Perfect Matchings and K_4^3-Tilings in Hypergraphs of Large Codegree*

Oleg Pikhurko†

Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
USA
Web: http://www.math.cmu.edu/~pikhurko

Abstract. For a k-graph F, let $t_l(n, m, F)$ be the smallest integer t such that every k-graph G on n vertices in which every l-set of vertices is included in at least t edges contains a collection of vertex-disjoint F-subgraphs covering all but at most m vertices of G. Let K_m^k denote the complete k-graph on m vertices.

The function $t_{k-1}(kn, 0, K_k^k)$ (i.e. when we want to guarantee a perfect matching) has been previously determined by Kühn and Osthus [9] (asymptotically) and by Rödl, Ruciński, and Szemerédi [13] (exactly). Here we obtain asymptotic formulae for some other l. Namely, we prove that for any $k \geq 4$ and $k/2 \leq l \leq k-2$,

$$t_l(kn, 0, K_k^k) = \left(\frac{1}{2} + o(1)\right)\binom{kn}{k-l}.$$

Also, we present various bounds in another special but interesting case: $t_2(n, m, K_4^3)$ with $m = 0$ or $m = o(n)$, that is, when we want to tile (almost) all vertices by copies of K_4^3, the complete 3-graph on 4 vertices.

Key words. Complete 3-graph on 4 vertices, Hajnal-Szemerédi Theorem, hypergraph codegree, hypergraph matching

1. Introduction

A *k-graph* (or a *k-uniform set system*) is a pair $G = (V, E)$, where the *edge set* E is a collection of k-subsets of the *vertex set* V. The *order* of G is $v(G) = |V|$ and the *size* of G is $e(G) = |E|$. In the obvious way we define the notions of isomorphism, subgraph, etc. (Note that we do not restrict the notion of subgraph to induced subgraphs only.) When the vertex set is not important, we may identify a k-graph G with its edge set. Thus, for example, $|G| = e(G)$ denotes the number of edges in G.

Let K_m^k denote the complete k-graph on m vertices. Also, we use the following notation: $[n] = \{1, \ldots, n\}$ and, for a set A, $\binom{A}{k} = \{B : B \subseteq A,\ |B| = k\}$. All logarithms are base $\mathrm{e} = 2.718\ldots$.

† Partially supported by the National Science Foundation, Grant DMS-0457512.

Given k-graphs F and G, an *F-tiling* in G is a collection of vertex-disjoint F-subgraphs of G (i.e. subgraphs isomorphic to F). For the single-edge k-graph K_k^k, a K_k^k-tiling is also called a *K_k^k-matching* or simply a *matching*.

For graphs (that is, 2-graphs) there is a large body of research on finding conditions on the minimal degree in terms of the order that guarantee an (almost) perfect F-tiling. This direction of research was motivated by the seminal paper of Hajnal and Szemerédi [6]. More recently, this problem was considered for general k-graphs, in which case it seems to become more difficult.

Let $G = (V, E)$ be a k-graph and let $0 \le l \le k$. The *neighborhood* of l distinct vertices $x_1, \ldots, x_l \in V$ (which we may interchangeably view as an l-set $X = \{x_1, \ldots, x_l\}$) is

$$N(x_1, \ldots, x_l) = N(X) = \left\{ D \in \binom{V(G) \setminus X}{k-l} : D \cup X \in E \right\},$$

that is, it consists of those $(k-l)$-sets $D \subseteq V(G) \setminus X$ for which $D \cup X$ is an edge of G. The *codegree* of X is $|N(X)|$ and the *l-codegree* of G is

$$\delta_l(G) = \min \left\{ |N(X)| : X \in \binom{V(G)}{l} \right\}.$$

Here is the main definition of this paper. Given a k-graph F and integers l, n, and m such that $0 \le m < n$ and $0 \le l < k$, let $t_l(n, m, F)$ be the smallest integer t such that every k-graph of order n and l-codegree at least t has an F-tiling of size at least $(n-m)/v(F)$ (that is, an F-tiling that covers all but at most m vertices).

Determining $t_l(n, m, F)$ is a very general problem. For example, if $m = n - v(F)$, then we want to guarantee at least one copy of a k-graph F. It is not hard to show that the function $t_1(n, n - v(F), F)$ determines the *Turán density* $\pi(F) = \lim_{n\to\infty} \mathrm{ex}(n, F)/\binom{n}{k}$, the last problem being notoriously hard for virtually all hypergraphs.

Here we are interested in almost perfect F-tilings, that is, the case when $m = o(n)$ (especially the case $m = 0$). Unlike the 2-graph case, there are not many results of this type for $k \ge 3$. Let us very briefly mention some of them, referring the reader to the original papers for all details. Kühn and Osthus [9] showed that, for fixed $k \ge 3$,

$$t_{k-1}(kn, 0, K_k^k) = \frac{1}{2}kn + O(\sqrt{n \log n}) \tag{1}$$

and that $t_{k-1}(n, m, K_k^k) = n/k + o(n)$ when $2k^2 \le m \le o(n)$. Rödl, Ruciński, and Szemerédi [12,13] were able to reduce or completely eliminate the error terms in the above estimates using their 'absorbing' technique. In particular, they determined $t_{k-1}(n, k-1, K_k^k)$ exactly for all $n \ge n_0(k)$. Kühn and Osthus [8] showed, among other things, that $t_2(4n, 0, F) = (1 + o(1))n$, where F is the (unique) 3-graph of order 4 and size 2.

Abassi (unpublished) asked if $t_2(4n, 0, K_4^3) \le 2n$. Czygrinow and Nagle [3, Theorem 2.1] disproved this by showing that $t_2(n, m, K_4^3) \ge 3n/5 - m + o(n)$. The latter paper motivated the author's interest in this area. The 3-graph K_4^3 is a very interesting case; for example, determining the Turán function $\mathrm{ex}(n, K_4^3)$ of K_4^3 is one of the most famous open problems in extremal combinatorics.

Here we present the following lower bounds, in particular improving the above bound of Czygrinow and Nagle [3] for all m.

Proposition 1. *There is a constant $C > 0$ such that for any $0 \le m < n$ we have*

$$t_2(n, m, K_4^3) \ge \frac{5}{8}\, n - \frac{1}{8}\, m - C\sqrt{n \log n}. \tag{2}$$

Also, for any $n = 8q + r$ with integers $q \ge 0$ and $r \in \{0, 4\}$, we have

$$t_2(n, 0, K_4^3) \ge 6q + r - 2 \ge \frac{3}{4}\, n - 2. \tag{3}$$

On the other hand, we have the following upper bounds.

Theorem 1. *For every $n \ge 15$,*

$$t_2(n, 14, K_4^3) \le \left\lceil \frac{2n + \lfloor (n-15)/4 \rfloor - 2}{3} \right\rceil \le \frac{3}{4}\, n - \frac{5}{4}. \tag{4}$$

Also, there is a constant C such that for all integers $n \ge 1$ we have

$$t_2(4n, 0, K_4^3) \le \alpha \times 4n + C\sqrt{n \log n}, \tag{5}$$

where

$$\alpha = \frac{2 + \sqrt{10}}{6} = 0.8603... \,.$$

It is unlikely that (5) is sharp. But, given the absence of any upper bounds on $t_2(4n, 0, K_4^3)$ in the literature, we feel we should present at least *some* non-trivial bound. Hopefully, the quest for improving (5) would lead to new ideas and better bounds.

In fact, the proof of (5) required estimating the function $t_2(4n, 0, K_4^4)$. After the author determined the latter function asymptotically, he realized that the proof extends to some other values of the parameters, giving the following result of independent interest.

Theorem 2. *Let $k \ge 3$ be fixed and let $k/2 \le l \le k-1$. Let n be large. Then*

$$\frac{1}{2}\binom{kn}{k-l} - O(n^{k-l-1}) \;\le\; t_l(kn, 0, K_k^k) \;\le\; \frac{1}{2}\binom{kn}{k-l} + O(n^{k-l-1/2}\sqrt{\log n}). \tag{6}$$

The lower bound in (6) is obtained as follows. Partition $[kn] = A \cup B$ into parts of sizes a and b such that $|a - b| \le 2$ and a is odd. Take as edges all k-tuples that intersect A in an even number of vertices. Since any K_k^k-matching uses an even number of vertices from A, there is no perfect K_k^k-matching.

For any k-graph G, $l \in [k]$, and $A \in \binom{V(G)}{l-1}$ with $|N(A)| = \delta_{l-1}(G)$, we have

$$\delta_{l-1}(G) = |N(A)| = \frac{1}{k-l+1} \sum_{x \in V(G)\setminus A} |N(A \cup \{x\})| \ge \frac{v(G) - l + 1}{k-l+1}\, \delta_l(G). \tag{7}$$

Hence, it is enough to prove the upper bound in (6) for $l = \lceil k/2 \rceil$ only. Theorem 2 includes the upper bound (1) of Kühn and Osthus [9], except the constants in our error terms are larger.

We prove the upper bound of Theorem 2 by modifying the ideas in [9]. It is not surprising that our proof shares many features with that from [9]. In particular, we consider a certain k-partite version of the problem first and then deduce Theorem 2.

For a k-partite k-graph G with vertex parts labeled as $V_1, \dots, V_k \subseteq V(G)$ and an index set $L \subseteq [k]$, let $\delta_L(G)$ be the minimum of $|N(X)|$ over all sets $X \subseteq V(G)$ such that $|X| = |L|$ and X intersects every V_i with $i \in L$ (in other words, $|X \cap V_i|$ is either 1 or 0 depending on whether $i \in L$ or $i \in [k] \setminus L$).

Theorem 3. *Let $k \geq 2$, $l \in [k-1]$, and $L \in \binom{[k]}{l}$ be fixed. Let n be sufficiently large and let*

$$\lambda = \sqrt{257kn\log n}. \tag{8}$$

Let $H = (V, E)$ be a k-uniform k-partite hypergraph with parts $V_1 \cup \cdots \cup V_k = V$ such that $|V_i| = n$ for each $i \in [k]$. If

$$\delta_L(H)n^l + \delta_{[k]\setminus L}(H)n^{k-l} \geq n^k + k\lambda n^{k-1}, \tag{9}$$

then H admits a perfect K_k^k-matching.

Note that in Theorem 3 we look at the codegrees of sequences of vertices of two special types only. The constant factor 1 in front of n^k in the right-hand side of (9) is easily seen to be sharp (see, for example, the construction in [9, Lemma 10]). We do not try to optimize the constants in the error term.

Recently, Aharoni, Georgakopoulos, and Sprüssel [1, Theorem 2] showed that every k-partite k-graph H with parts $V_1, \ldots, V_k$, each of size n, has a perfect K_k^k-matching provided $\delta_{[k-1]}(G) > n/2$ and $\delta_{\{2,\ldots,k\}}(G) \geq n/2$. They also conjectured [1, Conjecture 1] that, for any $l \in [k-1]$ the above assumptions can be weakened to $\delta_{[l]}(G) > n^{k-l}/2$ and $\delta_{\{l+1,\ldots,k\}}(G) \geq n^l/2$. Our Theorem 3 implies an asymptotic version of this conjecture for all large n (with the factor $1/2$ replaced by $1/2 + ((257k^3 \log n)/n)^{1/2}$ in one of the codegree conditions).

Section 5 contains some concluding remarks and open problems.

2. K_4^3-Tilings

Proof of Proposition 1. Let us first prove (2). Given n and m, let $V = [n]$. Partition $[n] = A \cup B$ with $|A| = \lfloor (n-m-1)/4 \rfloor$. Let G consist of all triples intersecting A plus a K_4^3-free 3-graph H on B with

$$\delta_2(H) > \frac{|B|}{2} - C\sqrt{n\log n}. \tag{10}$$

Such a 3-graph H was constructed by Czygrinow and Nagle [3, Proposition 2.3]. Let us briefly recall their construction. We generate a random tournament T on B and let a triple $\{x, y, z\} \in \binom{B}{3}$ with $x < y < z$ in the order induced from $[n]$ belong to H if of the two pairs $\{x, y\}$ and $\{x, z\}$ exactly one is directed toward x in the tournament T. It is not hard to see that H is always K_4^3-free. On the other hand, the codegree of any x, y in H (once, for simplicity, we fix the orientation of $\{x, y\}$) has the binomial distribution with parameters $(|B|-2, 1/2)$. By Chernoff's bounds, the probability that $|N_H(x, y)| \leq |B|/2 - C\sqrt{n\log n}$ is at most $o(|B|^{-2})$. (When there are two different hypergraphs on a common set of vertices, we may use the subscript to clarify in which one we take the neighborhood.) So (10) holds almost surely.

Every K_4^3-subgraph of G must intersect A. Thus, any K_4^3-tiling in G has at most $|A|$ elements and at least $m+1$ vertices remain uncovered. Since all triples intersecting A belong to G, we have $\delta_2(G) = \delta_2(H) + |A|$. By (10), this proves the required bound (2).

The following construction establishes (3). If $r = 4$, let $a_0 = 2q + 1$. Otherwise (if $r = 0$), we let a_0 be either $2q + 1$ or $2q - 1$, with both choices giving the same bound. Partition $[n] = A_0 \cup A_1 \cup A_2 \cup A_3$ into parts of sizes $a_0 + a_1 + a_2 + a_3 = n$, where a_1, a_2, a_3 are nearly equal, that is, $|a_i - a_j| \leq 1$ for $1 \leq i < j \leq 3$. Let G consist of all triples that satisfy one of the following (mutually exclusive) properties:

– have exactly two vertices in A_0;
– intersect each of A_0, A_i, A_j for some $1 \le i < j \le 3$;
– lie inside A_i for some $i \in [3]$;
– have two vertices in A_j and one vertex in A_i for some distinct $i, j \in [3]$.

Since A_0 spans no edge, no K_4^3-subgraph of G can intersect it in more than 2 vertices. Also, no K_4^3-subgraph can have exactly one vertex in A_0. Indeed, otherwise its other three vertices x_1, x_2, x_3 must come from A_1, A_2, A_3, one from each part, which contradicts the fact that such a triple $\{x_1, x_2, x_3\}$ is not an edge of G by definition. Thus every K_4^3 has an even number of vertices in A_0. So, a perfect K_4^3-tiling is impossible because $a_0 = |A_0|$ is odd. Also, the easy case analysis shows that

$$\delta_2(G) = a_0 + a_1 + a_2 + a_3 - \max(a_0, a_1, a_2, a_3) - 2 = 6q + r - 3,$$

where the 2-codegree of G can be achieved by a pair of vertices connecting some two of the parts A_1, A_2, A_3 or lying inside a largest set among A_1, A_2, A_3. The inequality $t_2(n, 0, K_4^3) \ge \delta_2(G) + 1$ gives (3). □

Remark 1. Czygrinow and Nagle [3, Conjecture 3.1] conjectured that the 2-codegree of an arbitrary K_4^3-free 3-graph of order n is at most $(\frac{1}{2} + o(1))n$. If this conjecture is false, then our bound (2) can be improved.

Proof of Theorem 1. We prove (4). Our proof is obtained by adopting the proof of Theorem 2.1 in Fischer [5].

Take an arbitrary 3-graph $G = (V, E)$ of order $n \ge 15$ and 2-codegree $a = \delta_2(G)$ with

$$a > \frac{2n + \lfloor (n-15)/4 \rfloor - 3}{3}. \tag{11}$$

Recall that K_m^3 is the complete 3-graph on m vertices. In particular, K_3^3 denotes the single 3-edge while K_2^3 denotes the 3-graph of order 2 without any edges. For a family $\mathcal{F}$ of 3-graphs, let an *$\mathcal{F}$-tiling* in G be a collection of vertex-disjoint subgraphs of G, each isomorphic to a member of $\mathcal{F}$. Define the *weight factors* to be $w_2 = 2$, $w_3 = 6$, and $w_4 = 11$. Let T be a $\{K_2^3, K_3^3, K_4^3\}$-tiling in G that maximizes the *total weight* $w(T) = w_2 l_2 + w_3 l_3 + w_4 l_4$, where l_i denotes the number of copies of K_i^3 in T. Since an arbitrary pair of vertices forms a K_2^3-subgraph (and $w_2 > 0$), at most one vertex of V is missed by the tiling T.

We claim that the K_4^3-subgraphs of T cover all but at most 14 vertices of V, which proves the theorem. Suppose on the contrary that the claim is false.

Let us say that a 3-graph $F \in T$ and a vertex $x \in V \setminus V(F)$ make a *connection* (denoted by $(F, x) \in \mathcal{C}$) if $v(F) \le 3$ and $V(F) \cup \{x\}$ spans a complete 3-graph. Thus, each connection produces another $\{K_2^3, K_3^3, K_4^3\}$-tiling when we move x to $V(F)$. We are going to explore this fact (and the maximality of T). For example, if a K_i^3-subgraph $F \in T$ with $i \le 3$ makes a connection with some x, then x belongs to a K_j^3-subgraph of T with $j > i$ for otherwise we can strictly increase the weight of T by moving x to F. (Note that $w_4 + w_2 - 2w_3 = 1$, $w_3 - 2w_2 = 2$, and other possible weight changes are all positive.)

Clearly, each K_2^3-subgraph of F makes at least $a = \delta_2(G)$ connections. Also, for each K_3^3-subgraph of F there are at least $3\delta_2(G) - 3 = 3a - 3$ edges that intersect it in exactly two vertices. Hence, if c is the number of connections made by a K_3^3-subgraph, then $3a-3 \le 3c+2(n-3-c)$, i.e.

$$c \ge 3a-2n+3.$$

Suppose first that $l_3 \ge 4$. Let $F_1, \dots, F_4$ be some K_3^3-subgraphs of T. As we have already observed, all connections created by F_i's belong to K_4^3-subgraphs of T. By our assumption, T has at most $\lfloor (n-15)/4 \rfloor$ K_4^3-subgraphs. Since $4(3a-2n+3) > 4\lfloor (n-15)/4 \rfloor$, the vertices of some K_4^3-subgraph $F \in T$ make in total at least 5 connections with $F_1, \dots, F_4$. Consider the bipartite graph B with parts $\{F_1, \dots, F_4\}$ and $V(F)$ whose edge sets consists of those pairs that make a connection. Since B has at least 5 edges, the König-Egerváry Theorem (see e.g. [4, Theorem 2.1.1]) or a direct analysis shows that B has two disjoint edges. This gives us distinct $i, j \in [4]$ and distinct $x, y \in V(F)$ such that $(F_i, x), (F_j, y) \in \mathcal{C}$. By moving x to F_i and y to F_j (thus reducing F to K_2^3), we increase the total weight of T by $2(w_4 - w_3) + (w_2 - w_4) = 1$, which contradicts the maximality of T.

Thus $l_3 \le 3$. Since the number of vertices of G that are not covered by K_4^3-subgraphs of T is at least 15 and at most $1 + 2l_2 + 3l_3$, we conclude that $l_2 \ge 3$. Let F_1, F_2, F_3 be some K_2^3-subgraphs of T. The vertices of no K_3^3-subgraph $F \in T$ make more than 3 connections to F_1, F_2, F_3. Indeed, otherwise (by König-Egerváry) we get distinct $i, j \in [3]$ and distinct $x, y \in V(F)$ such that $(F_i, x), (F_j, y) \in \mathcal{C}$. By moving x to F_i and y to F_j (thus eliminating F from T completely), we increase the total weight of T by $2(w_3-w_2)-w_3 = 2$, a contradiction to the maximality of T. Likewise, the vertices of no K_4^3-subgraph $F \in T$ make more than 8 connections to F_1, F_2, F_3. Indeed, otherwise (again, by König-Egerváry) we find distinct $x_1, x_2, x_3 \in V(F)$ such that $(F_i, x_i) \in \mathcal{C}$ for every $i \in [3]$. Thus by moving each x_i to F_i, we create three new K_3^3-subgraphs, increasing the weight of T by $3(w_3 - w_2) - w_4 = 1$, a contradiction.

Therefore, by estimating from above and below the number of connections created by F_1, F_2, F_3, we obtain that $3a \le 3l_3 + 8l_4$. Since T has at most $\lfloor (n-15)/4 \rfloor$ K_4^3-subgraphs and $l_3 \le 3$, we have $3l_3 + 8l_4 \le 9 + 8\lfloor (n-15)/4 \rfloor$. Thus we have

$$3a \le 9 + 8\lfloor (n-15)/4 \rfloor < \lfloor (n-15)/4 \rfloor + 2n - 3,$$

where it is routine to check that the second inequality holds for all $n \ge 15$. This contradicts (11) and proves (4).

Let us turn to (5). Fix large C and then let n be sufficiently large. Given a 3-graph G on $[4n]$ such that

$$a = \delta_2(G) \ge \alpha \times 4n + C\sqrt{n \log n},$$

let us form the 4-graph H on $[4n]$, where a 4-set A is an edge of H if and only if $G[A]$ is isomorphic to K_4^3.

It is enough to show that H admits a perfect K_4^4-matching. Take any distinct $x, y \in [4n]$. Let A be a set of exactly a vertices that form an edge of G with $\{x, y\}$. (Such a set A exists since $|N_G(x,y)| \ge a$.) Moreover, the neighborhood $N_G(x)$, when restricted to $\binom{A}{2}$ has minimum degree (as a 2-graph) at least $2a - 4n + 1$. The same applies to $N_G(y) \cap \binom{A}{2}$. Hence, the neighborhood $N_H(x,y)$ in the 4-graph H has size at least

$$\left| N_G(x) \cap N_G(y) \cap \binom{A}{2} \right| \ge 2 \times \frac{(2a-4n+1)a}{2} - \binom{a}{2} > \frac{1}{2}\binom{4n}{2} + C' n^{3/2}\sqrt{\log n}$$

for some C' that tends to infinity with C (in fact, we can take $C' = (2\sqrt{10} - o(1))C$).

Theorem 2 implies that H has a perfect matching, which gives the desired K_4^3-tiling in G. This finishes the proof of Theorem 1. □

Remark 2. It is easy to convert the proof of (4) into a polynomial-time algorithm that produces a K_4^3-tiling guaranteed by Theorem 1: just start with an arbitrary collection of $\lfloor n/2 \rfloor$ vertex-disjoint K_2^3-subgraphs and keep updating it (as stipulated in the proof) until at most 14 vertices remain uncovered by K_4^3-subgraphs. Since each update increases the total weight $w(T)$ at least by 1, the total number of updates is at most $w_4 n/4 - w_2 \lfloor n/2 \rfloor \leq 9n/4 + 1$.

The proof of (5) (see Section 4) gives a randomized polynomial time algorithm that succeeds with high probability.

3. Auxiliary Lemmas

The purpose of this section is to provide a few auxiliary results needed for the proofs of Theorems 2 and 3. First, we will use a special case of a more general result of McDiarmid [11, Theorem 1.1] (see also Talagrand [14, Theorem 5.1]).

Lemma 1. *Let c, r be positive reals. Let $Z : S_n \to \mathbb{R}_{\geq 0}$ be a function from permutations on $[n]$ to non-negative reals such that*

- *if $\sigma, \sigma' \in S_n$ differ only in two places then $|Z(\sigma) - Z(\sigma')| \leq c$;*
- *if $Z(\sigma) = s$ then there is a subset $W \subseteq [n]$ of size at most rs such that every $\sigma' \in S_n$ that coincides with σ on W satisfies $Z(\sigma') \geq s$.*

Let σ be a permutation from S_n, drawn uniformly at random, and let m be a median of the random variable $Z(\sigma)$. Then, for each $h \geq 0$, the probability of $Z(\sigma) \leq m - h$ satisfies the following inequality:

$$\Pr\{Z(\sigma) \leq m - h\} \leq 2 \exp\left(-\frac{h^2}{16rc^2m}\right). \quad \square$$

The above lemma allows us to deduce the following corollary rather easily.

Corollary 1. *Let G be an arbitrary subgraph of $K_{n,n}$ and let h satisfy*

$$h > 4np_0, \tag{12}$$

where we set $p_0 = 4\exp(-h^2/(256n))$. Let M be a perfect matching of $K_{n,n}$, chosen uniformly at random. Then

$$\Pr\left\{ \left| |M \cap E(G)| - \frac{e(G)}{n} \right| \geq h \right\} \leq p_0. \tag{13}$$

Proof. Let $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ be the parts of $K_{n,n}$. We generate M by taking a random permutation $\sigma \in S_n$ and letting

$$M = \{\, \{x_i, y_{\sigma(i)}\} : i \in [n] \,\}.$$

Let $Z(\sigma) = |M \cap E(G)|$ and let m be a median of $Z(\sigma)$. Clearly, the assumptions of Lemma 1 are satisfied with $c = 2$ and $r = 1$. In view of $m \leq n$, we have $\Pr\{Z(\sigma) \leq m - h/2\} \leq p_0/2$. The random variable $Z'(\sigma) = n - Z(\sigma)$ counts $|M \cap E(\overline{G})|$ so it also satisfies the assumptions of Lemma 1 with $c = 2$ and $r = 1$. By applying Lemma 1 to $Z'(\sigma)$ (with respect to a median $m' = n - m$) we obtain that

$$\Pr\{Z'(\sigma) \le m' - h/2\} = \Pr\{Z(\sigma) \ge m + h/2\} \le \frac{1}{2}\,p_0.$$

Thus the probability that $Z(\sigma) \notin I$ is at most p_0, where

$$I = \{x \in \mathbb{R} : |x - m| < h/2\}.$$

We claim that the expectation

$$\mu = \mathrm{E}(Z(\sigma)) = \frac{e(G)}{n}$$

lies in I. Suppose on the contrary that this is not true, say $\mu \ge m + h/2$. Consider the expectation of $X = \mu - Z(\sigma)$, which is of course 0. Those σ with $Z(\sigma) \le m$ contribute at least $(1/2) \times (h/2)$ to $\mathrm{E}(X)$. The contribution from σ with $m < Z(\sigma) \le m + h/2$ is non-negative. The remaining σ contribute at least $-p_0 n$. However, this contradicts the assumption (12) and the fact that $\mathrm{E}(X) = 0$.

It follows from $\mu \in I$ that if $|Z(\sigma) - \mu| \ge h$ then $Z(\sigma) \notin I$. As we have already demonstrated, the latter event has probability at most p_0. The corollary is proved. $\square$

Remark 3. Kühn and Osthus [9, Lemma 8] proved a result analogous to Corollary 1 but their proof (based on Brégman's theorem [2] on permanents) crucially relies on the assumption that the minimum degree of G on one side is larger than $n/2$. Also, another result of Kühn and Osthus [10, Theorem 1.1] implies a weaker version of (13) (namely, that for any positive constant c there is $\varepsilon = \varepsilon(c) > 0$ such that (13) holds when $n > n_0(c)$ $h = cn$, and $p_0 = \mathrm{e}^{-\varepsilon n}$). Since our proof is short and gives better error terms (namely, $h = O(\sqrt{n \log n})$ in our applications), we included it here.

Also, we will make use of the following result [9, Proposition 13] proved by the standard probabilistic tools.

Proposition 2. *For each integer $k \ge 2$ there exists an integer n_0 such that the following holds. Suppose $n \ge n_0$ and that H is a k-uniform hypergraph with kn vertices. Then there exists a partition $V_1, \dots, V_k$ of $V(H)$ into sets of size n such that for every $i \in [k]$ and every distinct $x_1, \dots, x_{k-1} \in V(H)$ we have*

$$\left| |N(x_1, \dots, x_{k-1}) \cap V_i| - \frac{|N(x_1, \dots, x_{k-1})|}{k} \right| \le 2k\sqrt{n \log n}. \quad \square$$

Let the symbol K denote the (generic) complete hypergraph and let $K[V_1, \dots, V_k]$ consist of all k-sets A such that $|A \cap V_i| = 1$ for each $i \in [k]$.

Corollary 2. *For any $k \ge 2$ there exists n_0 such that the following holds. Suppose $n \ge n_0$ and that H is a k-uniform hypergraph with kn vertices. Then there exists a partition $V_1, \dots, V_k$ of $V(H)$ into sets of size n such that for every $l \in [k-1]$, every l-set $X \subseteq V(H)$, and any distinct $i_1, \dots, i_m \in [k]$ we have*

$$\left| \frac{k^m}{m!}\, |N(X) \cap K[V_{i_1}, \dots, V_{i_m}]| - |N(X)| \right| \le 4k^{m+1} n^{m-1/2} \sqrt{\log n}, \tag{14}$$

where $m = k - l$.

Proof. Let n be sufficiently large so that Proposition 2 applies. Given $H = (V, E)$, take the partition $V(H) = V_1 \cup \cdots \cup V_k$ given by Proposition 2. Let us show that it has the required properties. Let $l \in [k-1]$. Take any l-set $X \subseteq V$. Let $U_i = V_i \setminus X$ and $U = V \setminus X$. Let $F = (U, N(X))$ be the m-graph on U having $N(X)$ for the edge set. Proposition 2 implies that for any $(m-1)$-set $Y \subseteq U$ and $i, j \in [k]$

$$-\lambda \le |N_F(Y) \cap U_i| - |N_F(Y) \cap U_j| \le \lambda, \tag{15}$$

where $\lambda = 4k\sqrt{n \log n}$.

Let the *distribution* of a set $A \subseteq U$ be the vector

$$\mathbf{d}(A) = (\,|A \cap U_1|, \ldots, |A \cap U_k|\,).$$

Let $\mathcal{D}$ consist of all k-vectors of non-negative integers with sum m. The *weight* of a vector $\mathbf{d} = (d_1, \ldots, d_k) \in \mathcal{D}$ is

$$w(\mathbf{d}) = \frac{m!}{d_1! \times \cdots \times d_k! \times k^m}.$$

This is exactly the limit of the probability that a random m-set has the distribution $\mathbf{d}$, when the part sizes are equal and tend to infinity. In particular,

$$\sum_{\mathbf{d} \in \mathcal{D}} w(\mathbf{d}) = 1. \tag{16}$$

For $\mathbf{d} \in \mathcal{D}$, let $F^{\mathbf{d}}$ consist of the m-sets in $E(F)$ whose distribution is $\mathbf{d}$.

Suppose that $\mathbf{d} = (d_1, \ldots, d_k)$ and $\mathbf{d}' = (d_1', \ldots, d_k')$, $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$, are at ℓ_1-distance 2 (that is, $\sum_{i=1}^k |d_i - d_i'| = 2$). Let $i, j \in [k]$ satisfy $d_i = d_i' + 1$ and $d_j = d_j' - 1$. Let us sum (15) over all sets $Y \in \binom{U}{m-1}$ with distribution

$$(\min(d_1, d_1'), \ldots, \min(d_k, d_k')) = (d_1, \ldots, d_{i-1}, d_i - 1, d_{i+1}, \ldots, d_k).$$

Since there are at most n^{m-1} choices of Y, we obtain $-\lambda n^{m-1} \le d_i |F^{\mathbf{d}}| - d_j' |F^{\mathbf{d}'}| \le \lambda n^{m-1}$. This implies that

$$\left| \frac{|F^{\mathbf{d}}|}{w(\mathbf{d})} - \frac{|F^{\mathbf{d}'}|}{w(\mathbf{d}')} \right| \le \lambda n^{m-1} \times \frac{k^m \prod_{h=1}^k \min(d_h!, d_h'!)}{m!} \le \lambda n^{m-1} \times \frac{k^m}{m}. \tag{17}$$

Now, every two distributions from $\mathcal{D}$ are at ℓ_1-distance at most $2m$ apart. It follows from (17) (and the Triangle Inequality) that for arbitrary $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$ we have

$$\left| \frac{|F^{\mathbf{d}}|}{w(\mathbf{d})} - \frac{|F^{\mathbf{d}'}|}{w(\mathbf{d}')} \right| \le \lambda n^{m-1} k^m. \tag{18}$$

Given distinct $i_1, \ldots, i_m \in [k]$ as the input to the corollary, define the distribution $\mathbf{d} \in \mathcal{D}$ by $d_{i_1} = \cdots = d_{i_m} = 1$ with all other d_i's being 0. Now, (16) and (18) imply that

$$|F| = \sum_{\mathbf{d}' \in \mathcal{D}} |F^{\mathbf{d}'}| \le \left(\frac{|F^{\mathbf{d}}|}{w(\mathbf{d})} + \lambda n^{m-1} k^m \right) \sum_{\mathbf{d}' \in \mathcal{D}} w(\mathbf{d}') = \frac{|F^{\mathbf{d}}|}{w(\mathbf{d})} + \lambda n^{m-1} k^m.$$

This proves one direction in (14). The other direction is proved analogously. $\square$

4. Perfect Matchings in k-Partite k-Graphs

Proof of Theorem 3. At the first reading, the reader may find it useful to run the proof in the special case $k = 3$ and $L = \{1, 2\}$, which illustrates the main idea well.

Assume without loss of generality that $L = [l]$. We will inductively produce K_2^2-matchings $M_1, \ldots, M_{l-1}$ where M_i is a perfect matching in the bipartite graph $K[V_i, V_{i+1}]$. We need to introduce some notation first.

Given $i \in [l-1]$ and $M_1, \ldots, M_i$, let T_i be the perfect K_i^i-matching in $K[V_1, \ldots, V_i]$ obtained by 'gluing' the edges of $M_1, \ldots, M_{i-1}$ together. Formally, T_1 is the perfect K_1^1-matching of V_1, which consists of singletons while, for $i \geq 2$,

$$T_i = \{\, \{x_1, \ldots, x_i\} : \forall j \in [i-1]\ \{x_j, x_{j+1}\} \in M_j \,\}.$$

Also, for a set $X \in K[V_{i+2}, \ldots, V_k]$, let $F_X \subseteq K[V_i, V_{i+1}]$ consist of those pairs $\{x_i, x_{i+1}\}$ with $x_i \in V_i$ and $x_{i+1} \in V_{i+1}$ such that $D \cup \{x_{i+1}\} \cup X \in E(H)$, where D is the element of the perfect K_i^i-matching T_i that contains x_i.

We require that the following *Property* $\mathcal{P}_i$ holds for every $i \in [l-1]$:

$$\left|\, |M_i \cap F_X| - \frac{|F_X|}{n} \,\right| \leq \lambda, \qquad \text{for every } X \in K[V_{i+2}, \ldots, V_k], \tag{19}$$

where λ is defined by (8).

If $l = 1$, then we just define T_1 to be the perfect K_1^1-matching of V_1 and there is nothing else to do.

Suppose that $l \geq 2$, $i \in [l-1]$, and we have already constructed $M_1, \ldots, M_{i-1}$ (and thus we also have the matching T_i). Let M_i be a random perfect matching of $K[V_i, V_{i+1}]$. For any fixed $X \in K[V_{i+2}, \ldots, V_k]$, the probability that (19) is false is at most

$$p_0 = 4 \exp\left(-\frac{\lambda^2}{256n}\right) = o(n^{-k}), \tag{20}$$

by Corollary 1 applied to F_X with $h = \lambda$. (Note that (12) is satisfied by (8) and (20).) Since there are $n^{k-i-1} \leq n^k$ choices of X, the union bound implies that with probability $1 - o(1)$ M_i satisfies Property $\mathcal{P}_i$. Fix any such M_i and, if $i < l-1$, repeat the argument for $i+1$.

Thus the required matchings $M_1, \ldots, M_{l-1}$ exist.

Now, let us prove that for any $i \in [l-1]$ and any $X \in K[V_{i+2}, \ldots, V_k]$, we have

$$\left|\, |N(X) \cap T_{i+1}| - \frac{|N(X)|}{n^i} \,\right| \leq i\lambda. \tag{21}$$

We prove this by induction on i, the case $i = 1$ being exactly Property $\mathcal{P}_1$. Suppose that $2 \leq i \leq l-1$ and that (21) holds for $i-1$. Let $X \in K[V_{i+2}, \ldots, V_k]$. We have, by Property $\mathcal{P}_i$ and the induction assumption, that

$$\begin{aligned} |N(X) \cap T_{i+1}| = |F_X \cap M_i| \ &\leq\ \lambda + \frac{1}{n}\,|F_X| \\ &= \lambda + \frac{1}{n} \sum_{x \in V_{i+1}} |N(X \cup \{x\}) \cap T_i| \\ &\leq \lambda + \frac{1}{n} \sum_{x \in V_{i+1}} \left(\frac{|N(X \cup \{x\})|}{n^{i-1}} + (i-1)\lambda \right) \ =\ i\lambda + \frac{|N(X)|}{n^i}. \end{aligned}$$

This proves one direction of (21). The other direction is proved in the same way. Hence, (21) holds for all $i \in [l-1]$.

We will need only the case $i = l-1$ of (21) which implies that for every $X \in K[V_{l+1}, \dots, V_k]$ we have

$$|N(X) \cap T_l| \geq \frac{|N(X)|}{n^{l-1}} - (l-1)\lambda \geq \frac{\delta_{[k]\setminus L}(H)}{n^{l-1}} - (l-1)\lambda. \tag{22}$$

Similarly, we build matchings from the other end. Namely, inductively for $j = k, k-1, \dots, l+2$ we consider a random matching $M_j' \subseteq K[V_{j-1}, V_j]$ and take a 'typical' one. Similarly to above, we show that there is a K_{k-l}^{k-l}-matching $T_{k-l}' \subseteq K[V_{l+1}, \dots, V_k]$ such that any l-set $Y \in K[V_1, \dots, V_l]$ forms an H-edge with at least

$$\frac{\delta_L(H)}{n^{k-l-1}} - (k-l-1)\lambda \tag{23}$$

edges of T_{k-l}'.

Finally, we consider the bipartite graph B with parts $U_1 = T_l$ and $U_2 = T_{k-l}'$, where we connect $X \in T_l$ to $Y \in T_{k-l}'$ if and only if $X \cup Y \in E(H)$. By (22), (23), and our assumption (9) we have

$$\delta_{\{1\}}(B) + \delta_{\{2\}}(B) \geq \frac{\delta_{[k]\setminus L}(H)}{n^{l-1}} - (l-1)\lambda + \frac{\delta_L(H)}{n^{k-l-1}} - (k-l-1)\lambda \geq n,$$

where e.g. $\delta_{\{1\}}(B)$ is the smallest B-degree of a vertex from U_1. Let us check that the bipartite graph B satisfies Hall's condition. Let X be an arbitrary non-empty subset of U_1 and let $\Gamma(X)$ consist of all vertices in U_2 that send at least one edge to X. If $|X| \leq \delta_{\{1\}}(B)$, then $|\Gamma(X)| \geq \delta_{\{1\}}(B) \geq |X|$, as desired. Otherwise, every $x \in U_2$ is connected to X (since $|U_1 \setminus X| < n - \delta_{\{1\}}(B) \leq \delta_{\{2\}}(B)$) and $|\Gamma(X)| = n \geq |X|$. Thus, by Hall's Marriage Theorem, B contains a perfect matching which, in turn, gives a perfect K_k^k-matching in H, proving Theorem 3. □

Proof of Theorem 2. We have already established the lower bound in (6). Also, by (7), it suffices to prove the upper bound in (6) for $l = \lceil k/2 \rceil$ only. Given $k \geq 3$, let $l = \lceil k/2 \rceil$ and let $C = C(k)$ be sufficiently large. Let n be arbitrary and let G be a k-graph on kn vertices with

$$\delta_l(G) \geq \frac{1}{2}\binom{kn}{k-l} + Cn^{k-l-1/2}\sqrt{\log n}. \tag{24}$$

(Since $C(k)$ is large, n is also forced to be large with respect to k.)

Apply Corollary 2 to G obtaining a balanced k-partite k-graph H with partition $V(H) = V_1 \cup \dots \cup V_k$. Let $L = [l]$ and $m = k-l$. By (14) and (24), we have

$$\delta_L(H) \geq \frac{m!}{k^m}\left(\delta_l(G) - 4k^{m+1}n^{m-1/2}\sqrt{\log n}\right) > \frac{n^{k-l}}{2} + k\lambda n^{k-l-1},$$

where λ is defined by (8). Likewise, $\delta_{[k]\setminus L}(H) \geq n^l/2 + k\lambda n^{l-1}$. Hence, the assumptions of Theorem 3 are satisfied, which implies that H has a perfect K_k^k-matching. This gives the required K_k^k-matching in $G \supseteq H$. □

5. Concluding Remarks

Many problems remain open. One is to close the gaps in our bounds on (almost) perfect K_4^3-tilings.

It is possible that there is an '$\Omega(n)$-jump' in $t_2(n,m,K_4^3)$ when m goes from 0 to 14 or that the ratio $t_2(n,3,K_4^3)/n$ tends to different limits as $n\to\infty$, depending on the residue of n modulo 4. Such phenomena do occur for $t_{k-1}(n,m,K_k^k)$ with $k\ge 3$, see [9,12,13]. Unfortunately, our bounds are not strong enough to show or refute this.

An interesting open problem on perfect matchings is to determine the *exact* value of $t_l(kn,0,K_k^k)$ for $k/2\le l\le k-2$ and all large n. Also, it is not clear what happens for the remaining values of l. Unfortunately, our method breaks down when $l<k/2$. The simplest open case is $t_1(3n,0,K_3^3)$. The author conjectured in the first version of the paper that $t_1(3n,0,K_3^3)=(\frac{1}{2}+o(1))\binom{3n}{2}$. This was disproved by Hán, Person, and Schacht [7] who showed that $t_1(3n,0,K_3^3)=(\frac{5}{9}+o(1))\binom{3n}{2}$. The lower bound is demonstrated by the following simple construction: partition $[3n]=A\cup B$ with $|A|=n-1$ and take all triples that intersect A (see also Aharoni, Georgakopoulos, and Sprüssel [1, Section 3] where a 3-partite version of this construction appears).

Hopefully, this paper (and the open problems stated here) will generate more interest and work in this difficult but fascinating area.

Note added in proof: After this paper was written, the author learned that Peter Keevash and Benny Sudakov had observed that $t_2(n,0,K_4^3)\ge(5/8-o(1))n$ several years ago, and Peter Keevash and Yi Zhao also proved (unpublished) that $2n/3-1\le t_2(n,0,K_4^3)\le(\frac{2+\sqrt{10}}{6}+o(1))n$, where n is a large integer divisible by 4.

Acknowledgements. The author is grateful to Anusch Taraz and Helmut Veith for organizing his visit to the Technical University of Munich and for their hospitality. Also, the author thanks the anonymous referees for the very careful reading of the manuscript and many useful remarks, in particular for drawing the author's attention to the related results [1,10].

References

1. R. Aharoni, A. Georgakopoulos, and P. Sprüssel. Perfect matchings in r-partite r-graphs. To appear in Europ. J. Combin, 2008.
2. L. M. Brègman. Certain properties of nonnegative matrices and their permanents. *Dokl. Akad. Nauk SSSR*, 211:27–30, 1973.
3. A. Czygrinow and B. Nagle. A note on codegree problems for hypergraphs. *Bull. ICA*, 32:63–69, 2001.
4. R. Diestel. *Graph Theory.* Springer, Berlin, 3d edition, 2006.
5. E. Fischer. Variants of the Hajnal-Szemerédi Theorem. *J. Graph Theory*, 31:275–282, 1999.
6. A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970.
7. H. Hàn, Y. Person, and M. Schacht. On perfect matchings in uniform hypergraphs with large minimum vertex degree. In preparation, 2008.

8. D. Kühn and D. Osthus. Loose Hamilton cycles in 3-uniform hypergraphs of large minimum degree. *J. Combin. Theory* (B), 96:767–821, 2006.

9. D. Kühn and D. Osthus. Matchings in hypergraphs of large minimum degree. *J. Graph Theory*, 51:269–280, 2006.

10. D. Kühn and D. Osthus. Multicolored Hamilton cycles and perfect matchings in pseudo-random graphs. *SIAM J. Discr. Math.*, 20:273–286, 2006.

11. C. McDiarmid. Concentration for independent permutations. *Combin. Prob. Computing*, 11:163–178, 2002.

12. V. Rödl, A. Ruciński, and E. Szemerédi. Perfect matchings in uniform hypergraphs with large minimum degree. *Europ. J. Combin.*, 27:1333–1349, 2006.

13. V. Rödl, A. Ruciński, and E. Szemerédi. Perfect matchings in large uniform hypergraphs with large minimum collective degree. Submitted, 2007.

14. M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, 81:73–205, 1995.

Received: ????
Final version received: ????