On a hypergraph matching problem

Noga Alon * Raphael Yuster [†]

Abstract

Let H = (V, E) be an *r*-uniform hypergraph and let $\mathcal{F} \subset 2^V$. A matching M of H is (α, \mathcal{F}) perfect if for each $F \in \mathcal{F}$, at least $\alpha |F|$ vertices of F are covered by M. Our main result is
a theorem giving sufficient conditions for an *r*-uniform hypergraph to have a $(1 - \epsilon, \mathcal{F})$ -perfect
matching. As a special case of our theorem we obtain the following result. Let K(n, r) denote
the complete *r*-uniform hypergraph with *n* vertices. Let *t* and *r* be fixed positive integers where $t \geq r \geq 2$. Then, K(n, r) can be packed with edge-disjoint copies of K(t, r) such that each
vertex is incident with only $o(n^{r-1})$ unpacked edges. This extends a result of Rödl [9].

1 Introduction

A hypergraph H is an ordered pair H = (V, E) where V is a finite set (the vertex set) and E is a family of distinct subsets of V (the edge set). A hypergraph is r-uniform if all edges have size r. In this paper we only consider r-uniform hypergraphs where $r \ge 2$ is fixed. A subset $M \subseteq E(H)$ is a matching if every pair of edges from M has an empty intersection. A matching is called perfect if |M| = |V|/r. A vertex $v \in V$ is covered by the matching M if some edge from M contains v. Let $\mathcal{F} \subset 2^V$ and let $0 \le \alpha \le 1$. A matching M is (α, \mathcal{F}) -perfect if for each $F \in \mathcal{F}$, at least $\alpha |F|$ vertices of F are covered by M. Thus, a $(1, \{V\})$ -perfect matching is simply a perfect matching.

A seminal result of Pippenger strengthening an earlier result of Frankl and Rödl [3] on near perfect coverings and matchings of uniform hypergraphs gives sufficient conditions for the existence of a $(1 - \epsilon, \{V\})$ -perfect matching. For $x, y \in V$ let d(x) denote the number of edges containing x (the *degree* of x) and let d(x, y) denote the number of edges that contain both x and y (the *co-degree* of x and y). Let $\Delta(H), \delta(H)$ and $\Delta_2(H)$ denote the maximum degree, minimum degree and maximum co-degree of H, respectively. The following is an unpublished result of Pippenger, strengthening a theorem proved in [3].

^{*}Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: nogaa@post.tau.ac.il. Research supported in part by a USA Israeli BSF grant, by a grant from the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

[†]Department of Mathematics, University of Haifa at Oranim, Tivon 36006, Israel. E–mail: raphy@research.haifa.ac.il

Theorem 1.1 (Pippenger) For an integer $r \ge 2$ and a real $\epsilon > 0$ there exists a real $\mu = \mu(r, \epsilon)$ so that the following holds: If the r-uniform hypergraph H on n vertices satisfies: (i) $\delta(H) \ge (1 - \mu)\Delta(H)$, (ii) $\Delta_2(H) < \mu\Delta(H)$, then H has a matching that covers all but at most ϵn vertices.

In other words, if H is nearly regular and the maximum co-degree is relatively small compared to the maximum degree then an almost perfect matching is guaranteed to exist. We note that the proof also applies to the analogous covering version. Furthermore, as noted in [4], the statement of Theorem 1.1 remains valid even if we allow a small fraction of the vertices to have degrees that deviate significantly from the average degree.

The main result in this short paper is a variant of Theorem 1.1, giving sufficient conditions for the existence of a $(1 - \epsilon, \mathcal{F})$ -perfect matching. We show that essentially the same conditions of Theorem 1.1 suffice to guarantee a $(1 - \epsilon, \mathcal{F})$ -perfect matching even if \mathcal{F} is quite large and even if the sizes of the elements of \mathcal{F} vary significantly.

For $\mathcal{F} \subset 2^V$ let $s(\mathcal{F}) = \min_{F \in \mathcal{F}} |F|$, and for a hypergraph H let $g(H) = \Delta(H)/\Delta_2(H)$.

Theorem 1.2 For an integer $r \ge 2$, a real C > 1 and a real $\epsilon > 0$ there exist a real $\mu = \mu(r, C, \epsilon)$ and a real $K = K(r, C, \epsilon)$ so that the following holds: If the r-uniform hypergraph H = (V, E) on n vertices satisfies:

(i) $\delta(H) \ge (1-\mu)\Delta(H)$, (ii) $g(H) > \max\{1/\mu, K(\ln n)^6\}$, then for every $\mathcal{F} \subset 2^V$ with $|\mathcal{F}| \le C^{g(H)^{1/(3r-3)}}$ and with $s(\mathcal{F}) \ge 5g(H)^{1/(3r-3)}\ln(|\mathcal{F}|g(H))$ there is $a (1-\epsilon, \mathcal{F})$ -perfect matching in H.

The proof of Theorem 1.2, which is presented in the following section, relies on probabilistic arguments and on a result of Pippenger and Spencer concerning the chromatic index of nearly regular hypergraphs with small co-degrees [8]. In the final section we describe some applications of Theorem 1.2. The main one is an extension of a result of Rödl on packing the complete *n*-vertex *r*-uniform hypergraph with a complete *r*-uniform hypergraph of fixed size [9]. We show that there is always such a packing for which every vertex of K(n, r) is incident with only $o(n^{r-1})$ unpacked edges.

2 Proof of the main result

The chromatic index of a hypergraph S, denoted q(S), is the smallest integer q such that the set of edges of S can be partitioned into q matchings. The following result of Pippenger and Spencer [8] gives sufficient conditions on S which guarantee that q(S) is very close to the maximum degree of S. **Lemma 2.1 (Pippenger and Spencer [8])** For an integer $r \ge 2$ and a real $\gamma > 0$ there exists a real $\beta = \beta(r, \gamma)$ so that the following holds: If an r-uniform hypergraph S has the following properties for some t:

(i) $(1 - \beta)t < d(x) < (1 + \beta)t$ holds for all vertices, (ii) $d(x, y) < \beta t$ for all distinct x and y, then $q(S) \le (1 + \gamma)t$.

Better estimates for the error term and some extensions have been proved subsequently in [6, 7]. **Proof of Theorem 1.2:** Fix an integer $r \ge 2$ and reals $\epsilon > 0$ and C > 1. Let $H(x) = -x \log x - (1-x) \log(1-x)$ be the entropy function. Let ζ be chosen such that $\zeta < (2^{-H(\epsilon/4)}C^{-1})^{1/\epsilon}$. Let γ be chosen sufficiently small so that $2(1 - (1 - \beta)/(1 + \gamma))/\epsilon \le \zeta$ where $\beta = \beta(r, \gamma)$ is the constant from Lemma 2.1. Let ρ be chosen such that $\rho + \rho^2 = \beta$. Let $\mu = \rho^{12}/(r-1)^6$. Let c_{ρ} be a (small) constant depending only on ρ . We choose c_{ρ} during the proof. Let $K = (r/c_{\rho})^6$.

Suppose H = (V, E) is an *n*-vertex *r*-uniform hypergraph satisfying the conditions of Theorem 1.2. Let $\mathcal{F} \subset 2^V$ where $|\mathcal{F}| \leq C^{g(H)^{1/(3r-3)}}$ and $s(\mathcal{F}) \geq 5g(H)^{1/(3r-3)} \ln(|\mathcal{F}|g(H))$. We need to show that H has a $(1 - \epsilon, \mathcal{F})$ -perfect matching.

Our proof proceeds as follows. In the first stage we randomly color the vertices with m colors. Our goal is to choose m large while still guaranteeing that with high probability, for each $i = 1, \ldots, m$, the subhypergraph S_i of H induced by the vertices colored i satisfies the conditions of Lemma 2.1. We also want the random coloring to guarantee, with high probability, that each $F \in \mathcal{F}$ has roughly |F|/m vertices in each color. By showing these, we may fix a coloring and fix S_1, \ldots, S_m having these properties. In the second stage we pick, for each S_i , an arbitrary edge-coloring with the properties guaranteed to exist by Lemma 2.1. We then pick a random color class, M_i , which is a matching of S_i . In the final stage, we prove that, with positive probability, $M_1 \cup \cdots \cup M_m$ is a $(1 - \epsilon, \mathcal{F})$ -perfect matching. We now describe each stage in detail.

Let $m = g(H)^{1/(3r-3)}$ (in the sequel we shall ignore floors and ceilings whenever appropriate as this does not affect the asymptotic nature of our result; hence me may assume m is an integer), and let p = 1/m. Each $v \in V$ selects a color from $\{1, \ldots, m\}$ at random. The choices made by distinct vertices are independent. Let V_i be the set of vertices with color i. Let S_i be the subhypergraph of H induced by V_i . For $x, y \in V_i$, let $d_i(x)$ be the degree of x in S_i and let $d_i(x, y)$ be their co-degree in S_i .

Claim 2.2 With probability at least 0.5, for some t and for all i = 1, ..., m, (i) $(1 - \beta)t < d_i(x) < (1 + \beta)t$ holds for all $x \in V_i$, (ii) $d_i(x, y) < \beta t$ for all distinct $x, y \in V_i$. **Proof:** Let $t = \Delta(H)g(H)^{-1/3}$. For this choice of t, the requirement (ii) clearly holds (deterministically). Indeed,

$$d_i(x,y) \le d(x,y) \le \Delta_2(H) = \Delta(H)/g(H) = tg(H)^{-2/3} < t\mu^{2/3} < t\rho < t\beta.$$

For $x \in V_i$, let A_x denote the event that $d_i(x) \notin [(1-\beta)t, (1+\beta)t]$. In order to prove (i) (and hence the claim), it suffices to show that A_x holds with probability less than 1/(2n).

Fix $x \in V_i$. For each $e \in E$ with $x \in e$, we have $\Pr[e \in S_i] = p^{r-1}$. Thus,

$$E[d_i(x)] = d(x)p^{r-1} = d(x)g(H)^{-1/3}$$

Let N(x) denote the set of edges of H incident with x. Consider the graph G_x whose vertex set is N(x). We connect $e, f \in N(x)$ in G_x if and only if $|e \cap f| \geq 2$. Notice that $\Delta(G_x) < (r-1)\Delta_2(H)$. In particular, $\chi = \chi(G_x) \leq (r-1)\Delta_2(H)$. Let R_1, \ldots, R_{χ} be a partition of N(x) such that R_j is a delta system with core x for $j = 1, \ldots, \chi$. Notice that $|R_1| + \cdots + |R_{\chi}| = |N(x)| = d(x)$. Let $N_i(x)$ be the set of edges of S_i incident with x and let $N_{i,j}(x) = N_i(x) \cap R_j$. Put $d_{i,j}(x) = |N_{i,j}(x)|$. Clearly, $d_i(x) = \sum_{j=1}^{\chi} d_{i,j}(x)$. Furthermore, $E[d_{i,j}(x)] = |R_j|p^{r-1}$. Given that $x \in V_i$, for any edge $e \in R_j$, the event that $e \in N_{i,j}(x)$ is *independent* of all the events $f \in N_{i,j}(x)$. Thus, by a large deviation inequality of Chernoff (cf. [1] Appendix A), there exists a positive constant c_{ρ} depending only on ρ such that

$$\Pr\left[\left|d_{i,j}(x) - |R_j|p^{r-1}\right| \ge \rho |R_j|p^{r-1}\right] < 2\exp(-c_\rho |R_j|p^{r-1}) = 2\exp(-c_\rho |R_j|g(H)^{-1/3}).$$
(1)

We call R_j large if $|R_j| > g(H)^{1/2}$. Otherwise, R_j is small. By (1) we have that for large R_j ,

$$\Pr\left[\left|d_{i,j}(x) - |R_j|p^{r-1}\right| \ge \rho |R_j|p^{r-1}\right] < 2\exp(-c_\rho g(H)^{1/6}) < \frac{1}{2\chi n}.$$
(2)

Notice that the last inequality follows from the fact that $g(H) > K(\ln n)^6$. The sum of the sizes of the small R_j is relatively small. Indeed,

$$\sum_{j, |R_j| \le g(H)^{1/2}} |R_j| \le \chi g(H)^{1/2} \le (r-1)\Delta_2(H)g(H)^{1/2}$$
(3)
= $(r-1)\Delta(H)g(H)^{-1/2} = (r-1)tg(H)^{-1/6}.$

We therefore have

$$\sum_{j, |R_j| > g(H)^{1/2}} |R_j| \ge d(x) - (r-1)tg(H)^{-1/6}.$$
(4)

It follows from (2), (3) and (4) that with probability at least 1 - 1/(2n),

$$d_i(x) = \sum_{j=1}^{\chi} d_{i,j}(x) \le d(x)(1+\rho)p^{r-1} + (r-1)tg(H)^{-1/6}$$

$$= d(x)(1+\rho)g(H)^{-1/3} + (r-1)tg(H)^{-1/6} \le (1+\rho+\rho^2)t = (1+\beta)t$$

(we used here the fact that $\rho^2 = (r-1)\mu^{1/6} > (r-1)g(H)^{-1/6}$) and also

$$\begin{aligned} d_i(x) &= \sum_{j=1}^{\chi} d_{i,j}(x) \ge (d(x) - (r-1)tg(H)^{-1/6})(1-\rho)p^{r-1} \\ &= d(x)(1-\rho)g(H)^{-1/3} - (r-1)(1-\rho)tg(H)^{-1/2} \ge (1-\rho)t(1-\mu-(r-1)g(H)^{-1/2}) \\ &> (1-\rho)t(1-\rho^2) > (1-\beta)t. \end{aligned}$$

In particular, A_x holds with probability at most 1/(2n).

For $F \in \mathcal{F}$, let $F_i = F \cap V_i$. We say that F is *deviating* if for some i, $|F_i| > 2|F|/m$. The following simple claim gives an upper bound for the probability that some F is deviating.

Claim 2.3 The probability that some $F \in \mathcal{F}$ is deviating is less than 0.5.

Proof: Fix $i \in \{1, ..., m\}$ and fix $F \in \mathcal{F}$. The expectation of $|F_i|$ is |F|/m. As each vertex chooses its color independently, we have by a Chernoff inequality (cf. [1]) that

$$\Pr\left[|F_i| - \frac{|F|}{m} > \frac{|F|}{m}\right] < \exp\left(-\frac{2}{27}\frac{|F|}{m}\right) < \exp\left(-\frac{2}{27}\frac{s(\mathcal{F})}{m}\right).$$

By the last inequality and the assumption that $s(\mathcal{F}) \geq 5g(H)^{1/(3r-3)} \ln(|\mathcal{F}|g(H))$, the probability that some $F \in \mathcal{F}$ is deviating is less than

$$|\mathcal{F}|m\exp(-2s(\mathcal{F})/(27m)) < 0.5.$$

By Claim 2.2 and Claim 2.3 we may fix a coloring of the vertices of H such that all the subhypergraphs S_i satisfy the conditions of Lemma 2.1 and such that for each $F \in \mathcal{F}$, the number of vertices colored i is at most 2|F|/m. For each S_i let $q_i = q(S_i)$ and let $M(i, 1), \ldots, M(i, q_i)$ be a partition of the edges of S_i to q_i matchings. By Lemma 2.1 we have $q_i \leq (1+\gamma)\Delta(H)g(H)^{-1/3}$. For each $i = 1, \ldots, m$ we pick at random, and independently, a matching M(i, j). Let $M = \bigcup_{i=1}^m M(i, j)$. Notice that M is a matching of H. Let $F \in \mathcal{F}$. In order to complete the proof of Theorem 1.2 it suffices to prove the following claim.

Claim 2.4 With probability greater than $1 - 1/|\mathcal{F}|$, at least $(1 - \epsilon)|F|$ vertices of F are covered by M.

Proof: Let $F_i = F \cap V_i$ for i = 1, ..., m. We say that F_i is badly covered by M if more than $\epsilon |F_i|/2$ vertices of F_i are uncovered by M(i, j). Let ℓ denote the number of badly covered F_i . We first notice that if $\ell \leq \epsilon m/4$ then M covers at least $(1 - \epsilon)|F|$ vertices of F. Indeed, the total size of all the badly covered subsets is at most $2\ell |F|/m$. The remaining subsets have total size at least $|F|(1 - 2\ell/m)$ and hence at least $(1 - \epsilon/2)|F|(1 - 2\ell/m) > (1 - \epsilon)|F|$ vertices of F are covered by M. It remains to show that the probability that F has more than $\epsilon m/4$ badly covered subsets is less than $1/|\mathcal{F}|$. Notice that if $i \neq j$, the event that F_i is badly covered is independent of the event that F_j is badly covered. Thus, if ζ is an upper bound for the probability that F_i is badly covered, and ζ is independent of i, it suffices to prove that

$$\binom{m}{\epsilon m/4} \zeta^{\epsilon m} < \frac{1}{|\mathcal{F}|}.$$

Since $d_i(x) \ge (1-\beta)t$ for all $x \in V_i$ and since $q_i \le (1+\gamma)t$ we have that the expected number of vertices of F_i covered by M(i,j) is at least $\frac{1-\beta}{1+\gamma}|F_i|$. Therefore, the probability that there are less than $(1-\epsilon/2)|F_i|$ vertices of $|F_i|$ covered by M(i,j) is at most $2(1-(1-\beta)/(1+\gamma))/\epsilon \le \zeta$. However, recall that ζ was chosen such that $\zeta < (2^{-H(\epsilon/4)}C^{-1})^{1/\epsilon}$. We therefore have

$$\binom{m}{\epsilon m/4} \zeta^{\epsilon m} < \frac{1}{C^m} < \frac{1}{|\mathcal{F}|}$$

By Claim 2.3 we have that with positive probability, M is an $(1 - \epsilon, \mathcal{F})$ -perfect matching. We have therefore completed the proof of Theorem 1.2.

It is easy to implement the proof of Theorem 1.2 as a polynomial time randomized algorithm. In fact, all the details of the proof are easily seen to be algorithmic and the only "black box" that is used is Lemma 2.1. Fortunately, Grable [5] gave an algorithmic proof of the Pippenger-Spencer theorem. It is possible to derandomize the algorithm using the method of conditional probabilities [1]. The only obstacle is that when we implement Claim 2.3 (constructing the vertex coloring) and Claim 2.4 (constructing the matchings M(i, j)), the method of conditional probabilities requires that we scan all elements of \mathcal{F} in every step. Thus, the derandomized algorithm is only polynomial in $|\mathcal{F}| + n$ (one may argue that this is fine since the input should contain a list of all elements of \mathcal{F} , but in practice one should think of \mathcal{F} as being described implicitly). In applications where $|\mathcal{F}|$ is bounded by a polynomial in n (such as the applications given in the next section), the derandomized algorithm runs in polynomial (in n) time.

3 Remarks and Applications

• Already a special case of Theorem 1.2 enables us to prove the following strengthening of a theorem of Rödl [9]. Let K(n, k) denote the complete k-uniform hypergraph with n vertices. A

K(t,k)-packing of K(n,k) is a set of edge-disjoint copies of K(t,k) in K(n,k). Let p(n,t,k) denote the maximum size of a K(t,k)-packing of K(n,k). Clearly, $p(n,t,k) \leq {\binom{n}{k}}/{\binom{t}{k}}$. Solving a longstanding conjecture of Erdős and Hanani, Rödl proved that for every $\epsilon > 0$, and for fixed integers t, k with t > k > 1, if n is sufficiently large then $p(n,t,k) > (1-\epsilon) {\binom{n}{k}}/{\binom{t}{k}}$. Notice that such an almost-optimal packing may still be "unfair" to some vertices. In fact, a vertex might still have $\Theta(n^{k-1})$ incident edges that are unpacked. Using Theorem 1.2 we are able to prove that there always exists an almost-optimal packing which is fair.

Theorem 3.1 Let t, k be fixed integers with t > k > 1, and let $\epsilon > 0$. For n sufficiently large there is a K(t,k)-packing of K(n,k) such that each vertex is incident with at most $\epsilon \binom{n-1}{k-1}$ unpacked edges.

Proof: We apply Theorem 1.2 with ϵ , $r = \binom{t}{k}$ and, say, C = 1.1. Given K(n,k) we create another hypergraph H = H(n,k) as follows. The vertices of H are the edges of K(n,k). The edges of H are the K(t,k) copies of K(n,k). Notice that H is r-uniform and has $N = \binom{n}{k}$ vertices. Also, $\Delta(H) = \delta(H) = \binom{n-k}{t-k}$. Notice also that any two edges of K(n,k) appear together in at most $\binom{n-k-1}{t-k-1}$ copies of K(t,k). Thus, $\Delta_2(H) = \binom{n-k-1}{t-k-1}$ and g(H) = (n-k)/(t-k). For each vertex $v \in \{1,\ldots,n\}$ of K(n,k) let F_v be the set of edges incident with v. Note that F_v is also a subset of vertices of H with $|F_v| = \binom{n-1}{k-1}$. Let $\mathcal{F} = \{F_v \mid v = 1,\ldots,n\}$. Thus, $|\mathcal{F}| = n$ and $s(\mathcal{F}) = \binom{n-1}{k-1}$. Let K and μ be the constants from Theorem 1.2. It is easy to see that for n sufficiently large (and hence N sufficiently large), the conditions of Theorem 1.2 are satisfied. Therefore, H has a $(1 - \epsilon, \mathcal{F})$ -perfect matching. This, in turn, implies that there is a K(t,k)-packing of K(n,k) such that each vertex is incident with at most $\epsilon\binom{n-1}{k-1}$ unpacked edges.

• A similar reasoning enables us to obtain the following strengthening of the main result of Frankl and Füredi in [2].

Theorem 3.2 Let $H = (U, \mathcal{F})$ be a fixed k-uniform hypergraph with $|\mathcal{F}| = f$ edges and fix $\epsilon > 0$. Then for all sufficiently large n there is a family \mathcal{H} of copies $H_1 = (U_1, \mathcal{F}_1), H_2 = (U_2, \mathcal{F}_2), \ldots$ of H in the complete k-uniform hypergraph K on n vertices such that $|U_i \cap U_j| \leq k$ for all $i \neq j$, If $|U_i \cap U_j| = k$ and $U_i \cap U_j = B$ then $B \notin \mathcal{F}_i, B \notin \mathcal{F}_j$. Each vertex of K is incident with less than ϵn^{k-1} edges that do not belong to any member of \mathcal{H} .

The proof is by choosing a random subhypergraph K' of K, where each edge is chosen, randomly and independently, with probability $p = 1 - \delta$, for an appropriate small δ . We next consider the hypergraph whose vertices are the edges of the random subhypergraph K', where each *induced* copy of H in K' forms an edge. One can now complete the proof by applying Theorem 1.2 to this last hypergraph, with the obvious choice of the members of \mathcal{F} . We omit the details.

• A possible modification in the proof of Theorem 1.2 enables one to formulate a version of Theorem 1.2 in which g(H) does not have to grow with n, assuming the family \mathcal{F} satisfies appropriate regularity assumptions that will enable us to apply the Local Lemma. We omit the details.

References

- N. Alon and J. H. Spencer, The Probabilistic Method, Second Edition, Wiley, New York, 2000.
- [2] P. Frankl and Z. Füredi, Colored packing of sets, Combinatorial design theory, 165–177, North-Holland Math. Stud., 149, North-Holland, Amsterdam, 1987.
- [3] P. Frankl and V. Rödl, Near perfect coverings in graphs and hypergraphs, European J. Combinatorics 6 (1985), 317–326.
- [4] Z. Füredi, Matchings and covers in hypergraphs, Graphs and Combinatorics 4 (1988), 115–206.
- [5] D. Grable, Nearly-perfect hypergraph packing is in NC, Information Processing Letters 60 (1996), 295–299.
- [6] J. Kahn, Asymptotically good list colorings, J. Combin. Theory, Ser. A 73 (1996), 1–59.
- [7] M. Molloy and B. Reed, Near optimal list colorings, Random Structures and Algorithms 17 (2000), 376–402.
- [8] N. Pippenger and J. Spencer, Asymptotic behavior of the chromatic index for hypergraphs, J. Combin. Theory, Ser. A 51 (1989), 24–42.
- [9] V. Rödl, On a packing and covering problem, Europ. J. of Combin. 6 (1985), 69–78.