Matchings in hypergraphs of large minimum degree

Daniela Kühn Deryk Osthus

Abstract

It is well known that every bipartite graph with vertex classes of size n whose minimum degree is at least $n/2$ contains a perfect matching. We prove an analogue of this result for hypergraphs. We also prove several related results which guarantee the existence of almost perfect matchings in r-uniform hypergraphs of large minimum degree. Our bounds on the minimum degree are essentially best possible.

1 Introduction

The so called 'marriage theorem' of Hall provides a necessary and sufficient condition for the existence of a perfect matching in a bipartite graph. For hypergraphs there is no analogue of this result—up to now only partial results are known. For example, Conforti et al. [4] extended Hall's theorem to balanced hypergraphs and Haxell [8] extended Hall's theorem to a sufficient condition for the existence of a hypergraph matching which contains a given set of vertices. Moreover, there are many results about the existence of almost perfect matchings in hypergraphs which are pseudo-random in some sense. Most of these are based on an approach due to Rödl (see e.g. [1] for an introduction to the topic or Vu [15] for more recent results). For random r-uniform hypergraphs, the threshold for a perfect matching is still not known. There are several partial results, see e.g. Kim [11].

A simple corollary of Hall's theorem for graphs states that every bipartite graph with vertex classes A and B of size n whose minimum degree is at least $n/2$ contains a perfect matching. (This can also be proved directly by considering a maximal matching.) The condition on the minimum degree is best possible. The first theorem of this paper provides an analogue of this result for r-uniform r-partite hypergraphs. So instead of two vertex classes and a set of edges joining them (as in the graph case), we now have r vertex classes and a set of (unordered) r-tuples, each of whose vertices lies in a different vertex class (see Section 2 for the precise definition). A natural way to define the minimum degree of an r-uniform r-partite hypergraph $\mathcal{H}$ is the following. Given $r-1$ distinct vertices $x_1, \dots, x_{r-1}$ of $\mathcal{H}$, the *neighbourhood* $N_{r-1}(x_1, \dots, x_{r-1})$ *of* $x_1, \dots, x_{r-1}$ in $\mathcal{H}$ is the set of all those vertices x which form a hyperedge together with $x_1, \dots, x_{r-1}$. The *minimum degree* $\delta'_{r-1}(\mathcal{H})$ is defined to be the minimum $|N_{r-1}(x_1, \dots, x_{r-1})|$ over all tuples $x_1, \dots, x_{r-1}$ of vertices lying in different vertex classes of $\mathcal{H}$.

Theorem 1 *Suppose that $\mathcal{H}$ is an r-uniform r-partite hypergraph with vertex classes of size $n \geq 1000$ which satisfies $\delta'_{r-1}(\mathcal{H}) \geq n/2 + \sqrt{2n \log n}$. Then $\mathcal{H}$ has a perfect matching.*

Theorem 1 is best possible up to the error term $\sqrt{2n \log n}$ (see Lemma 10). Surprisingly, a simple argument already shows that a significantly smaller minimum degree guarantees a matching which covers *almost all* vertices of $\mathcal{H}$:

Theorem 2 *Suppose that $\mathcal{H}$ is an r-uniform r-partite hypergraph with vertex classes of size n which satisfies $\delta'_{r-1}(\mathcal{H}) \geq n/r$. Then $\mathcal{H}$ has a matching which covers all but at most $r-2$ vertices in each vertex class of $\mathcal{H}$.*

Again, the bound on the minimum degree in Theorem 2 is essentially best possible: if we reduce it by εn, then we cannot even guarantee a matching which covers all but εn vertices in each vertex class (see Lemma 12 for a more precise assertion). Actually, instead of Theorem 2, we will prove a more general statement (Theorem 11).

Finally, we use Theorems 1 and 2 to obtain analogues for r-uniform hypergraphs $\mathcal{H}$ which are not necessarily r-partite. The minimum degree $\delta_{r-1}(\mathcal{H})$ of such a hypergraph $\mathcal{H}$ is defined similarly as before except that we now take the minimum $|N_{r-1}(x_1, \dots, x_{r-1})|$ over *all* $(r-1)$-tuples of distinct vertices of $\mathcal{H}$.

Theorem 3 *For every integer $r \geq 3$ there exists an integer $n_0 = n_0(r)$ such that for every $n \geq n_0$ the following holds. Suppose that $\mathcal{H}$ is an r-uniform hypergraph with $|\mathcal{H}| = rn$ vertices which satisfies $\delta_{r-1}(\mathcal{H}) \geq |\mathcal{H}|/2 + 3r^2\sqrt{n \log n}$. Then $\mathcal{H}$ has a perfect matching.*

Theorem 4 *For every integer $r \geq 3$ there exists an integer $n_0 = n_0(r)$ such that for every $n \geq n_0$ the following holds. Suppose that $\mathcal{H}$ is an r-uniform hypergraph with $|\mathcal{H}| = n$ vertices which satisfies $\delta_{r-1}(\mathcal{H}) \geq n/r + 3r^2\sqrt{n \log n}$. Then $\mathcal{H}$ has a matching which covers all but at most $2r^2$ vertices of $\mathcal{H}$.*

Again the bounds on the minimum degree are best possible up to lower order terms (see Lemmas 15 and 17).

Recently, Rödl, Ruciński and Szemerédi [13] proved the related result that for all positive ε there exists an integer n_0 such that for all $n \geq n_0$ every 3-uniform hypergraph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \geq n/2 + \varepsilon n$ contains a tight Hamilton cycle. (A *tight Hamilton cycle* in a 3-uniform hypergraph $\mathcal{H}$ is a cyclic ordering of its vertices such that every 3 consecutive vertices form a hyperedge.) An immediate corollary is that if n is divisible by 3, then $\mathcal{H}$ contains a perfect matching. We believe that the main advantage of Theorem 3 is that it provides a much shorter proof of this corollary. Also, a related result was proved and used as a tool in [13]: Let $\mathcal{H}$ be a 3-uniform hypergraph on n vertices and consider the auxiliary graph G whose edges are all the pairs x, y of vertices of $\mathcal{H}$ with $|N_2(x, y)| < n/2$. If the maximum degree of G is small, then $\mathcal{H}$ has an almost perfect matching. The argument used in the proof of Theorem 2 gives a simple proof of this fact and reduces the necessary minimum degree from $n/2$ to $n/3$ in the 3-uniform case (see Theorems 11 and 16).

From an algorithmic point of view, there is also a major difference between matching problems for graphs and hypergraphs. A largest matching in a graph can be found in polynomial time (see e.g. [12]), whereas Karp proved that one cannot find a maximum matching in an r-partite r-uniform hypergraph in polynomial time unless P=NP (see e.g. [6], this is also known as the r-dimensional matching problem). In fact, Kann [10] proved that the problem is even MaxSNP-complete, so unless P=NP it is not even possible to approximate the optimal solution within a given factor $1 + \varepsilon$ in polynomial time. The best known approximation algorithm is due to Hurkens and Schrijver [7] and has approximation ratio $r/2 + \varepsilon$ for any given $\varepsilon > 0$.

On the other hand, our proofs can easily be reformulated as polynomial time algorithms (which are randomized in most cases). For example, if the minimum degree is a little larger than $|\mathcal{H}|/2$, there is a randomized polynomial time algorithm which finds a perfect matching with high probability. It would be interesting to know whether this can be achieved without randomization. This is certainly the case for the argument we use to obtain Theorem 4 from Theorem 2—it can easily be derandomized using standard techniques (see e.g. [1, Chapter 15]). Hence an almost perfect matching can be found in polynomial time if the minimum degree is at least $|\mathcal{H}|/r$.

Finally, we present some open questions which are immediately suggested by the above results: Obviously, it would be desirable to eliminate the gaps between the upper and the lower bounds on the minimum degree. Also, it would be interesting to know whether one can obtain similar results if one adopts the following alternative definition of minimum degree: The neighbourhood of a vertex x of an r-uniform hypergraph is the set of all those $(r-1)$-tuples of vertices which form a hyperedge together with x and the minimum degree is the size of the smallest neighbourhood.

This paper is organized as follows. In Section 2 we introduce some notation and collect tools which are needed in the proof of Theorem 1. In Section 3 we consider minimum degree conditions for the existence of perfect matchings in r-partite r-uniform hypergraphs. In Section 4 we consider minimum degree conditions for almost perfect matchings in such hypergraphs. In the final section we then derive the corresponding results for r-uniform hypergraphs which are not necessarily r-partite by using a simple probabilistic argument.

2 Notation and tools

In this paper, all logarithms are base e, where e denotes the Euler number. We write $|G|$ for the number of vertices in a graph G. We denote the degree of a vertex $x \in G$ by $d_G(x)$ and the set of its neighbours by $N_G(x)$. We often write $G = (A, B)$ for a bipartite graph G with vertex classes A and B.

A *hypergraph* $\mathcal{H}$ consists of a set V of *vertices* together with some set E of subsets of V. The elements of E are called the *hyperedges of* $\mathcal{H}$. We write $|\mathcal{H}|$ for the number of vertices of $\mathcal{H}$. $\mathcal{H}$ is *r-uniform* if all its hyperedges are r-element sets. If $\mathcal{H}$ is an r-uniform hypergraph, we will also refer to its hyperedges as *r-tuples of* $\mathcal{H}$. An r-uniform hypergraph $\mathcal{H}$ is called *r-partite* if the vertex set of $\mathcal{H}$ can be partitioned into r classes, $V_1, \dots, V_r$ say, such that every hyperedge meets every V_i in precisely one vertex. The V_i are the *vertex classes* of $\mathcal{H}$. A *matching* in $\mathcal{H}$ is a set M of disjoint hyperedges of $\mathcal{H}$. M is *perfect* if every vertex of $\mathcal{H}$ lies in some hyperedge belonging to M.

We use the following version of Stirling's inequality (the bound is a weak form of a result of Robbins, see e.g. [2]):

Proposition 5 *For all integers $n \geq 1$ we have*

$$\left(\frac{n}{\mathrm{e}}\right)^n \leq n! \leq 3\sqrt{n}\left(\frac{n}{\mathrm{e}}\right)^n. \tag{1}$$

Moreover, we need the following result of Brégman [3] about the permanent of a 0-1 matrix which was conjectured by Minc. (A short proof of it was given by Schrijver [14], see also [1]). We formulate this result in terms of the number of perfect matchings of a bipartite graph.

Theorem 6 *Every bipartite graph $G = (A, B)$ contains at most*

$$\prod_{a \in A} (d_G(a)!)^{1/d_G(a)}$$

perfect matchings.

An application of Stirling's inequality (Proposition 5) to Theorem 6 immediately yields the following.

Corollary 7 *Let $G = (A, B)$ be a bipartite graph with $|A| = |B| = n$ and $d_G(a) \geq n/3$ for every $a \in A$. Let m denote the number of perfect matchings in G. Then*

$$m \leq 27n^{3/2} \prod_{a \in A} \frac{d_G(a)}{\mathrm{e}}.$$

The following lemma will be used in the proof of Theorem 1.

Lemma 8 *Suppose that $G = (A, B)$ is a bipartite graph with $|A| = |B| = n \geq 1000$ and such that $d_G(a) \geq n/2 + \sqrt{2n \log n}$ for all $a \in A$. Let M be a perfect matching in the complete bipartite graph with vertex classes A and B chosen uniformly at random. Then the probability that M contains at most $n/2$ edges of G is at most $1/(2n)$.*

Proof. Set $y := \lceil n/2 + \sqrt{2n \log n}\rceil - n/2$ and choose any spanning subgraph H of G such that $d_H(a) = n/2 + y$ for all $a \in A$. Given a set $A' \subseteq A$, we denote by $H_{A'}$ the bipartite graph with vertex classes A and B in which every vertex $a \in A'$ is joined to all the vertices $b \in N_H(a)$ while every vertex $a \in A \setminus A'$ is joined to all the vertices $b \in B \setminus N_H(a)$. Given $0 \leq k \leq n/2$, let m_k denote the number of perfect matchings M' in the complete bipartite graph between A and B which contain precisely k edges of H. Every such matching M' can be obtained by first fixing a k-element set $A' \subseteq A$ and then choosing a perfect matching in the graph $H_{A'}$. (Thus the elements of A' correspond to the k endvertices of the edges in $M' \cap E(H)$.) As $n \geq 1000$, we have $\delta(H_{A'}) \geq n/2 - y \geq n/3$, and we can apply Corollary 7 to $H_{A'}$. Altogether, this yields

$$\begin{aligned} m_k &\le \binom{n}{k} \cdot 27n^{3/2} \cdot \left(\frac{n/2+y}{\mathrm{e}}\right)^k \left(\frac{n/2-y}{\mathrm{e}}\right)^{n-k} \\ &\le 2^n \cdot 27n^{3/2} \cdot \mathrm{e}^{-n}(n/2+y)^{n/2}(n/2-y)^{n/2} \\ &= 2^n \cdot 27n^{3/2} \cdot \mathrm{e}^{-n}(n^2/4-y^2)^{n/2} \\ &\le 27n^{3/2} \cdot \left(\frac{n}{\mathrm{e}}\right)^n \cdot \mathrm{e}^{-2y^2/n} \overset{(1)}{\le} \frac{27n!}{n^{5/2}}. \end{aligned}$$

Now consider a perfect matching M in the complete bipartite graph between A and B chosen uniformly at random. Since there are $n!$ such perfect matchings and $n \ge 1000$, the probability that M contains at most $n/2$ edges of H is at most

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \frac{m_k}{n!} < \frac{1}{2n}.$$

This proves the lemma since H is a subgraph of G. □

3 Perfect matchings in r-uniform r-partite hypergraphs

In Theorem 9 below, we prove a slightly strengthened version of the 3-uniform case of Theorem 1. This will then be used to derive the general case of Theorem 1.

Theorem 9 *Let $\mathcal{H}$ be a 3-uniform 3-partite hypergraph with vertex classes A, B and C of size $n \ge 1000$. Suppose that $|N_2(a,b)| \ge n/2 + \sqrt{2n\log n}$ for all pairs $a \in A$, $b \in B$ and $|N_2(a,c)| \ge n/2 + \sqrt{2n\log n}$ for all pairs $a \in A$, $c \in C$. Then $\mathcal{H}$ has a perfect matching.*

Proof. Given a perfect matching M in the complete bipartite graph with vertex classes A and B, we define an auxiliary graph G_M as follows. The vertex classes of G_M are C and M. Vertices $c \in C$ and $ab \in M$ are neighbours in G_M whenever abc is a hyperedge of $\mathcal{H}$. Clearly, $\mathcal{H}$ contains a perfect matching if G_M does. Thus it suffices to show that there exists a choice for M such that G_M has minimum degree at least $n/2$. Then G_M contains a perfect matching by Hall's theorem.

Consider a perfect matching M in the complete bipartite graph with vertex classes A and B which is chosen uniformly at random. By the above, it suffices to show that with positive probability the minimum degree of G_M is at least $n/2$. Since $|N_2(a,b)| \ge n/2 + \sqrt{2n\log n}$ for all pairs $a \in A$, $b \in B$, all the vertices in $M \subseteq V(G_M)$ have degree at least $n/2$ in G_M. Thus we only have to show that for every vertex $c \in C$ the probability that $d_{G_M}(c) < n/2$ is at most $1/(2n)$. But this immediately follows from Lemma 8. Indeed, let G_c denote the bipartite graph with vertex classes A and B such that ab is an edge of G_c whenever abc is a hyperedge of $\mathcal{H}$. Then the degree of c in G_M

is precisely the number of edges in $M \cap E(G_c)$. But for every vertex $a \in A$ we have $d_{G_c}(a) = |N_2(a,c)| \geq n/2 + \sqrt{2n \log n}$. Thus Lemma 8 implies that $|M \cap E(G_c)| \leq n/2$ with probability at most $1/(2n)$. □

Proof of Theorem 1. We will derive Theorem 1 from Theorem 9. Suppose that $r \geq 4$ and let $V_1, \dots, V_r$ denote the vertex classes of $\mathcal{H}$. We first choose any perfect matching M' in the complete $(r-2)$-uniform $(r-2)$-partite hypergraph with vertex classes $V_3, \dots, V_r$. Consider the auxiliary 3-uniform 3-partite hypergraph $\mathcal{H}'$ whose vertex classes are V_1, V_2 and M' and in which $v_1 \in V_1$, $v_2 \in V_2$ and $m \in M'$ from a hyperedge if and only if v_1 and v_2 form a hyperedge of $\mathcal{H}$ together with all the $r-2$ vertices belonging to m. Thus a perfect matching in $\mathcal{H}'$ corresponds to a perfect matching in $\mathcal{H}$. But $\mathcal{H}'$ contains a perfect matching since it satisfies the assumptions of Theorem 9. (Put $A := M'$, $B := V_1$ and $C := V_2$.) □

The following lemma shows that the bound on the minimum degree in Theorem 1 is best possible up to the error term $\sqrt{2n \log n}$.

Lemma 10 *For all integers $r \geq 3$ and every $n \geq 1$ there exists an r-uniform r-partite hypergraph $\mathcal{H}$ with vertex classes of size n which satisfies $\delta'_{r-1}(\mathcal{H}) \geq n/2 - 1$ but does not contain a perfect matching.*

Proof. The vertex classes of our hypergraph $\mathcal{H}$ will be n-element sets $V_1, \dots, V_r$. For each i we choose a set $V'_i \subseteq V_i$ such that $n/2 - 1 \leq |V'_i| \leq n/2 + 1$ and $|V'_1 \cup \dots \cup V'_r|$ is odd. Clearly, this is always possible. The hyperedges of $\mathcal{H}$ will be those r-tuples of vertices which meet each V_i in exactly one vertex and which additionally meet an even number (i.e. possibly none) of the sets $V'_1, \dots, V'_r$. It is easily seen that $\delta'_{r-1}(\mathcal{H}) \geq n/2 - 1$. However, any matching covers an even number of vertices in $V'_1 \cup \dots \cup V'_r$ since each hyperedge of $\mathcal{H}$ contains an even number of these vertices. Hence, as $|V'_1 \cup \dots \cup V'_r|$ is odd, there cannot exist a perfect matching. □

Drake [5] showed that one can also obtain a lower bound of $n/2$ for all n where $n/2$ is odd if r is odd.

4 Almost perfect matchings in r-uniform r-partite hypergraphs

Instead of proving Theorem 2, we will prove the following more general assertion, which implies that Theorem 2 is 'robust' in the sense that we still get an almost perfect matching if the degree is slightly smaller and/or the degree condition fails for only a small fraction of the $(r-1)$-tuples.

Theorem 11 *Let $k \geq 1$, $\ell \geq 0$ be integers and let $\mathcal{H}$ be an r-uniform r-partite hypergraph with vertex classes of size n. Put*

$$\delta' := \begin{cases} \lceil n/r \rceil - \ell & \text{if } n \equiv 0 \bmod r \quad \text{or} \quad n \equiv r-1 \bmod r \\ \lfloor n/r \rfloor - \ell & \text{otherwise.} \end{cases}$$

Suppose that there are fewer than k^{r-1} tuples $x_1, \dots, x_{r-1}$ of vertices in $\mathcal{H}$ such that all the x_i lie in different vertex classes and $|N_{r-1}(x_1, \dots, x_{r-1})| < \delta'$. Then $\mathcal{H}$ has a matching which covers all but at most $(r-1)k + \ell r - 1$ vertices in each vertex class of $\mathcal{H}$.

The proof of Theorem 11 is short and elementary—the idea is to consider a matching of maximum size. This then will turn out to have the required properties.

Proof of Theorem 11. Let $V_1, \dots, V_r$ denote the vertex classes of $\mathcal{H}$. Choose any matching M of $\mathcal{H}$ whose size is maximum. Let $V_i' \subseteq V_i$ be the set of all those vertices which are not covered by M. Then $|V_1'| = \dots = |V_r'| =: s$. Suppose that $s \geq (r-1)k + \ell r \geq (r-1)k$. Then for every $i = 1, \dots, r$ one can find a set A_i which consists of exactly k vertices from each V_j' $(j \neq i)$ and avoids V_i and such that all the A_i are disjoint from each other. Thus each A_i contains vertices $x_1^i, \dots, x_{r-1}^i$ lying in different vertex classes for which $|N_{r-1}(x_1^i, \dots, x_{r-1}^i)| \geq \delta'$. By the choice of M, $N_{r-1}(x_1^i, \dots, x_{r-1}^i)$ lies entirely in $V_i \setminus V_i'$ and thus meets at least δ' of the elements of the matching M. Since $r\delta' > n - (r-1)k - \ell r \geq n - s = |M|$, there exist indices $i \neq j$ such that $N_{r-1}(x_1^i, \dots, x_{r-1}^i)$ and $N_{r-1}(x_1^j, \dots, x_{r-1}^j)$ meet the same element of M, m say. Let M' be the matching obtained from M by deleting m and adding the hyperedge consisting of $x_1^i, \dots, x_{r-1}^i$ together with the unique vertex in $m \cap V_i$ as well as adding the hyperedge consisting of $x_1^j, \dots, x_{r-1}^j$ together with the unique vertex in $m \cap V_j$. Then $|M'| = |M| + 1$, contradicting the choice of M. $\square$

The following lemma shows that the minimum degree in Theorem 2 (and in the $k = 1$, $\ell = 0$ case of Theorem 11) cannot be reduced. Moreover, it implies that if we reduce the minimum degree in Theorem 2 by εn, then one cannot even guarantee a matching which covers all but at most εn vertices in each vertex class.

Lemma 12 *Given integers $q \geq 1$, $r \geq 3$ and $n \geq rq$, suppose that $n = sr + t$ where $s, t \in \mathbb{N}$ and $0 \leq t < r$. Put*

$$C := \begin{cases} rq & \text{if } t = 0 \\ r(q-1) + t & \text{otherwise.} \end{cases}$$

There exists an r-uniform r-partite hypergraph $\mathcal{H}$ with vertex classes of size n such that $\delta_{r-1}'(\mathcal{H}) = \lceil n/r \rceil - q$ and such that every matching in $\mathcal{H}$ avoids at least C vertices in each vertex class of $\mathcal{H}$.

Proof. Put $\delta := \lceil n/r \rceil - q$. Let $V_1, \dots, V_r$ be disjoint n-element sets. For each i let V_i' be any δ-element subset of V_i. Consider the r-uniform hypergraph whose vertex classes are $V_1, \dots, V_r$ and whose hyperedges are precisely those r-tuples which meet each V_i in exactly one vertex and which additionally meet at least one of the sets $V_1', \dots, V_r'$. Then $\delta_{r-1}'(\mathcal{H}) = \delta$. But since every hyperedge of $\mathcal{H}$ meets at least one of the sets V_i', every matching in $\mathcal{H}$ has at most $r\delta$ elements and thus avoids at least $n - r\delta = C$ vertices in each of the vertex classes. $\square$

Note that Lemma 12 implies that Theorem 11 with $\ell = 0$ and $k = 1$ is also sharp in the sense that if $n = sr + (r-2)$, then there exists an r-uniform r-partite hypergraph with vertex classes of size n which has minimum degree $\lfloor n/r \rfloor$ and where every matching misses $r-2$ vertices in each vertex class.

5 Matchings in general r-uniform hypergraphs

To derive Theorem 3 from Theorem 1 we show that the vertex set of every r-uniform hypergraph $\mathcal{H}$ as in Theorem 3 can be partitioned into r vertex classes of equal size such that the r-partite subhypergraph thus obtained satisfies the conditions of Theorem 1. Indeed, a straightforward argument (see Proposition 13) shows that a random partition of the vertex set of $\mathcal{H}$ into classes of equal size will have the desired properties. To work out the details, we need the following definition. Suppose that $\mathcal{H}$ is an r-uniform hypergraph whose number of vertices is divisible by r, $|\mathcal{H}| = rn$ say. Given a set $N \subseteq V(\mathcal{H})$, we say that a partition $V_1, \dots, V_r$ of the vertex set of $\mathcal{H}$ *splits* N *fairly* if $|V_i| = n$ and

$$\left| |N \cap V_i| - \frac{|N|}{r} \right| \leq 2r\sqrt{n \log n} \tag{2}$$

for every $i \leq r$.

Proposition 13 *For each integer $r \geq 2$ there exists an integer $n_0 = n_0(r)$ such that the following holds. Suppose that $n \geq n_0$ and that $\mathcal{H}$ is an r-uniform hypergraph with rn vertices. Then there exists a partition $V_1, \dots, V_r$ of the vertex set of $\mathcal{H}$ which splits all neighbourhoods $N_{r-1}(x_1, \dots, x_{r-1})$ fairly.*

Proposition 13 follows from a straightforward application of the following large deviation bound for the hypergeometric distribution (see e.g. [9, Thm. 2.10, Cor. 2.3 and Cor. 2.4]).

Lemma 14 *Given $q \in \mathbb{N}$ and sets $N \subseteq V$ with $|V| \geq q$, let Q be a subset of V which is obtained by successively selecting q elements of V uniformly at random without repetitions. Let $X := |N \cap Q|$.*

(i) *For all $0 < \alpha \leq 3/2$ we have $\mathbb{P}(|X - \mathbb{E}X| \geq \alpha \mathbb{E}X) \leq 2\mathrm{e}^{-\frac{\alpha^2}{3}\mathbb{E}X}$.*

(ii) *If $\alpha' \geq \frac{3}{2}\mathbb{E}X$, we have $\mathbb{P}(X \geq \alpha') \leq \mathrm{e}^{-c'\alpha'}$, where c' is an absolute constant.*

Proof of Proposition 13. Choose a partition $V_1, \dots, V_r$ with $|V_1| = \dots = |V_r| = n$ of the vertex set $V(\mathcal{H})$ of $\mathcal{H}$ uniformly at random from the set of all such partitions. Clearly, the probability that V_i equals a fixed n-element subset of $V(\mathcal{H})$ is the same as when V_i was obtained by successively selecting n elements of $V(\mathcal{H})$ uniformly at random without repetitions. Thus we may apply Lemma 14 with any one of the V_i taking the role of Q. Set $\gamma := 2r\sqrt{(\log n)/n}$.

Consider first any neighbourhood $N_{r-1}(x_1, \dots, x_{r-1}) =: N$ whose size is at least $2r\gamma n/3$. Since $\mathbb{E}(|N \cap V_i|) = |N|/r \le n$, Lemma 14(i) implies that

$$\begin{aligned}\mathbb{P}(|\,|N \cap V_i| - |N|/r\,| \ge \gamma n) &= \mathbb{P}(|\,|N \cap V_i| - |N|/r\,| \ge (\gamma nr/|N|)|N|/r) \\ &\le 2\mathrm{e}^{-\frac{\gamma^2 n^2}{3}\frac{r}{|N|}} \le 2\mathrm{e}^{-\frac{\gamma^2 n}{3}} \le \mathrm{e}^{-r^2 \log n} < \frac{1}{r} \cdot \binom{|\mathcal{H}|}{r-1}^{-1}.\end{aligned}$$

If $|N| \le 2r\gamma n/3$, we get the same bound using Lemma 14(ii). Indeed, since $\mathbb{E}(|N \cap V_i|) = |N|/r \le 2\gamma n/3$, Lemma 14(ii) implies that

$$\begin{aligned}\mathbb{P}(|\,|N \cap V_i| - |N|/r\,| \ge \gamma n) &= \mathbb{P}(|N \cap V_i| \ge |N|/r + \gamma n) \\ &\le \mathbb{P}(|N \cap V_i| \ge \gamma n) \\ &\le \mathrm{e}^{-c'\gamma n} = \mathrm{e}^{-c' 2r\sqrt{\log n}\sqrt{n}} < \frac{1}{r} \cdot \binom{|\mathcal{H}|}{r-1}^{-1}.\end{aligned}$$

Hence with probability $> 1 - 1/r$ the set V_i satisfies

$$|\,|N_{r-1}(x_1, \dots, x_{r-1}) \cap V_i| - |N_{r-1}(x_1, \dots, x_{r-1})|/r\,| \le \gamma n \tag{3}$$

for all neighbourhoods $N_{r-1}(x_1, \dots, x_{r-1})$. Thus the probability that all the partition sets V_i satisfy (3) is positive. Therefore there exists an outcome $V_1, \dots, V_r$ with this property. This is a partition of $V(\mathcal{H})$ as required. □

Proof of Theorem 3. First apply Proposition 13 to obtain a partition of the vertex set of $\mathcal{H}$ into $V_1, \dots, V_r$ which splits all the neighbourhoods $N_{r-1}(x_1, \dots, x_{r-1})$ fairly. Let $\mathcal{H}'$ be the r-uniform r-partite subhypergraph of $\mathcal{H}$ defined in this way. (So the hyperedges of $\mathcal{H}'$ are those hyperedges of $\mathcal{H}$ which meet every V_i in precisely one vertex.) Then $\delta'_{r-1}(\mathcal{H}') \ge n/2 + 3r\sqrt{n \log n} - 2r\sqrt{n \log n} \ge n/2 + \sqrt{2n \log n}$. Thus Theorem 1 implies that $\mathcal{H}'$ (and hence also $\mathcal{H}$) has a perfect matching. □

The following lemma implies that Theorem 3 becomes false if the minimum degree is 'a bit below' $|\mathcal{H}|/2$. It would be interesting to know whether there are even counterexamples $\mathcal{H}$ with $\delta_{r-1}(\mathcal{H}) \ge |\mathcal{H}|/2 - 1$.

Lemma 15 *For all integers $r \ge 3$ and every odd $n \ge r$ there exists an r-uniform hypergraph $\mathcal{H}$ with $2rn$ vertices which satisfies*

$$\delta_{r-1}(\mathcal{H}) = \begin{cases} rn - (r-1) & \text{if } r \text{ is odd} \\ rn - r & \text{if } r \text{ is even} \end{cases}$$

but does not contain a perfect matching.

Proof. We first consider the case when r is odd. Let A and B be two rn-element sets. The vertex set of our hypergraph $\mathcal{H}$ will be $A \cup B$. The hyperedges of $\mathcal{H}$ will be those r-tuples of distinct vertices which meet A in an even (and thus possibly empty) number of vertices. Then $\delta_{r-1}(\mathcal{H}) = rn - (r-1)$ and every matching covers only an even number of vertices in A (and thus cannot cover all of them).

In the case when r is even we proceed similarly except that A is now an $(rn+1)$-element set and B is an $(rn-1)$-element set. □

Instead of Theorem 4, we will prove the following strengthening.

Theorem 16 *For every integer $r \geq 3$ there exists an integer $n_0 = n_0(r)$ such that for every $n \geq n_0$ the following holds. Let $k \geq 1$ and $\ell \geq 0$ be integers and suppose that $\mathcal{H}$ is an r-uniform hypergraph with $|\mathcal{H}| = rn$ vertices which has less than k^{r-1} tuples $x_1, \dots, x_{r-1}$ of distinct vertices satisfying $|N_{r-1}(x_1, \dots, x_{r-1})| < |\mathcal{H}|/r - \ell + 3r^2\sqrt{n \log n}$. Then $\mathcal{H}$ has a matching which covers all but at most $r^2 k + \ell r$ vertices of $\mathcal{H}$.*

Theorem 4 follows immediately from the $k = 1$ case. Indeed, if the number of vertices of $\mathcal{H}$ is not divisible by r, then apply Theorem 16 to the hypergraph which is obtained from $\mathcal{H}$ by deleting up to $r - 1$ vertices.

Theorem 16 can be derived from Theorem 11 in the same way as Theorem 3 was derived from Theorem 1.

Proof of Theorem 16. First apply Proposition 13 to obtain a partition of the vertex set of $\mathcal{H}$ into $V_1, \dots, V_r$ which splits all neighbourhoods $N_{r-1}(x_1, \dots, x_{r-1})$ fairly. Let $\mathcal{H}'$ be the r-uniform r-partite subhypergraph of $\mathcal{H}$ defined in this way. Then there are less than k^{r-1} tuples $x_1, \dots, x_{r-1}$ such that all the x_i lie in different vertex classes of $\mathcal{H}'$ and which satisfy $|N_{r-1}(x_1, \dots, x_{r-1})| < n/r - \ell/r + 3r\sqrt{n \log n} - 2r\sqrt{n \log n}$. Since the right hand side is at least $\lceil n/r \rceil - \lfloor \ell/r \rfloor$, Theorem 11 implies that $\mathcal{H}'$ (and hence also $\mathcal{H}$) has a matching which avoids at most $r(r-1)k + r^2(\ell/r) - r \leq r^2 k + \ell r$ vertices. □

The following lemma implies that if we reduce the minimum degree in Theorem 4 by $2\varepsilon|\mathcal{H}|$, then one cannot guarantee a matching which leaves less than $r\varepsilon|\mathcal{H}|$ vertices uncovered, provided that $|\mathcal{H}|$ is sufficiently large compared to r and ε.

Lemma 17 *For all integers $r \geq 3$, $q \geq 1$ and every $n \geq q, r$ there exists an r-uniform hypergraph $\mathcal{H}$ with rn vertices which satisfies $\delta_{r-1}(\mathcal{H}) = n - q$ but does not contain a matching which avoids less than rq of its vertices.*

Proof. Let A be an $(n - q)$-element set and let B be an $(rn - (n - q))$-element set. The vertex set of our hypergraph will be $A \cup B$. The hypergedges of $\mathcal{H}$ will be those r-tuples which meet A in at least one vertex. Thus $\delta_{r-1}(\mathcal{H}) \geq n - q$. However, every matching in $\mathcal{H}$ covers at most $|B| - rq$ vertices in B since every hyperedge has at most $r - 1$ vertices in B and $(r-1)|A| = rn - (n - q) - rq = |B| - rq$. □

References

[1] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[2] B. Bollobás, *Random Graphs* (2nd edition), Cambridge studies in Advanced Mathematics 73, Cambridge University Press 2001.

[3] L.M. Brégman, Some properties of nonnegative matrices and their permanents, *Soviet Mathematics Doklady* **14** (1973), 945–949.

[4] M. Conforti, G. Cornuéjols, A. Kapoor and K. Vušković, Perfect matchings in balanced hypergraphs, *Combinatorica* **16** (1996), 325–329.

[5] D. Drake, in preparation.

[6] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, Freeman 1979.

[7] C. Hurkens and A. Schrijver, On the size of systems of sets every t of which have an SDR, with an application to the worst-case ratio of heuristics for packing problems, *SIAM J. Disc. Math.* **2** (1989), 68-72.

[8] P.E. Haxell, A condition for matchability in hypergraphs, *Graphs and Comb.* **11** (1995), 245–248.

[9] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience 2000.

[10] V. Kann, Maximum bounded 3-dimensional matching is MAX SNP-complete, *Inform. Process. Lett.* **37** (1991), 27–35.

[11] J.H. Kim, Perfect matchings in random uniform hypergraphs, *Random Structures & Algorithms* **23** (2003), 111–132.

[12] L. Lovász and M.D, Plummer, *Matching theory*, North-Holland 1986.

[13] V. Rödl, A. Ruciński and E. Szemerédi, preprint 2004.

[14] A. Schrijver, A short proof of Minc's conjecture, *J. Combin. Theory A* **25** (1978), 80–83.

[15] V.H. Vu, New bounds on nearly perfect matchings in hypergraphs: higher codegrees do help, *Random Structures & Algorithms* **17** (2000), 29–63.

Version 24 September 2004

Daniela Kühn & Deryk Osthus
School of Mathematics and Statistics
Birmingham University
Edgbaston
Birmingham B15 2TT
UK
E-mail addresses: {kuehn,osthus}@maths.bham.ac.uk