, v_2\} \in E(G)$ with the complete bipartite graph $K_{m,m}$ on vertex sets U_{v_1} and U_{v_2}.

A k-partite graph $G = (V_1, \ldots, V_k; E)$ is *balanced* if $|V_1| = \cdots = |V_k|$. The *natural bipartite subgraphs* of G are those induced by the pairs (V_i, V_j), and which we denote by $G[V_i, V_j]$. For a k-partite graph $G = (V_1, \ldots, V_k; E)$, we define the *minimum bipartite degree*, $\hat{\delta}_k(G)$, to be the smallest minimum degree among all of the natural bipartite subgraphs of G, that is,

$$\hat{\delta}_k(G) = \min_{1 \leq i < j \leq k} \delta(G[V_i, V_j]).$$

Now we can state the conjecture that inspired this work, a slightly weaker version of which appeared in [6].

Conjecture 3. *Fix an integer $k \geq 3$. If G is a balanced k-partite graph on kn vertices such that $\hat{\delta}_k(G) \geq (k-1)n/k$, then either G has a perfect K_k-tiling or both k and n/k are odd integers and G is isomorphic to the fixed graph $\Gamma_{k,n}$.*

The exceptional graphs $\Gamma_{k,n}$, where n is an integer divisible by k, are due to Catlin [3] who called them "type 2 graphs". The graph $\Gamma_{k,k}$ has vertex set $\{h_{ij} : i, j \in \{1, \ldots, k\}\}$ and h_{ij} is adjacent to $h_{i'j'}$ if $i \neq i'$ and either $j = j' \in \{k-1, k\}$ or $j \neq j'$ and at least one of j, j' is in $\{1, \ldots, k-2\}$. For n divisible by k, the graph $\Gamma_{k,n}$ is the blow-up graph $\Gamma_{k,k}(n/k)$.

We notice that if G satisfies the minimum bipartite degree condition in Conjecture 3, then its minimum degree $\delta(G)$ can still be as small as $(k-1)\left(\frac{k-1}{k}\right)n = \left(\frac{k-1}{k}\right)^2(kn)$, which is not enough to apply Theorem 1 directly.

The case of $k = 2$ of Conjecture 3 is an immediate corollary of the classical matching theorem due to König [22] and Hall [11]. Fischer [8] observed that if G is a balanced k-partite graph on kn vertices with $\hat{\delta}_k(G) \geq (1 - 1/2(k-1))\,n$, then G has a perfect K_k-tiling.

Some partial results were obtained, for $k = 3$, by Johansson [13] and, for $k = 3, 4$, by Fischer [8]. The case of $k = 3$ was settled for n sufficiently large by Magyar and the first author [25], and the case of $k = 4$ was settled for n sufficiently large by Szemerédi and the first author [26]. The results in [25, 26] each have as a key lemma a variation of the results of Fischer. However, it seems that such techniques are impossible for $k \geq 5$. An interesting result toward proving Conjecture 3 for general k is due to Csaba and Mydlarz [6] who proved that if G is a balanced k-partite graph on kn vertices, $\hat{\delta}_k(G) \geq \frac{q_k}{q_k+1}n$ and n is large enough, then G has a perfect K_k-tiling. Here, $q_k := k - \frac{3}{2} + \frac{1}{2}\sum_{i=1}^{k}\frac{1}{i} = k + O(\log k)$.

Recently, Keevash and Mycroft [14] proved that, for any $\gamma > 0$, if n is large enough, then $\hat{\delta}_k(G) \geq (k-1)n/k + \gamma n$ guarantees a perfect K_k-tiling in a balanced k-partite graph G on kn vertices. Their result is a consequence of a more general theorem on hypergraph matching, the proof of which uses the hypergraph regularity method and a hypergraph version of the Blow-up Lemma. Very shortly thereafter, Lo and Markström [24] proved the same result using methods from linear programming and the so-called "absorbing method". This effort culminated in [15], in which Keevash and Mycroft proved Conjecture 3.

In this paper, we are interested in more general problem of tiling k-partite balanced graphs by a fixed k-colorable graph H. More precisely, if H is a k-colorable graph and n obeys certain natural divisibility conditions, we look for a condition on $\hat{\delta}_k(G)$ to ensure that every balanced k-partite graph G on kn vertices satisfying this condition has a perfect H-tiling.

Zhao [31] found that the minimum degree required to perfectly tile a balanced bipartite graph on $2n$ vertices with copies of $K_{h,h}$ (h divides n) is $n/2 + C(h)$, where $C(h)$ differs sharply as to whether n/h is odd or even. Zhao and the first author [27, 28] showed similar results for tiling with $K_{h,h,h}$. Hladký and Schacht [12] and then Czygrinow and DeBiasio [7] improved the results of [31] by finding the minimum degree for copies of $K_{s,t}$, where $s + t$ divides n. Bush and Zhao [2] proved a Kühn-Osthus-type result by finding the asymptotically best-possible minimum degree condition in a balanced bipartite graph on $2n$ vertices in order to ensure its perfect H-tiling, for any bipartite H. All results are for n sufficiently large.

1.3. Main Result

We prove a multipartite version of the Alon-Yuster theorem (Theorem 2). Let K_h^k denote a k-partite graph with h vertices in each partite set. For example, the complete bipartite graph $K_{h,h}$ would be denoted K_h^2. Since the partite sets can be rotated, it is easy to see that any k-chromatic graph H of order h perfectly tiles the graph K_h^k. Hence, the following theorem gives a sufficient condition for a perfect H-tiling.

Theorem 4. *Fix an integer $k \geq 2$, an integer $h \geq 1$ and $\gamma \in (0,1)$. If n is sufficiently large, divisible by h, and G is a balanced k-partite graph on kn vertices with $\hat{\delta}_k(G) \geq \left(\frac{k-1}{k} + \gamma\right) n$, then G has a perfect K_h^k-tiling.*

Our proof relies on the regularity method for graphs and linear programming and it differs from approaches in [14, 24].

1.4. Structure of the Paper

In Section 2, we prove a fractional version of the multipartite Hajnal-Szemerédi theorem. This is the main tool in proving Theorem 4. Section 3 is the main proof and Section 4 gives the proofs of the supporting lemmas. We finish with Section 5, which has some concluding remarks.

2. Linear Programming

In this section, we shall prove a fractional version of Conjecture 3.

Definition 5. *For any graph G, let $\mathcal{T}_k(G)$ denote the set of all copies of K_k in G. The* ***fractional K_k-tiling number*** *$\tau_k^*(G)$ is defined as:*

$$\begin{aligned} \tau_k^*(G) = \max \quad & \sum_{T \in \mathcal{T}_k(G)} w(T) & & \quad (1)\\ \text{s.t.} \quad & \sum_{\substack{T \in \mathcal{T}_k(G) \\ V(T) \ni v}} w(T) \leq 1, & \forall v \in V(G), & \\ & w(T) \geq 0, & \forall T \in \mathcal{T}_k(G). & \end{aligned}$$

From the Duality Theorem of linear programming (see [29, Section 7.4]), we obtain that

$$\begin{aligned} \tau_k^*(G) = \min \quad & \sum_{v \in V(G)} x(v) & & \quad (2)\\ \text{s.t.} \quad & \sum_{v \in V(T)} x(v) \geq 1, & \forall T \in \mathcal{T}_k(G), & \\ & x(v) \geq 0, & \forall v \in V(G). & \end{aligned}$$

Let w^* be a function that achieves an optimal solution to (1). If there exists a vertex $v \in V(G)$ such that $\sum_{T\in\mathcal{T}_k(G),V(T)\ni v} w^*(T) < 1$, then we call v a *slack vertex* or just say that v *is slack*. Similarly, if x^* is a function that achieves an optimal solution to (2) and there exists a $T \in \mathcal{T}_k(G)$ such that $\sum_{v\in V(T)} x^*(v) > 1$, then we say that T *is slack*.

Remark 6. *Consider an optimal solution to (1), call it w^*. We may assume that $w^*(T)$ is rational for each $T \in \mathcal{T}_k(G)$. To see this, observe that the set of feasible solutions is a polyhedron for which each vertex is the solution to a system of equations that result from setting a subset of the constraints of the program (1) to equality. (For more details, see [4, Theorem 18.1].) Since the objective function achieves its maximum at such a vertex (See [9, Section 3.2].) we may choose an optimal solution $w^*(T)$ with rational entries.*

Now we can state and prove a fractional version of the multipartite Hajnal-Szemerédi Theorem.

Theorem 7. *Let $k \geq 2$. If G is a balanced k-partite graph on kn vertices such that $\hat{\delta}_k(G) \geq (k-1)n/k$, then $\tau_k^*(G) = n$.*

Proof. Setting $x(v) = 1/k$ for all vertices $v \in V(G)$ gives a feasible solution x to (2), and so $\tau_k^*(G) \leq \sum_{v\in V(G)} x(v) = n$. We establish that $\tau_k^*(G) \geq n$ by induction on k.

Base Case. $k = 2$**.** This case follows from the fact that Hall's matching condition implies that a balanced bipartite graph on $2n$ vertices with minimum degree at least $n/2$ has a perfect matching. Setting $w(e)$ equal to 1 if edge e is in the matching and equal to 0 otherwise, gives a feasible solution to (1), thus establishing that $\tau_2^*(G) \geq n$.

Induction step. $k \geq 3$**.** Now we assume $k \geq 3$ and suppose, for any balanced $(k-1)$-partite graph G' on a total of $(k-1)n'$ vertices with $\hat{\delta}_{k-1}(G') \geq \frac{k-2}{k-1}n'$, that $\tau_{k-1}^*(G') \geq n'$.

Let w^* be an optimal solution to (1). Let x^* be an optimal solution corresponding to (2) such that $x^*(z) = 0$ whenever vertex z is slack. This is guaranteed by the Complementary Slackness Theorem [29, Section 7.9]. Denote by $\mathcal{S}$ the set of slack vertices, and, for $i \in [k]$, set $\mathcal{S}_i = \mathcal{S} \cap V_i$. If some $\mathcal{S}_i = \emptyset$, then V_i having no slack vertices gives that $\sum_{T\ni v} w^*(T) = 1$ for each $v \in V_i$. Since each $T \in \mathcal{T}_k(G)$ has exactly one vertex in $V_k(G)$, then $\tau_k^*(G) = n$. Hence, we may assume that every $\mathcal{S}_i$ is non-empty.

Denote $[k] := \{1, \ldots, k\}$. For every $i \in [k]$, fix some $z_i \in \mathcal{S}_i$, choose exactly $n' := \left\lceil \frac{k-1}{k} n \right\rceil$ neighbors of z_i in each V_j, $j \in [k] - \{i\}$, and denote by G_i the subgraph of G induced on these $(k-1)n'$ neighbors.

Observe that the set of weights $\{x^*(v) : v \in V(G_i)\}$ must be a feasible solution to the minimization problem (2) defined by the $(k-1)$-partite graph G_i. This is because every copy of K_{k-1} in G_i extends to a copy of K_k in G containing the vertex z_i and the sum of the weights of the vertices on that K_{k-1} must be at least 1 because $x^*(z_i) = 0$. Hence, we have that $\sum\limits_{v \in V(G_i)} x^*(v) \geq \tau^*_{k-1}(G_i)$.

Each vertex of G_i has at most $n - n'$ neighbors outside of $V(G_i)$ in each of its classes. Thus,

$$\hat{\delta}_{k-1}(G_i) \geq n' - (n - n') = \frac{k-2}{k-1}n' + \left(\frac{k}{k-1}n' - n\right) \geq \frac{k-2}{k-1}n'.$$

So, for every i, we may apply the inductive hypothesis to G_i and conclude that $\tau^*_{k-1}(G_i) = n'$.

Combining the previous two observations with the fact that each vertex v is in at most $k - 1$ of the subgraphs G_i, we get

$$(k-1)\tau^*_k(G) = (k-1) \sum_{v \in V(G)} x^*(v) \geq \sum_{i=1}^{k} \sum_{v \in V(G_i)} x^*(v) \geq \sum_{i=1}^{k} \tau^*_{k-1}(G_i) = kn'.$$

So, $\tau^*_k(G) \geq \frac{k}{k-1}n' = \frac{k}{k-1}\left\lceil \frac{k-1}{k}n \right\rceil \geq n$. This concludes the proof of Theorem 7. $\square$

3. Proof of Theorem 4

First, we will have a sequence of constants and the notation $a \gg b$ means that the constant b is sufficiently small compared to a. We fix $k \geq 2$ and $h \geq 1$ and let

$$\min\{k^{-1}, h^{-1}, \gamma\} \gg d \gg \varepsilon' \gg \zeta \gg n^{-1}, \tag{3}$$

We have an additional parameter ε and specify that $\varepsilon = (\varepsilon')^5/16$.

3.1. Applying the Regularity Lemma

We are going to use a variant of Szemerédi's Regularity Lemma. Before we can state it, we need a few basic definitions. If G is a graph with $S \subset V(G)$ and $x \in V(G)$, then $\deg_G(x, S)$ (or $\deg(x, S)$ if G is understood) denotes $|N(x) \cap S|$.

For disjoint vertex sets A and B in some graph, let $e(A, B)$ denote the number of edges with one endpoint in A and the other in B. Further, let the *density* of the pair (A, B) be $d(A, B) = e(A, B)/|A||B|$. The pair (A, B) is *ε-regular* if $X \subseteq A$, $Y \subseteq B$, $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ imply $|d(X, Y) - d(A, B)| \leq \varepsilon$.

We say that a pair (A, B) is (ε, δ)-*super-regular* if it is ε-regular and $\deg(a, B) \geq \delta|B|$ for all $a \in A$ and $\deg(b, A) \geq \delta|A|$ for all $b \in B$.

The degree form of Szemerédi's Regularity Lemma (see, for instance, [21]) is sufficient here, modified for the multipartite setting.

Theorem 8. *For every integer $k \geq 2$ and every $\varepsilon > 0$, there is an $M = M(k, \varepsilon)$ such that if $G = (V_1, \ldots, V_k; E)$ is a balanced k-partite graph on kn vertices and $d \in [0, 1]$ is any real number, then there is an integer ℓ, a subgraph $G' = (V_1, \ldots, V_k; E')$ and, for $i = 1, \ldots, k$, partitions of V_i into clusters $V_i^{(0)}, V_i^{(1)}, \ldots, V_i^{(\ell)}$ with the following properties:*

(P1) $\lceil \varepsilon^{-1} \rceil \leq \ell \leq M$,

(P2) $|V_i^{(0)}| \leq \varepsilon n$ *for* $i \in [\ell]$,

(P3) $|V_i^{(j)}| = L \leq \varepsilon n$ *for* $i \in [k]$ *and* $j \in [\ell]$,

(P4) $\deg_{G'}(v, V_{i'}) > \deg_G(v, V_{i'}) - (d + \varepsilon)n$ *for all* $v \in V_i$, $i \neq i'$, *and*

(P5) *all pairs* $(V_i^{(j)}, V_{i'}^{(j')})$, $i, i' \in [k]$, $i \neq i'$, $j, j' \in [\ell]$, *are* ε*-regular in* G', *each with density either* 0 *or exceeding* d.

We omit the proof of Theorem 8, which follows from the proof given in [30].

Given a balanced k-partite graph G on kn vertices with $\hat{\delta}_k(G) \geq \left(\frac{k-1}{k} + \gamma\right) n$, and given d and ε, we construct the *reduced graph* G_r on $k\ell$ vertices corresponding to the clusters $V_i^{(j)}$, $1 \leq i \leq k$, $1 \leq j \leq \ell$, obtained from Theorem 8. Each edge of G_r corresponds to an ε-regular pair with density at least d in G'. Observe that G_r is k-partite and balanced. Lemma 9 shows that G_r has a similar minimum-degree condition to that of G.

Lemma 9. *Let G be a balanced k-partite graph G on kn vertices with $\hat{\delta}_k(G) \geq \left(\frac{k-1}{k} + \gamma\right) n$. Then, for the reduced graph G_r defined as above, we have $\hat{\delta}_k(G_r) \geq \left(\frac{k-1}{k} + \gamma - ((k+2)\varepsilon + d)\right) \ell$. Furthermore, if $(k+2)\varepsilon + d \leq \gamma/2$, then*

$$\hat{\delta}_k(G_r) \geq \left(\frac{k-1}{k} + \gamma/2\right) \ell.$$

The proof of Lemma 9 is immediate (see [6]).

3.2. Partitioning the clusters

We first apply the fractional version of the k-partite Hajnal-Szemerédi Theorem (Theorem 7) to G_r and obtain that the value of $\tau_k^*(G_r)$ is equal to ℓ. Consider a corresponding optimal solution w^* to the linear program (1) as it is applied to G_r. By Remark 6, we may fix a corresponding solution $w^*(T)$ that is rational for every $T \in \mathcal{T}_k(G_r)$. We will call this w^* a *rational-entry solution for* G_r and denote by $D(G_r)$ the common denominator of all of the entries of w^*.

Since the linear program (1) depends only on G_r and the number of such reduced graphs is only dependent on $M(k, \varepsilon)$, the number of possible linear programs is only dependent only on k and ε. For each possible linear program we fix one rational-entry solution.

Therefore, the least common multiple of all of the common denominators $D(G_r)$ for these reduced graphs is a function only of k and ε. Call it $D = D(k,\varepsilon)$. In sum, D has the property that for every reduced graph G_r, there is a rational-entry solution w^* of the linear program (1) such that $D \cdot w^*(T)$ is an integer for every $T \in \mathcal{T}_k(G_r)$.

The next step is to partition, uniformly at random, each set $V_i^{(j)}$ into D parts of size $h\lfloor L/(Dh)\rfloor$ as well as a single (possibly empty) set of size $L - Dh\lfloor L/(Dh)\rfloor < Dh$. The vertices of the latter set of less than Dh vertices will be added to the corresponding leftover set, $V_i^{(0)}$. The resulting leftover set $\tilde{V}_i^{(0)}$ has size less than $\varepsilon n + Dh\ell < 2\varepsilon n$.

Thus, for $L' = h\lfloor L/(Dh)\rfloor$, we obtain $k(D\ell)$ clusters $\tilde{V}_i^{(j)}$, $i \in [k]$, $j \in [D\ell]$, such that each of them has size exactly L'. This new partition has the following properties:

(P1$'$) $\ell' = D\ell$,

(P2$'$) $|\tilde{V}_i^{(0)}| \le 2\varepsilon n$ for $i \in [k]$,

(P3$'$) $|\tilde{V}_i^{(j)}| = L' = h\lfloor L/(Dh)\rfloor$ for $i \in [k]$ and $j \in [\ell']$,

(P4$'$) $\deg_{G'}(v, V_{i'}) > \deg_G(v, V_{i'}) - (d+\varepsilon)n$ for all $i, i' \in [k]$, $i \ne i'$, $v \in V_i$ and

Now we prove that a property similar to property (P5) holds.

(P5$'$) all pairs $(\tilde{V}_i^{(j)}, \tilde{V}_{i'}^{(j')})$, $i, i' \in [k]$, $i \ne i'$, $j, j' \in [\ell']$ are ε'-regular in G', each with density either 0 or exceeding $d' := d - \varepsilon$.

Recall from (3) that $\varepsilon = (\varepsilon')^5/16$ and, consequently, $\varepsilon' = (16\varepsilon)^{1/5}$.

The upcoming Lemma 10, a slight modification of a similar lemma by Csaba and Mydlarz [6, Lemma 14], implies that, in fact, (P5$'$) holds with probability going to 1 as $n \to \infty$. The proof follows easily from theirs and so we omit it.

Lemma 10 (Random Slicing Lemma). *Let $0 < d < 1$, $0 < \varepsilon < \min\{d/4, (1-d)/4, 1/9\}$ and D be a positive integer. There exists a $C = C(\varepsilon, D) > 0$ such that the following holds: Let (X,Y) be an ε-regular pair of density d with $|X| = |Y| = DL'$. If X and Y are randomly partitioned into sets $A_1, \dots, A_D$, and $B_1, \dots, B_D$, respectively, each of size L', then, with probability at least $1 - \exp\{-C \cdot DL'\}$, all pairs (A_i, B_j) are $(16\varepsilon)^{1/5}$-regular with density at least $d - \varepsilon$.*

Using Lemma 10, the property (P5$'$) holds with probability at least $1 - \binom{k}{2}\ell^2 \exp\{-CDL'\} = 1 - \binom{k}{2}\ell^2 \exp\{-O(L)\}$. Since $\ell \le M = M(k,\varepsilon)$ and $L \ge n(1-\varepsilon)/M$, then for every sufficiently large n, a partition satisfying (P1$'$)-(P5$'$) exists (with high probability). We fix a partition that satisfies (P1$'$)-(P5$'$). The sets $\tilde{V}_j^{(j)}$ are called *sub-clusters*.

To understand this new partition, we define its reduced graph G'_r with vertex set $\bigcup_{i=1}^k \{u_i^{(1)}, \dots, u_i^{(\ell')}\}$. The vertex $u_i^{(j)}$ corresponds to the cluster $\tilde{V}_i^{(j)}$. The vertices $u_i^{(j)}$ and $u_{i'}^{(j')}$ are adjacent in G'_r if and only if the pair $(\tilde{V}_i^{(j)}, \tilde{V}_{i'}^{(j')})$ is ε'-regular with density at least d'. The graph G'_r clearly has the following properties:

- G'_r is k-partite and balanced on $k\ell'$ vertices. We denote its partite sets $U'_i = \{u_i^{(1)}, \ldots, u_i^{(\ell')}\}$, $i \in [k]$.
- $\hat{\delta}_k(G'_r) \geq \left(\frac{k-1}{k} + \gamma/2\right)\ell'$.

The usefulness of G'_r is that it has a K_k-tiling, which is derived from the fractional K_k-tiling of G_r:

Fact 11. *The reduced graph G'_r has a perfect K_k-tiling.*

Proof of Fact 11. Observe first that, by (P5) and (P5′), G'_r is simply the blow-up graph $G_r(D)$. Let w^* be the previously-chosen rational-valued solution to the linear program (1) as applied to G_r.

Consider some $T \in \mathcal{T}_k(G_r)$ with vertices $\{v_1, \ldots, v_k\}$. Observe that, by the definition of D, $Dw^*(T)$ is an integer. Then, we take $Dw^*(T)$ of the vertices from U_{v_1}, $Dw^*(T)$ of the vertices from U_{v_2} and so on until taking $Dw^*(T)$ of the vertices from U_{v_k}. This selection produces $Dw^*(T)$ vertex-disjoint copies of K_k in G'_r.

By the constraint inequalities in (1), the total number of vertices used from U_v is

$$\sum_{T \in \mathcal{T}_k(G_r), V(T) \ni v} Dw^*(T) \leq D = |U_v|,$$

hence the process never fails. The total number of vertex-disjoint K_k-s that are created in this way is $\sum_{T \in \mathcal{T}_k(G_r)} Dw^*(T) = D\ell = \ell'$. This uses each of the $k\ell'$ vertices of G'_r. $\square$

Since G'_r has a perfect tiling, we may re-index its vertices so that vertices of G'_r (the vertices of G'_r correspond to the sub-clusters of G) with the same upper-index are in the same copy of the tiling from Fact 11. More precisely,

- for $j = 1, \ldots, \ell'$, the k-tuple $(u_1^{(j)}, \ldots, u_k^{(j)})$ forms a K_k in G'_r. We refer to the k-tuples $(\tilde{V}_1^{(j)}, \ldots, \tilde{V}_k^{(j)})$ as *columns.*[2]

3.3. Making the cliques super-regular

In preparation for using the Blow-up Lemma (Lemma 18 below), we need to make each k-tuple $(\tilde{V}_1^{(j)}, \ldots, \tilde{V}_k^{(j)})$, $j \in [\ell']$, pairwise super-regular by placing some vertices from the corresponding sub-clusters into the respective leftover set. This is easy to do by a simple fact which is proven in Section 4:

[2]We visualize the vertex sets V_i as being horizontal, like rows in a matrix, so it is natural to think of these k-tuples as columns.

Fact 12. *Let $\varepsilon' > 0$ and $\varepsilon' < d'/(2k+2)$. Let $(A_1, \ldots, A_k)$ be a k-tuple that is pairwise ε'-regular of density at least d' with $|A_1| = \cdots = |A_k| = L'$. There exist subsets $A'_i \subset A_i$ for $i \in [k]$ such that $|A_i| = h\lceil(1-(k-1)\varepsilon')L'/h\rceil$ and each pair of $(A'_1, \ldots, A'_k)$ is $(2\varepsilon', d' - k\varepsilon')$-super-regular (with density at least $d' - \varepsilon'$).*

Fact 12 follows from well-known properties of regular pairs. We apply it to each k-tuple $(\tilde{V}_1^{(j)}, \ldots, \tilde{V}_k^{(j)})$, $j \in [\ell']$. We do not rename the sets $\tilde{V}_i^{(j)}$ since they only shrink in magnitude only by $(k-1)\varepsilon' L'$. Consequently,

- the leftover sets $\tilde{V}_i^{(0)}$, $1 \leq i \leq k$, are of size at most $2\varepsilon n + (k-1)\varepsilon' L' \ell' < k\varepsilon' n$,
- each pair $(\tilde{V}_i^{(j)}, \tilde{V}_{i'}^{(j)})$, $i \neq i'$, is $(2\varepsilon', d'/2)$-super-regular, and
- each pair $(\tilde{V}_i^{(j)}, \tilde{V}_{i'}^{(j')})$ is $2\varepsilon'$-regular with density either 0 or at least $d' - \varepsilon'$, regardless of whether or not $j = j'$.

If we use the Blow-up Lemma (Lemma 18) at this point, we would obtain a K_h^k-tiling that covers every vertex of G except those in the leftover sets. The remainder of the proof is to establish that we can, in fact, ensure that the leftover vertices can be absorbed by the sub-clusters and we can obtain a K_h^k-tiling that covers all the vertices of G.

3.4. Preparing for absorption

In order to absorb the vertices from the leftover sets, we need to prepare some copies of K_h^k throughout G that may be included in the final K_h^k-tiling. Their purpose is to ensure that, after inserting vertices from the leftover sets to the sub-clusters, the number of vertices in each of the sub-clusters can be balanced so that the Blow-up Lemma (Lemma 18) can be used. These copies of K_h^k will be specially designated and colored either red or blue according to their role.

The Reachability Lemma (Lemma 13) is how we transfer the imbalance of the sizes of one column to the first column.

Lemma 13 (Reachability Lemma). *Let G'_r be a balanced k-partite graph with partite sets $U'_i = \{u_i^{(j)} : j \in [\ell']\}$, $i \in [k]$. Let $\hat{\delta}_k(G'_r) \geq \frac{k-1}{k}\ell' + 2$. Then, for each $i \in [k]$ and $j \in \{2, \ldots, \ell'\}$, there is a pair (T_1, T_2) of copies of K_k such that their symmetric difference is $\{u_i^{(1)}, u_i^{(j)}\}$ and T_1 and T_2 contain no additional vertices from $\{u_1^{(1)}, \ldots, u_k^{(1)}, u_1^{(j)}, \ldots, u_k^{(j)}\}$. See Figure 3.4.*

Proof of Lemma 13. Without loss of generality, it suffices to prove the lemma for $i = 1$ and $j = \ell'$. The vertices $u_1^{(1)}$ and $u_1^{(\ell')}$ have at least $\ell' - 2(\ell' - \hat{\delta}_k(G'_r)) \geq \ell' - 2\left(\ell' - \frac{k-1}{k}\ell' - 2\right) = \left(\frac{k-2}{k}\right)\ell' + 4$ common neighbors in each of $U'_2, \ldots, U'_k$. Hence, one can choose a sequence $w_2, \ldots, w_k$ of vertices so that, for $i = 2, \ldots, k$,

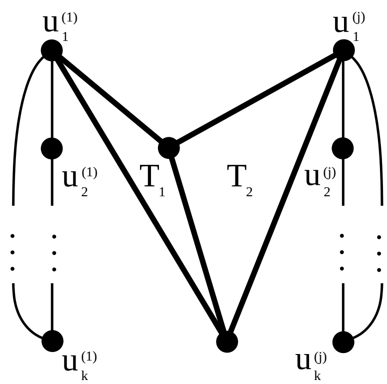

Figure 1: Diagram for T_1 and T_2 formed in reaching $u_1^{(1)}$ from $u_1^{(j)}$.

w_i is in $U_i' - \{u_i^{(1)}, u_i^{(\ell')}\}$ and is a common neighbor of $u_1^{(1)}, u_1^{(\ell')}, w_2, \ldots, w_{i-1}$. Note that at each stage, the number of available choices for w_i is at least

$$\left(\frac{k-2}{k}\ell' + 4\right) - (i-2)\left(\ell' - \frac{k-1}{k}\ell' - 2\right) - 2 = \frac{k-i}{k}\ell' + 2i - 2.$$

This quantity is positive since $1 \le i \le k$ and $k \ge 2$. $\square$

In preparation to insert the vertices, we create a set of special vertex-disjoint copies of K_h^k.

Lemma 14. *There exist disjoint sets $X_i(j) \subset \tilde{V}_i^{(1)}$, $i \in [k]$, $j \in [\ell'] - \{1\}$, such that for every $i \in [k]$, $j \in [\ell'] - \{1\}$:*

(1) $|X_i(j)| = 3h\zeta n$.

(2) For every $v \in X_i(j)$, there exist two vertex-disjoint copies of K_h^k, call them $\mathcal{R}(v)$ and $\mathcal{B}(v)$, such that

- *(i) $\mathcal{R}(v)$ contains v,*
- *(ii) $\mathcal{R}(v)$ contains $h-1$ vertices from $\tilde{V}_i^{(j)}$ and $\mathcal{B}(v)$ contains h vertices from $\tilde{V}_i^{(j)}$, and*
- *(iii) for every $i' \neq i$, there exists a $j' \notin \{1, j\}$ such that both $\mathcal{R}(v)$ and $\mathcal{B}(v)$ each have h vertices from $\tilde{V}_{i'}^{(j')}$.*

(3) The $2|X_i(j)|$ copies of K_h^k, namely $\mathcal{R}(v)$ and $\mathcal{B}(v)$ for all $v \in X_i(j)$, are all pairwise-disjoint.

Proof of Lemma 14. The proof will proceed as follows: We will have some arbitrary order on the pairs $\{(i,j) : i \in [k], j \in [\ell'] - \{1\}\}$ and dynamically define

$$X = \bigcup_{(i',j') \prec (i,j)} \bigcup_{v \in X_{i'}(j')} \left(V(\mathcal{R}(v)) \cup V(\mathcal{B}(v))\right).$$

That is, X is the set of all vertices belonging to a $\mathcal{R}(v)$ or a $\mathcal{B}(v)$ for all (i', j') that precede the current (i, j).

We will show that, for all $v \in X_i^{(j)}$ the vertex-disjoint $\mathcal{R}(v)$ and $\mathcal{B}(v)$ can be found among vertices not in X, as long as $|X| \le \zeta^{1/2} L'$.

Fix $i \in [k]$ and $j \in [\ell']$. Let (T_1, T_2) be a pair of K_k-s in G'_r from Lemma 13 for these values of i and j. Consider the subgraph F of G' induced on the sub-clusters $\tilde{V}_{i'}^{(j')}$ such that $u_{i'}^{(j')}$ form $V(T_2)$. Since T_2 is a K_k in the reduced graph G'_r, every pair of sub-clusters in this subgraph is ε'-regular with density at least d'. Since $|X| \le \zeta^{1/2} L'$, $|\tilde{V}_{i'}^{(j')} - X| \ge \frac{1}{2} |\tilde{V}_{i'}^{(j')}|$ and it follows from the definition of regularity that each pair of sub-clusters of $F - X$ is $2\varepsilon'$-regular with density at least $d' - \varepsilon'$. By the Key Lemma (Lemma 2.1 from [21]), $F - X$ contains at least $3h\zeta n$ vertex-disjoint copies of K_h^k as long as $3h\zeta n \ll \varepsilon' L'$. This is satisfied because

$$3h\zeta n \overset{(\mathrm{P2}')}{\le} 3h\zeta \frac{\ell' L'}{1 - 2\varepsilon} \le 4h\zeta \ell' L' = 4h\zeta \ell D L' \le 4h\zeta M D L' \overset{(3)}{\ll} \varepsilon' L'.$$

In the above inequality, we use the fact that $\ell' = \ell \cdot D \le M \cdot D$ and M and D depend only on k and ε. In addition, $\zeta \ll \varepsilon$. We refer to these $3h\zeta n$ copies of K_h^k as *blue copies of* K_h^k and we add their vertices to X.

In a similar fashion, let F now be the graph induced on the sub-clusters $\tilde{V}_{i'}^{(j')}$ such that $u_{i'}^{(j')}$ is in $V(T_1) \cup V(T_2)$. The graph $F - X$ also satisfies the assumptions of the Key Lemma and therefore we can find $3h\zeta n$ copies of K_h^k in such a way that each copy has one vertex in $\tilde{V}_i^{(1)} - X$ and $h - 1$ vertices in $\tilde{V}_i^{(j)} - X$. The remaining vertices of K_h^k are in the sub-clusters of $V(T_1) \cap V(T_2)$. We refer to these $3h\zeta n$ copies of K_h^k as *red copies of* K_h^k and add their vertices to X. For each red copy of K_h^k, we put its unique vertex in $V_i^{(1)}$ into $X_i(j)$ and call this copy $\mathcal{R}(v)$. For each $v \in X_i(j)$, let $\mathcal{B}(v)$ be a distinct blue copy of K_h^k as found above.

For this process to work, we need to ensure that $|X| \le \zeta^{1/2} L'$ at each step. This is true because each member of each $X_i(j)$ corresponds to two K_h^k-s which have a total of $2hk$ vertices and, hence, $|X| \le 2hk \sum_i \sum_j |X_i(j)| = 2hk(k\ell') \cdot (3h\zeta n) = 6h^2 k^2 \ell' \zeta n \ll \zeta^{1/2} L'$. □

We color the vertices of each $\mathcal{R}(v)$ red and the vertices of each $\mathcal{B}(v)$ blue.

3.5. Nearly-equalizing the sizes of the sub-clusters

Let us summarize where we are: We have a designated first column (we call the first column the *receptacle column* and its sub-clusters *receptacle sub-clusters*) with each sub-cluster of size L' and each sub-cluster having the same number of red vertices, which is at most $\zeta^{1/2} L'$. Each such red vertex is in a different vertex-disjoint red copy of K_h^k. In the remaining columns, each sub-cluster has L' original vertices, of which at most $\zeta^{1/2} L'$ are colored red and at most $\zeta^{1/2} L'$ are colored blue. The total number of red vertices in each V_i is the same multiple of h. Moreover, in every column, every pair of sub-clusters is

$(2\varepsilon', d'/2)$-super regular. Finally, for each $i \in [k]$, there is a leftover set $\tilde{V}_i^{(0)}$ of size at most $k\varepsilon' n$.

We shall now re-distribute the vertices from leftover sets $\tilde{V}_i^{(0)}$, $i \in [k]$, to non-receptacle sub-clusters in such a way that the size of leftover sets becomes $O(n)$ and each non-receptacle sub-cluster will contain exactly $h\lceil (1-d'/4)(L'/h)\rceil$ non-red vertices. These two properties will be essential for our procedure for finding perfect K_h^k-tiling to work.

We say that a vertex $v \in V_i$ *belongs in the sub-cluster* $\tilde{V}_i^{(j)}$ if v is adjacent to at least $(d'/2)L'$ vertices in each of the other sub-clusters $\tilde{V}_{i'}^{(j)}$, $i \neq i'$, in the j-th column.

Fact 15. *For every $i \in [k]$, we can partition the leftover set $\tilde{V}_i^{(0)}$ into subsets $Y_i^{(2)}, \dots, Y_i^{(\ell')}$ where, for every $j \in \{2, \dots, \ell'\}$, the members of $Y_i^{(j)}$ belong in sub-cluster $\tilde{V}_i^{(j)}$ and*

$$\left|Y_i^{(j)}\right| \leq \frac{k\varepsilon' n}{(1/k+\gamma/2)\ell'} \leq k^2\varepsilon' L'.$$

The number of red vertices in each sub-cluster may vary, but it is always less than $\zeta^{1/2}L'$. Hence, after applying Fact 15, the number of non-red vertices in each sub-cluster is in the interval $\big((1-\zeta^{1/2})L', (1+k^2\varepsilon')L'\big)$. Fact 15 is proved in Section 4.

Next, we wish to remove copies of K_h^k in such a way that the number of non-red vertices in each non-receptacle sub-cluster is the same and there are new leftover sets of size $O(\zeta n)$. This is accomplished via Lemma 16. After we insert vertices via Fact 15 and remove some to create a (much smaller) leftover set via Lemma 16, the sets $\tilde{V}_i^{(j)}$ will be slightly changed into sets $\hat{V}_i^{(j)}$ for $i \in [k]$ and $j \in \{0, 1, \dots, \ell'\}$.

Lemma 16. *For each $i \in [k]$, there exist disjoint vertex sets $\hat{V}_i^{(0)}, \hat{V}_i^{(1)}, \dots, \hat{V}_i^{(\ell')}$ in V_i such that the following occurs:*

- $|\hat{V}_i^{(0)}| \leq 3h\zeta n$,
- $\hat{V}_i^{(1)} = \tilde{V}_i^{(1)}$*, has exactly $(\ell'-1)3h\zeta n$ red vertices and exactly L' vertices total,*
- *for $j \in \{2, \dots, \ell'\}$, $\hat{V}_i^{(j)} \subset \tilde{V}_i^{(j)}$ and $\hat{V}_i^{(j)}$ contains all red and blue vertices of $\tilde{V}_i^{(j)}$,*
- *for $j \in \{2, \dots, \ell'\}$, $\hat{V}_i^{(j)}$ contains exactly $h\left\lceil \left(1-\frac{d'}{4}\right)\frac{L'}{h}\right\rceil$ non-red vertices, and*
- *the graph induced by $V(G') - \bigcup_{i=1}^k \bigcup_{j=0}^{\ell'} \hat{V}_i^{(j)}$ is spanned by the union of vertex-disjoint copies of K_h^k.*

Proof of Lemma 16. In this proof, we will remove some copies of K_h^k to thin the graph so that the sub-clusters satisfy the conditions above. We shall do this by taking the reduced graph G_r and creating an auxiliary graph A_r and then we apply Theorem 7 to A_r. From the resulting fractional K_k-tiling in A_r, we will produce a family of vertex-disjoint K_h^k-s in G that we shall remove.

From Section 3.2, recall that $D = D(k, \varepsilon)$ was the least common multiple of a common denominator of a rational-valued solution to linear program (1) over all balanced k-partite graphs with at most $k \cdot M = k \cdot M(k, \varepsilon)$ vertices. In a similar way, we may define $D_0 = D_0(k, \varepsilon, \zeta)$ to be the least common multiple of the common denominator of a rational-valued solution to linear program (1) over all balanced k-partite graphs with at most $\frac{d'}{3h\zeta}\ell' \leq \frac{1}{3h\zeta}D(k, \varepsilon)M(k, \varepsilon)$ vertices in each class.

Now we will define the auxiliary reduced graph A_r by blowing up the vertices and edges of the subgraph of G'_r induced by $V(G'_r) - \{u_1^{(1)}, u_2^{(1)}, \ldots, u_k^{(1)}\}$. The number of copies of each vertex, however, will not be the same. For $i \in [k]$ and $j \in \{2, \ldots, \ell'\}$, define $\nu(u_i^{(j)})$ to be the number of non-red vertices in sub-cluster $\tilde{V}_i^{(j)}$.

For $V(A_r)$, replace each vertex $u_i^{(j)}$ with the following number of copies: either the ceiling or the floor of

$$\frac{\nu(u_i^{(j)}) - \lceil (1 - d'/4)L' \rceil}{hD_0 \lceil \zeta L'/D_0 \rceil} - 1.$$

The choice of ceiling or floor is made arbitrarily, but only to ensure that the resulting graph is balanced. This is always possible because $\sum_{j=2}^{\ell'} \nu(u_i^{(j)})$ is the same for all $i \in [k]$. For $E(A_r)$, we replace each edge in G'_r by a complete bipartite graph and each nonedge by an empty bipartite graph.

First, we need to check that the number of vertices of A_r is not too large. Since $(1-\zeta^{1/2})L' \leq \nu(u_i^{(j)}) \leq (1+k^2\varepsilon')L'$, the number of vertices in each partite set of A_r is at most

$$\begin{aligned}
\sum_{j=2}^{\ell'} & \left\lceil \frac{\nu(u_i^{(j)}) - \lceil (1 - d'/4)L' \rceil}{hD_0 \lceil \zeta L'/D_0 \rceil} - 1 \right\rceil \\
& \leq (\ell' - 1) \left\lceil \frac{(1 + k^2\varepsilon')L' - (1 - d'/4)L'}{h\zeta L'} - 1 \right\rceil \\
& < (\ell' - 1) \frac{d'/4 + k^2\varepsilon'}{h\zeta}. \qquad (4)
\end{aligned}$$

This quantity is at most $\frac{d'}{3h\zeta}\ell'$ because $\varepsilon' \ll d'$.

Second, we need to check that each vertex A_r has sufficiently large degrees in order to apply Theorem 7. We observe that if u were adjacent to $u_i^{(j)}$ in G'_r, then every copy of u in $V(A_r)$ is adjacent to at least

$$\left\lfloor \frac{(1 - \zeta^{1/2})L' - \lceil (1 - d'/4)L' \rceil}{hD_0 \lceil \zeta L'/D_0 \rceil} - 1 \right\rfloor \geq \frac{d'/4 - 2\zeta^{1/2}}{h\zeta}$$

copies of $u_i^{(j)}$ in $V(A_r)$. So, each vertex in $V(A_r)$ is adjacent to at least

$$\left[\left(\frac{k-1}{k}+\frac{\gamma}{2}\right)\ell'-1\right]\frac{d'/4-2\zeta^{1/2}}{h\zeta}\geq(\ell'-1)\left(\frac{k-1}{k}+\frac{\gamma}{3}\right)\frac{d'}{4h\zeta}$$

vertices in each of the other partite sets of $V(A_r)$. By (4), every partite set of $V(A_r)$ has size at most $(\ell'-1)\frac{d'/4+k^2\varepsilon'}{h\zeta}$. Using (3), the proportion of neighbors of a vertex in $V(A_r)$ in any other vertex class is at least

$$\frac{(\ell'-1)\left(\frac{k-1}{k}+\frac{\gamma}{3}\right)\frac{d'}{4h\zeta}}{(\ell'-1)\frac{d'/4+k^2\varepsilon'}{h\zeta}}\geq\frac{k-1}{k}.$$

So, we can apply Theorem 7 to the auxiliary reduced graph A_r and obtain an optimal solution to linear program (1) with the property that $D_0 w(T)$ is an integer for every $T\in\mathcal{T}_k(A_r)$.

As in Fact 11, this implies that the blow-up graph $A_r(D_0)$ must have a perfect K_k-tiling. For each K_k in this tiling, we will remove $\lceil\zeta L'/D_0\rceil$ vertex-disjoint copies of K_h^k from the uncolored vertices of the corresponding sub-clusters of G_r'.

It is easy to find such vertex-disjoint copies of K_h^k in a k-tuple. Observe that every sub-cluster has at most $k^2\varepsilon' L'$ uncolored vertices added to the sub-cluster. Moreover, a set of $hD_0\lceil\zeta L'/D_0\rceil$ vertices will be removed from a sub-cluster at most $d'/(2\zeta)$ times as long as $\varepsilon'\ll d'$. So, there will always be at least $\lceil(1-d'/4)L'\rceil-k^2\varepsilon'L'-hd'L'/2\geq(1-d')L'$ uncolored vertices from the original sub-cluster. Using the Slicing Lemma (Fact 19), any pair of them form a $2(2\varepsilon')$-regular pair. As long as $\zeta\ll\varepsilon'\ll d'$, we could apply, say, the Key Lemma from [21] to ensure the existence of at most $\lceil\zeta L'/D_0\rceil$ vertex-disjoint copies of K_h^k in the k-tuple.

So, the total number of vertices removed from sub-cluster $V_i^{(j)}$ is

$$hD_0\lceil\zeta L'/D_0\rceil\times\left\lfloor\frac{\nu(u_i^{(j)})-\lceil(1-d'/4)L'\rceil}{hD_0\lceil\zeta L'/D_0\rceil}-1\right\rceil,$$

where $\lfloor\cdot\rceil$ is either the floor or ceiling of its argument.

Removing these copies of K_h^k has the effect of making the number of uncolored vertices in each sub-cluster nearly identical, that is, within $h\zeta L'$ of each other. For $i\in[k]$, place into the new leftover set of V_i at most $hD_0\lceil\zeta L'/D_0\rceil-1$ uncolored vertices from each sub-cluster to ensure that every sub-cluster retains either $\lceil(1-d'/4)L'\rceil$ or $\lceil(1-d'/4)L'\rceil+hD_0\lceil\zeta L'/D_0\rceil$ uncolored vertices, depending on whether the ceiling or floor function was chosen for rounding. In the latter case, place an additional $hD_0\lceil\zeta L'/D_0\rceil$ uncolored vertices from the sub-cluster to the leftover set.

Summarizing:

- We placed into each leftover set at most $2hD_0\lceil\zeta L'/D_0\rceil$ vertices from each sub-cluster, so each new leftover set $\hat{V}_i^{(0)}$ has a size of at most $\ell'\cdot 2hD_0\lceil\zeta L'/D_0\rceil\leq 3h\zeta n$.

- The sets $\tilde{V}_i^{(1)}$ are unchanged.
- For $j \in \{2, \ldots, \ell'\}$, $\hat{V}_i^{(j)}$ is formed by removing uncolored vertices from $\tilde{V}_i^{(j)}$.
- For $j \in \{2, \ldots, \ell'\}$, the number of non-red vertices in $\hat{V}_i^{(j)}$ is explicitly prescribed to be $h \lceil (1 - d'/4)(L'/h) \rceil$ because later we need it to be divisible by h.
- The vertices that are removed are all in vertex-disjoint copies of K_h^k.

□

3.6. Inserting the leftover vertices and construction of perfect K_h^k-tiling

We first insert the leftover vertices from $\bigcup_{i=1}^k \hat{V}_i^{(0)}$ to non-receptacle sub-clusters in such a way that we shall be able to find a perfect K_h^k-tiling in every column using the Blow-up Lemma. That is, each sub-cluster in the column will have the same number of vertices (divisible by h) and each pair of sub-clusters will be super-regular.

Suppose that vertex $w \in \hat{V}_i^{(0)}$ belongs in the sub-cluster $\tilde{V}_i^{(j)}$, $j \in \{2, \ldots, \ell'\}$. We then take any $v \in X_i^{(j)}$ and the red and blue copies $\mathcal{R}(v), \mathcal{B}(v)$ of K_h^k guaranteed by Lemma 14. We uncolor the vertices of $\mathcal{R}(v)$, remove the vertices of $\mathcal{B}(v)$ from their respective sub-clusters and place $\mathcal{B}(v)$ aside to be included in the final tiling of G. We also add w to the sub-cluster $\tilde{V}_i^{(j)}$ and remove v from $X_i^{(j)}$.

Each time this procedure is undertaken, the number of non-red vertices in each non-receptacle sub-cluster does not change and it is equal to $h \lceil (1 - d'/4)(L'/h) \rceil$.

After doing this procedure for every vertex in the leftover sets, we remove all the remaining (unused) red copies of of K_h^k and place them aside to be included in the final tiling of G. The sub-clusters in the first (receptacle) column have the same number of non-red vertices as each other and the number of non-red vertices in each receptacle sub-cluster has the same congruency modulo h as n does. That is, if we remove $n - h\lfloor n/h \rfloor$ non-red vertices from each receptacle sub-cluster, the remaining number of non-red vertices is divisible by h.

The non-red vertices in each receptacle sub-cluster form pairwise $(4\varepsilon', d'/4)$-super-regular pairs, this follows from the Slicing Lemma (Fact 19) because no vertices were added to these sub-clusters. So we focus on the non-receptacle sub-clusters.

Throughout this proof, in every non-receptacle sub-cluster, at most $\varepsilon' L'$ vertices were colored red and at most $\varepsilon' L'$ red vertices will be uncolored (i.e., they become non-red). In addition, the non-red vertices in any non-receptacle sub-cluster will have cardinality exactly $h\lceil (1 - d'/4)L'/h \rceil$. Recall that the original sub-clusters formed $(2\varepsilon', d')$-super-regular pairs in each column. There were at most $k^2\varepsilon' L'$ new vertices added to each sub-cluster, each of which were adjacent to at least $(d'/2)L'$ vertices in each of the original sub-clusters of the column. The next lemma will imply that the non-red vertices in every non-receptacle column will form super-regular pairs.

Fact 17. *Let (A, B) be an $(\varepsilon_1, \delta_1)$-super-regular pair. Furthermore, let $A' \supset A$ and $B' \supset B$ be such that $|A' - A| \leq \varepsilon_2 |A|$ and $|B' - B| \leq \varepsilon_2 |B|$. If*

- *every vertex in $A' - A$ has at least $\delta_2 |B|$ neighbors in B and*
- *every vertex in $B' - B$ has at least $\delta_2 |A|$ neighbors in A,*

then the pair (A', B') is $(\varepsilon_0, \delta_0)$-super-regular, where $\delta_0 = \frac{\min\{\delta_1, \delta_2\}}{(1+\varepsilon_2)^2}$ and $\varepsilon_0 = \varepsilon_1 + \varepsilon_2$.

We apply Fact 17 with $\varepsilon_1 = 2\varepsilon'$, $\delta_1 = d'$, $\varepsilon_2 = k\varepsilon'$ and $\delta_2 = d'/2$. Consequently, we use $\varepsilon_1 + \varepsilon_2 \leq (k+2)\varepsilon' \leq \sqrt{\varepsilon'}$ and $\frac{\min\{\delta_1, \delta_2\}}{(1+\varepsilon_2)^2} = \frac{d'/2}{(1+k\varepsilon')^2} \geq d'/3$ to conclude that the augmented pairs in each column are $(\sqrt{\varepsilon'}, d'/3)$-super-regular.

Finally, to finish the tiling, apply the Blow-up Lemma to non-red vertices in each non-receptacle column (recall that the number of such vertices is the same and is divisible by h). We can also apply the Blow-up Lemma to the non-red vertices in the receptacle column as well, because the sizes of those sets are divisible by h.

Lemma 18 (Blow-up Lemma, Komlós-Sárközy-Szemerédi [19]). *Given a graph R of order r and positive parameters δ, Δ, there exists an $\varepsilon_{\mathrm{BL}} > 0$ such that the following holds: Let N be an arbitrary positive integer, and let us replace the vertices of R with pairwise disjoint N-sets $V_1, V_2, \ldots, V_r$ (blowing up). We construct two graphs on the same vertex-set $V = \bigcup V_i$. The graph $R(N)$ the graph which is the blow-up of R by N and a sparser graph G is constructed by replacing the edges of R with some $(\varepsilon_{\mathrm{BL}}, \delta)$-super-regular pairs. If a graph H with maximum degree $\Delta(H) \leq \Delta$ can be embedded into $R(N)$, then it can be embedded into G.*

Our K_h^k-tiling consists of

(i) the copies of K_h^k that are outside of the sets $\hat{V}_i^{(j)}$, as established in Lemma 16,

(ii) the red copies of K_h^k that were not uncolored in the process of absorbing vertices from the leftover sets $\hat{V}_i^{(0)}$ to non-receptacle sub-clusters, and

(iii) the copies of K_h^k found by applying the Blow-up Lemma to the non-red vertices in each column.

This is the tiling of G with n/h copies of K_h^k.

What remains to show is that we can choose our constants to satisfy (3) so that all inequalities in our proof will be satisfied for sufficiently large n. Indeed, for given $\gamma > 0$ and h, we let $d = \gamma/4$. We also set $R = K_k$, $r = k$, $\Delta = (k-1)h$ and $\delta = \gamma/12$ and apply Lemma 18 to obtain $\varepsilon_{\mathrm{BL}}$. Now we define $\varepsilon' = \min\{\varepsilon_{18}^2, d/(12k^2)\}$ and we let $\varepsilon = \min\{(\varepsilon')^5/16, d/4(k+2)\}$. Finally, we set $\zeta = 1/(12h^2k^2M(k,\varepsilon)^2D(k,\varepsilon)^2)$, where $M(k,\varepsilon)$ comes from the Regularity Lemma (Theorem 8) and $D(k,\varepsilon)$ is defined in Section 3.2. This concludes the proof of Theorem 4.

4. Proofs of Facts

For convenience, we restate the facts to be proven.

Fact 15. *For every $i \in [k]$, we can partition the leftover set $\tilde{V}_i^{(0)}$ into subsets $Y_i^{(2)}, \ldots, Y_i^{(\ell')}$ where, for every $j \in \{2, \ldots, \ell'\}$, the members of $Y_i^{(j)}$ belong in sub-cluster $\tilde{V}_i^{(j)}$ and $|Y_i^{(j)}| \le \frac{k\varepsilon' n}{(1/k+\gamma/2)\ell'} \le k^2\varepsilon' L'$.*

Proof of Fact 15. First, we show that each vertex belongs in at least $(1/k + \gamma/2)\ell'$ sub-clusters. To see this, let x be the number of sub-clusters in $V_{i'}$, $i' \neq i$ such that v is adjacent to less than $(d'/2)L'$ vertices of that sub-cluster. Then, since $n - \ell' L' \le 2\varepsilon n$,

$$x\frac{d'}{2}L' + (\ell' - x)L' + (n - \ell' L') \ge \left(\frac{k-1}{k} + \gamma\right)n.$$

From this it is easy to derive that with d', ε' small enough relative to γ, it is the case that $x < (1/k - \gamma/2)\ell'$. By a simple union bound, the number of sub-clusters in which v belongs is greater than $\ell' - (k-1)(1/k - \gamma/2)\ell' \ge (1/k + \gamma/2)\ell'$. Hence, there are at least $(1/k + \gamma/2)\ell'$ sub-clusters outside of the receptacle column in which v belongs.

Sequentially and arbitrarily assign $v \in \tilde{V}_i^{(0)}$ to $Y_i^{(j)}$ if both v belongs in $\tilde{V}_i^{(j)}$ and $|Y_i^{(j)}| < \frac{k\varepsilon' n}{(1/k+\gamma/2)\ell'}$. Since the size of $\tilde{V}_i^{(0)}$ is at most $k\varepsilon' n$, we can always find a place for v. □

Fact 12. *Let $\varepsilon' > 0$ and $\varepsilon' < d'/(2k+2)$. Let $(A_1, \ldots, A_k)$ be a k-tuple that is pairwise ε'-regular of density at least d' with $|A_1| = \cdots = |A_k| = L'$. There exist subsets $A_i' \subset A_i$ for $i \in [k]$ such that $|A_i| = h\lceil (1-(k-1)\varepsilon')L'/h \rceil$ and each pair of $(A_1', \ldots, A_k')$ is $(2\varepsilon', d' - k\varepsilon')$-super-regular (with density at least $d' - \varepsilon'$).*

Proof of Fact 12. We use the so-called Slicing Lemma [21, Fact 1.5].

Fact 19 (Slicing Lemma [21]). *Given ε, α, d such that $0 < \varepsilon < \alpha < 1$ and $d, 1-d \ge \max\{2\varepsilon, \varepsilon/\alpha\}$. Let (A, B) be an ε-regular pair with density d, $A' \subset A$ with $|A'| \ge \alpha|A|$ and $B' \subset B$ with $|B'| \ge \alpha|B|$. Then (A', B') is ε_0-regular with $\varepsilon_0 = \max\{2\varepsilon, \varepsilon/\alpha\}$ and density in $[d-\varepsilon, d+\varepsilon]$.*

It follows from the ε'-regularity of (A_i, A_j) that all but at most $\varepsilon'|A_i|$ vertices of A_i have at least $(d' - \varepsilon')|A_j|$ neighbors in A_j. So, there is a set $A_i' \subset A_i$ of size $(1-(k-1)\varepsilon')|A_i|$ such that each vertex of A_i' has at least $(d'-\varepsilon')|A_j|$ neighbors in A_j for every $j \neq i$ and, consequently, at least $(d'-\varepsilon')|A_j| - (k-1)\varepsilon'|A_j| = (d' - k\varepsilon')|A_j'|$ neighbors in A_j' for every $j \neq i$.

Since $\varepsilon' < d'/(2k+2)$ and $(1-(k-1)\varepsilon') > 1/2$, the Slicing Lemma (Fact 19) with $\alpha = 1/2$ and $\varepsilon' < d'/(2k+2)$ gives that each pair (A'_i, A'_j) is $(2\varepsilon', d - k\varepsilon')$-super-regular. $\square$

Fact 17. Let (A, B) be an $(\varepsilon_1, \delta_1)$-super-regular pair. Furthermore, let $A' \supset A$ and $B' \supset B$ be such that $|A' - A| \le \varepsilon_2 |A|$ and $|B' - B| \le \varepsilon_2 |B|$. If every vertex in $A' - A$ has at least $\delta_2 |B|$ neighbors in B and every vertex in $B' - B$ has at least $\delta_2 |A|$ neighbors in A, then the pair (A', B') is $(\varepsilon_0, \delta_0)$-super-regular, where $\varepsilon_0 = \varepsilon_1 + \varepsilon_2$ and $\delta_0 = \frac{\min\{\delta_1, \delta_2\}}{(1+\varepsilon_2)^2}$.

Proof of Fact 17. First we establish the minimum degree condition. Each of the vertices in A is adjacent to at least $\delta_1 |B| = \delta_1 \frac{|B|}{|B'|} |B'|$ vertices in B'. Each of the vertices in $A' - A$ is adjacent to at least $\delta_2 |B| = \delta_2 \frac{|B|}{|B'|} |B'|$ neighbors in B'. Similar conditions hold for vertices in B'.

Since

$$\delta_0 \le \frac{\min\{\delta_1, \delta_2\}}{(1+\varepsilon_2)^2} \le \min\left\{\delta_1 \frac{|A|}{|A'|}, \delta_2 \frac{|A|}{|A'|}, \delta_1 \frac{|B|}{|B'|}, \delta_2 \frac{|B|}{|B'|}\right\},$$

each vertex $a \in A'$ has at least $\delta_0 |B'|$ neighbors in B' and each vertex $b \in B'$ has at least $\delta_0 |A'|$ neighbors in A'.

Now, consider any $X' \subseteq A'$ and $Y' \subseteq B'$ such that $|X'| \ge \varepsilon_0 |A'|$ and $|Y'| \ge \varepsilon_0 |B'|$. Consider $X = X' - (A' - A)$ and $Y = Y' - (B' - B)$. Note that

$$|X| \ge |X'| - \varepsilon_2 |A| \ge \varepsilon_0 |A'| - \varepsilon_2 |A| \ge \varepsilon_1 |A|.$$

Similarly, $|Y| \ge \varepsilon_1 |B|$ and so $d(X, Y) \ge \delta_1$.

Consequently,

$$d(X', Y') \ge d(X, Y) \frac{|X||Y|}{|X'||Y'|} \ge \frac{\delta_1}{(1+\varepsilon_2)^2} \ge \delta_0,$$

and the pair is $(\varepsilon_0, \delta_0)$-super-regular. $\square$

5. Concluding Remarks

The common denominator $D = D(k, \varepsilon)$ used in Section 3.2 can, in principle, be astronomically large, as it is the common denominator of values of rational-valued solutions for all balanced k-partite graphs on at most $M = M(k, \varepsilon)$ vertices. We chose this value for the convenience of the proof. Indeed, the constant M is quite large itself and so D is not so large, relatively speaking.

We could utilize a much smaller integer value of D by choosing a D such that if w^* is the rational-valued solution of (1), then for every $v \in V(G_r)$ and every $T \in \mathcal{T}_k(G_r)$ for which $V(T) \ni v$, we assign $\lfloor D w^*(T) \rfloor$ vertices of G'_r to copies of T. Because $D w^*(T)$ is not necessarily an integer, we end up with $D - \sum_{V(T) \ni v} \lfloor D w^*(T) \rfloor$ unused vertices. Choose D large enough to ensure that this is always small ($O(\varepsilon M^{k-1})$ suffices), and they can be placed in the leftover set.

We should also note that, asymptotically, Conjecture 3 is stronger than the Hajnal-Szemerédi Theorem. That is, if G is a graph on kn vertices with minimum degree at least $\left(\frac{k-1}{k}+\gamma\right)kn$, then a random partition of the vertex set into k equal parts gives a k-partite graph $\tilde{G}$ with $\hat{\delta}_k(\tilde{G}) \geq \left(\frac{k-1}{k}+\gamma\right)kn - O(\sqrt{n}\log n)$ and applying Conjecture 3 would give a K_k-tiling in $\tilde{G}$ and, hence, G itself.

6. Acknowledgements

The authors would like to thank the London School of Economics and Political Science. Part of this research had been done as part of a visit by the first author. We express our appreciation to Hong-Jian Lai who has made the publications of Paul Catlin available online (`http://www.math.wvu.edu/~hjlai/`). Thanks also to Bernhard von Stengel for helpful comments on linear programming and to Richard Mycroft for useful conversations. Special thanks to an anonymous referee whose comments improved the paper.

References

[1] N. Alon and R. Yuster, H-factors in dense graphs. *J. Combin. Theory Ser. B* **66** (1996), no. 2, 269–282.

[2] A. Bush and Y. Zhao, Minimum degree thresholds for bipartite graph tiling. *J. Graph Theory* **70** (2012), no. 1, 92–120.

[3] P.A. Catlin, On the Hajnal-Szemerédi theorem on disjoint cliques. *Utilitas Math.* **17** (1980), 163–177.

[4] V. Chvátal, *Linear Programming.* A series of Books in the Mathematical Sciences. *W. H. Freeman and Company, New York*, 1983. xiii+478 pp.

[5] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph. *Acta Math. Acad. Sci. Hungar.* **14** (1963), 423–439.

[6] B. Csaba and M. Mydlarz, Approximate multipartite version of the Hajnal-Szemerédi theorem. *J. Combin. Theory Ser. B* **102** (2012), no. 2, 395–410.

[7] A. Czygrinow and L. DeBiasio, A note on bipartite graph tiling, *SIAM J. Discrete Math.*, **25** (2011), no. 4, 1477–1489.

[8] E. Fischer, Variants of the Hajnal-Szemerédi theorem, *J. Graph Theory* **31** (1999), no. 4, 275–282.

[9] S. I. Gass, *Linear programming. Methods and applications.* Third edition, McGraw-Hill Book Co., New York-London-Toronto, Ont. 1969 xiii+358 pp.

[10] A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős, *Combinatorial theory and its applications, II (Proc. Colloq. Balatonfüred, 1969)*, pp. 601–623. *North-Holland*, Amsterdam, 1970.

[11] P. Hall, On representation of subsets. *J. London Math. Soc.* **10** (1935), 26–30.

[12] J. Hladký and M. Schacht, Note on bipartite graph tilings. *SIAM J. Discrete Math.* **24** (2010), no. 2, 357–362.

[13] R. Johansson, Triangle factors in a balanced blown-up triangle. *Discrete Math.* **211** (2000), no. 1–3, 249–254.

[14] P. Keevash and R. Mycroft, A geometric theory for hypergraph matching, *Mem. Amer. Math. Soc.* **233** (2015), no. 1098 vi+95 pp.

[15] P. Keevash and R. Mycroft, A multipartite Hajnal-Szemerédi theorem, *J. Combin. Theory Ser. B* **114** (2015), 187–236.

[16] H.A. Kierstead and A.V. Kostochka, A short proof of the Hajnal-Szemerédi theorem on equitable colouring. *Combin. Probab. Comput.* **17** (2008), no. 2, 265–270.

[17] H.A. Kierstead, A.V. Kostochka and M. Mydlarz and E. Szemerédi, A fast algorithm for equitable coloring, *Combinatorica* **30** (2010), no. 2, 217–224.

[18] J. Komlós, Tiling Turán theorems. *Combinatorica* **20** (2000), no. 2, 203–218.

[19] J. Komlós, G. Sárközy and E. Szemerédi, Blow-up lemma. *Combinatorica* **17** (1997), no. 1, 109–123.

[20] J. Komlós, G. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture. Combinatorics (Prague, 1998). *Discrete Math.* **235** (2001), no. 1–3, 255–269.

[21] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, pp. 295–352, Bolyai Soc. Math. Stud., 2, *János Bolyai Math. Soc., Budapest*, 1996.

[22] D. König, Graphen und Matrizen, *Mat. Lapok* **38** (1931), 116–119.

[23] D. Kühn and D. Osthus, The minimum degree threshold for perfect graph packings. *Combinatorica* **29** (2009), no. 1, 65–107.

[24] A. Lo and K. Markström, A multipartite version of the Hajnal-Szemerédi theorem for graphs and hypergraphs, *Combin. Probab. Comput.* **22** (2013), no. 1, 97–111.

[25] Cs. Magyar and R. Martin, Tripartite version of the Corrádi-Hajnal theorem, *Discrete Math.* **254** (2002), no. 1–3, 289–308.

[26] R. Martin and E. Szemerédi, Quadripartite version of the Hajnal-Szemerédi theorem, *Discrete Math.* **308** (2008), no. 19, 4337–4360.

[27] R. Martin and Y. Zhao, Tiling tripartite graphs with 3-colorable graphs, *Electron. J. Combin.* **16** (2009), no. 1, Research Paper 109, 16pp.

[28] R. Martin and Y. Zhao, Tiling tripartite graphs with 3-colorable graphs: The extreme case. Preprint.

[29] A. Schrijver, *Theory of linear and integer programming.* Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, Ltd., Chichester, 1986. xii+471 pp.

[30] E. Szemerédi, Regular partitions of graphs. *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, pp. 399–401, Colloq. Internat. CNRS, 260, CNRS, Paris, 1978.

[31] Y. Zhao, Bipartite graph tiling, *SIAM J. Discrete Math.* **23** (2009), no. 2, 888–900.