Nearly perfect matchings in regular simple hypergraphs

Noga Alon [*] Jeong-Han Kim [†] Joel Spencer [‡]

Abstract

For every fixed $k \geq 3$ there exists a constant c_k with the following property. Let H be a k-uniform, D-regular hypergraph on N vertices, in which no two edges contain more than one common vertex. If $k > 3$ then H contains a matching covering all vertices but at most $c_k N D^{-1/(k-1)}$. If $k = 3$, then H contains a matching covering all vertices but at most $c_3 N D^{-1/2} \ln^{3/2} D$. This improves previous estimates and implies, for example, that any Steiner Triple System on N vertices contains a matching covering all vertices but at most $O(N^{1/2} \ln^{3/2} N)$, improving results by various authors.

1 Introduction

A *Hypergraph* is a pair (V, H), where V is a finite set of *vertices* and H is a finite family of subsets of V, called *edges*. It is *k-uniform* if every edge contains precisely k vertices. The *degree* $deg(x)$ is the number of edges containing x and the *codegree* $codeg(x, y)$, for two distinct vertices x and y, is the number of edges containing both x and y. The hypergraph is *D-regular* if the degree of each of its vertices is D and it is *simple* if the codegree of every pair of vertices is at most 1. A *matching* (or *packing*) in a hypergraph is a collection of pairwise disjoint edges.

The "semi-random" method was initiated in [1] and led to the pioneering work of Rödl [12] in which he introduced his "nibble" technique. Extensions of his result by various researchers including Frankl and Rödl [5] and Pippenger and Spencer [11] followed. One of the early extensions is Pippenger's result that for every fixed k and $\epsilon > 0$ there is some $\delta = \delta(k, \epsilon) > 0$, and $D_0 = D_0(k, \epsilon)$ such that any k-uniform hypergraph on N vertices in which all the degrees are between $(1-\delta)D$ and D, where $D > D_0$, and all the codegrees are at most δD contains a matching covering all vertices but at most ϵN. The proof does not supply, however, a good bound for the dependence of D_0 and δ on

[*] AT & T Bell Labs, Murray Hill, NJ 07974, USA and Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: noga@math.tau.ac.il. Research supported in part by a USA-Israel BSF grant.

[†] AT & T Bell Labs, Murray Hill, NJ 07974, USA. Email: jhkim@research.att.com.

[‡] Department of Mathematics and Computer Science, Courant Institute, NYU, NY, USA. Email:spencer@cs.nyu.edu. Research supported in part by a USA-Israel BSF Grant.

ϵ, and such a bound is desirable in various applications, such as the one mentioned in Corollary 1.2 below.

In the present paper we restrict our attention to the special case of regular, simple hypergraphs. For this case, a recent result of Grable [7] implies that for any $\epsilon > 0$, any k-uniform, D-regular hypergraph on N vertices contains a matching covering all vertices but at most $O(N(D/\ln N)^{-1/(2k-1+\epsilon)})$, provided D is sufficiently large as a function of k and ϵ. Our main result is the following stronger result.

Theorem 1.1 *Let (V, H) be a simple k-uniform, D-regular hypergraph on N vertices. If $k > 3$ there exists a matching P covering all vertices but at most*

$$O(ND^{-1/(k-1)})$$

If $k = 3$ there exists a matching P covering all vertices but at most

$$O(ND^{-1/2}\ln^{3/2} D)$$

Here and throughout the constants implicit in O and o-type notations depend only on k and are uniformly bounded as $N, D \to \infty$. The proof of the main result (and of a slightly stronger version of it, that merely assumes the hypergraph is nearly regular) is described in the next section. Our central idea is to take, repeatedly, a "random bite" out of (V, H). This differs from a "nibble" in that it removes a positive proportion of V, and contains a mechanism that keeps the remaining hypergraph nearly regular. The details require some careful analysis, described in Section 2. One key to the analysis is the use of *martingales* to bound large deviations. The martingale inequality used is given in Section 3.

Several simple corollaries of the main result, and some related questions are described in the final Section 4. Here we only state one such result.

A *Steiner Triple System* is a 3-uniform, simple hypergraph in which every pair of vertices is contained in precisely one edge. Improving previous results by various authors (see [10], [15] [16]), Brouwer [3] proved that any Steiner Triple System on n vertices contains a matching which covers all vertices but at most $5n^{2/3}$. Since a Steiner Triple System is clearly $d = (n-1)/2$-regular, Theorem 1.1 supplies the following asymptotic improvement.

Corollary 1.2 *Any Steiner Triple System on n vertices contains a matching covering all vertices but at most $O(n^{1/2}\ln^{3/2} n)$.*

2 Proof of the main result

We first define a *bite* without reference to randomness.

Bite. A bite is determined by a choice of edges $X \subseteq H$, called chosen edges, and a choice of vertices $W \subseteq V$ called wasted vertices. Given X, W we set

$$M = \{E \in X : E \cap E' = \emptyset \text{ for all other } E' \in X\}$$

and call all $E \in M$ isolated edges. These are clearly disjoint. We then set

$$V^* = V - \bigcup M - W$$

and

$$H^* = H|_{V^*},$$

where $H|_{V^*}$ denotes the induced subhypergraph of H on V^*.

The Big Picture. Begin with $(V, H) = (V_0, H_0)$. At stage i given (V_i, H_i) we take an appropriate bite X_i, W_i giving M_i and a new $(V_{i+1}, H_{i+1}) = (V^*, H^*)$. We continue until we reach some (V_ω, H_ω) when we quit. Then $P = \cup_{i<\omega} M_i$ is our packing. All vertices not in $\bigcup P$ are either in V_ω or in some W_i so

$$|V - \bigcup P| \leq |V_\omega| + \sum_{i<\omega} |W_i|$$

(A $w \in W_i$ might well be in $\bigcup M_i$ so this is an overcount.) Because the bite has to be iterated the difficulty lies in giving conditions on an intermediate value of (V_i, H_i) so that we can take a "good-sized bite" and still have (V_{i+1}, H_{i+1}) meet those same conditions. The exact statement is given in Theorem 2.1 below.

Waste. Why, you might well ask, should we ever "waste" vertices $w \in V$. The original (V, H) is regular but as we take successive random bites the values of $\deg(x)$ (in (V_i, H_i)) have a tendency to get further and further apart. The (suitably random) wastage W is a stabilization mechanism, designed to keep the values $\deg(x)$ reasonably close together.

Degrees. Returning to the bite of (V, H) we define, for $x \in V^*$, $\deg^*(x)$ to be the degree of x in V^*, the number of $E \in H^*$ with $x \in H^*$. For analysis we consider $\deg^*(x)$ to be the number of $E \in H$ with $x \in E$ such that $E - \{x\} \subseteq V^*$. This is defined for all $x \in V$ and the definitions coincide when $x \in V^*$.

Limited Effect. Suppose smallbite and bigbite were identical except for one $E \in H$, chosen in bigbite but not in smallbite. If E is isolated in bigbite then it is in M for bigbite but not smallbite. On the other side, if $E' \cap E \neq \emptyset$ and E' is isolated in smallbite then E' is in M for smallbite but not bigbite. This can hold for at most k sets E', one for each vertex of E, as two E' sharing a common vertex with E cannot both be isolated. Thus M differs in at most $k = O(1)$ places between smallbite and bigbite. Hence $|V^*|$ differs by at most $k(k+1) = O(1)$. For any $x \in V$ the sets $E - \{x\}$, $x \in E \in H$ are mutually disjoint as, critically, H is simple. Thus the values $\deg^*(x)$ differ by at most $k(k+1) = O(1)$ between smallbite and bigbite. More easily, now suppose smallbite and bigbite differ only in that $x \in W$ for bigbite but not for smallbite. This affects V^* by at most one vertex (x itself) and $\deg^*(y)$ by at most one and then only for those y sharing an edge with x.

Random Bite. Let (V, H) be given along with parameters D, Δ such that

$$D - \Delta \leq \deg(x) \leq D$$

for all $x \in V$. We take X random with

$$\Pr[E \in X] = \frac{1}{D}$$

for all $E \in H$, the events $E \in X$ being mutually independent. For $v \in V$ set

$$p(v) = \Pr[v \in E \text{ for some } E \in M]$$

This is given precisely by

$$p(v) = \sum_{v \in E} \Pr[E \in M] = \sum_{v \in E} \frac{1}{D}(1 - \frac{1}{D})^{f(E)}$$

where $f(E)$ is defined as the number of $E' \in H$ overlapping E. As H is simple no E' overlaps E in more than one vertex so

$$f(E) = \sum_{v' \in E} (\deg(v') - 1)$$

Define p^* as the maximal $p(v)$ over all $v \in V$. Now let W be a random subset of V with the events $w \in V$ mutually independent and

$$\Pr[w \in W] = c(w)$$

(c for compensatory) where $c(w)$ is that value satisfying

$$p(w) + c(w) - p(w)c(w) = p^*$$

With the probability space now defined $X, M, \deg^*(x), V^*, H^*$ become random variables. Our definition assures

$$\Pr[w \in V^*] = 1 - p(w) - c(w) + p(w)c(w) = 1 - p^*$$

precisely for all $w \in V$. That, indeed, is the purpose of the wastage W. Note a subtle point, that the value $c(w)$ is determined by (V, H) and does not depend on the actual value of X. Hence the probability space defined by the random bite can be (and will be) considered as one generated from mutually independent events of the forms $E \in X$ and $w \in W$.

Theorem 2.1 *Let $k \geq 3$ be fixed. Then there exist K, D_{min} so that: Let (V, H) be a k-uniform simple hypergraph on N vertices such that $D - \Delta \leq \deg(x) \leq D$ for all $x \in V$ where $D \geq D_{min}$ and*

$$\Delta = K D^{1/2} \ln^{1/2} D$$

Then there exists a bite X, W with

$$|W| = O(N D^{-1/2} \ln^{1/2} D)$$

with (V^*, H^*) *having* N^* *vertices,* $D^* - \Delta^* \leq \deg^*(x) \leq D^*$ *for all* $x \in V^*$ *with*

$$N^* < N(1 - e^{-k})(1 + O(D^{-0.4}))$$

$$\frac{D^*}{D} = \left(\frac{N^*}{N}\right)^{k-1} (1 + O(D^{-0.4}))$$

and, critically,

$$\Delta^* = K(D^*)^{1/2} \ln^{1/2} D^*$$

with the same K.

The relationship between N^* and D^* could certainly be tightened but it will suffice for our purposes. It is the near-regularity of (V^*, H^*) that will be most important. We shall show that with positive probability the random bite X, W defined above satisfies these conditions.

Little Compensation. We first note that all $f(E) = kD + O(\Delta)$ so that all

$$p(v) = (D + O(\Delta))\frac{1}{D}(1 - \frac{1}{D})^{kD+O(\Delta)} = (1 + O(\Delta/D))e^{-k} = (1 + O(D^{-1/2} \ln^{1/2} D))e^{-k}$$

so that

$$p^* = (1 + O(D^{-1/2} \ln^{1/2} D))e^{-k}$$

(in particular, $p^* = e^{-k} + o(1)$) and the compensatory

$$c(w) = O(D^{-1/2} \ln^{1/2} D)$$

for all $w \in V$. Now

$$E[|W|] = \sum_w c(w) = O(ND^{-1/2} \ln^{1/2} D)$$

and $|W|$ is the sum of mutually independent indicator random variables. We apply Chernoff bounds (see, e.g., the appendix of [2]) to give

$$\Pr[|W| > c_1 ND^{-1/2} \ln^{1/2} D] = e^{-\Omega(ND^{-1/2} \ln^{1/2} D)} \tag{1}$$

for an appropriately large constant c_1.

Vertices. We have precisely

$$E[|V^*|] = \sum_v \Pr[v \in V^*] = (1 - p^*)N$$

Our probability space is determined by at most ND/k choices of whether $X \in E$ and N choices of whether $w \in W$. The first type is yes with probability D^{-1}, the second with probability $O(D^{-1/2} \ln^{1/2} D)$. By limited effect, each choice can change $|V^*|$ by only a constant. The Martingale Inequality of Section 3 then gives

$$\Pr[|\ |V^*| - (1 - p^*)N\ | > \alpha N^{1/2}] < e^{-\Omega(\alpha^2)}$$

for $\alpha = o(N^{1/2})$. We set $\alpha = N^{1/2}D^{-1}$ (not needing the tightest result here) so that

$$\Pr[|\ |V^*| - (1-p^*)N\ | > ND^{-1}] < e^{-\Omega(ND^{-2})} \tag{2}$$

It is worth noting that $\alpha = N^{1/2}D^{-0.4}$ would also suffice here, but (2) will do as well.

Now we turn to degrees. Fix $x \in V$ and set $Z = \deg^*(x)$. By Linearity of Expectation

$$E[Z] = \sum_{x \in E} \Pr[E - \{x\} \subseteq V^*] \tag{3}$$

Claim 1: Fix $E \in H$ containing x. Then

$$\Pr[E - \{x\} \subseteq V^*] = (1 + O(D^{-1}))(1-p^*)^{k-1}$$

We set $E^- = E - \{x\}$ for convenience. As $\Pr[y \in V^*] = 1 - p^*$ we need to show the events $y \in V^*$ are "nearly" $(k-1)$-wise independent. We reach this in stages.

Suppose $1 \le l \le k-1$ and let $\prod, \sum, \wedge$ be over $1 \le i \le l$. First let $F_1, \ldots, F_l \in H$ be disjoint. Then $\prod \Pr[F_i \in M]$ and $\Pr[\wedge F_i \in M]$ differ only in that $\Pr[F \notin X]$ is multiply counted in $\prod$ for the $O(1)$ edges F overlapping two or more of the F_i. Thus

$$\Pr[\wedge F_i \in M] = (1 + O(D^{-1})) \prod \Pr[F_i \in M]$$

Now let $y_1, \ldots y_l \in E^-$, and compare

$$\prod p(y_i) = \sum_{y_i \in F_i} \prod \Pr[F_i \in M],$$

where the sum ranges over all l-tuples of edges F_i satisfying $y_i \in F_i$, with

$$\Pr[\wedge y_i \in \bigcup M] = \Pr[E \in M] + \sum_{y_i \in F_i}^{*} \Pr[\wedge F_i \in M]$$

where $\sum^*$ is over all choices of l disjoint edges F_i with $y_i \in F_i$. Then $\sum^*$ is missing $O(D^{l-1})$ of the $\sim D^l$ terms $(F_1, \ldots, F_l)$ with $y_i \in F_i$ of $\sum$. The terms of $\sum$ are comparable. Deleting $O(D^{l-1})$ terms changes it by a $1 + O(D^{-1})$ factor, the change from $\prod \Pr$ to $\Pr[\wedge]$ is another $1 + O(D^{-1})$, and $\Pr[E \in M] = O(D^{-1})$ is a third $1 + O(D^{-1})$ factor so that

$$\prod[\wedge y_i \in \bigcup M] = (1 + O(D^{-1})) \prod p(y_i)$$

Let further $z_1, \ldots z_s \in E^-$ be distinct (though possibly $z_j = y_i$). Then

$$\Pr[\wedge y_i \in \bigcup M \wedge \wedge z_j \in W] = (1 + O(D^{-1})) \prod p(y_i) \times \prod c(z_j)$$

as $\Pr[z_j \in W] = c(z_j)$ and this event is independent of all other choices. As this product is at most one we can rewrite

$$\Pr[\wedge y_i \in \bigcup M \wedge \wedge z_j \in W] = \prod p(y_i) \times \prod c(z_j) + O(D^{-1})$$

The event $E^- \subseteq V^*$ may be written as a Boolean combination of the "atomic" events $y \in \bigcup M$ and $z \in W$ over all $y, z \in E^-$. By Inclusion-Exclusion $\Pr[E^- \subseteq V^*]$ can be written as the sum and difference (with $3^{k-1} = O(1)$ terms) of probabilities of conjunctions of atomic events and so the probability is within an additive $O(D^{-1})$ of what it would be if the atomic events were independent. But in that case we *would* have $\Pr[E^- \subseteq V^*] = \prod \Pr[y \in V^*] = (1-p^*)^{k-1}$. Thus $\Pr[E^- \subseteq V^*] = (1-p^*)^{k-1} + O(D^{-1})$. As $(1-p^*)^{k-1} \sim (1-e^{-k})^{k-1}$ is bounded away from zero we can rewrite this in the form on the Claim. □

From Claim 1 and (3) $Z = \deg^*(x)$ has

$$E[Z] = \deg(x)(1-p^*)^{k-1}(1 + O(D^{-1}))$$

so the bounds on $\deg(x)$ of Theorem 2.1 give from above

$$D(1-p^*)^{k-1} + O(1) \geq E[Z]$$

and from below

$$E[Z] \geq D(1-p^*)^{k-1} - (K(1-p^*)^{k-1} + o(1))D^{1/2}\ln^{1/2} D$$

This is an important savings, the discrepancy Δ in the degrees in (V, H) is multiplied by $(1-p^*)^{k-1} \sim (1-e^{-k})^{k-1}$ in going to (V^*, H^*). This is counterbalanced by the new variation in the degrees due to the randomness, which we now address.

Claim 2: For $\alpha = o(D^{1/2})$

$$\Pr[|Z - E[Z]| > \alpha D^{1/2}] < e^{-\Omega(\alpha^2)}$$

Call $E \in H$ primary if it contains x, secondary if it overlaps a primary edge, tertiary if it overlaps a secondary edge. In case of multiple designation take the earliest. Z is determined by these edges and W. In the sense of Section 3 we give a strategy for "Paul" to determine Z. Paul first asks "Carole" which primary and secondary E are in X. Then call a secondary E important if $E \in X$ and no other overlapping secondary E is in X. There are only $O(D)$ such E, at most one containing each y sharing an edge with x. Paul then asks if $E' \in X$ for each tertiary E' that overlaps an important E. There are only $O(D^2)$ such E'. Now which secondary E are in M is determined. Finally Paul asks if $y \in W$ for each of the $O(D)$ vertices y sharing an edge with x. Now Z is determined precisely. By Limited Effect each query has effect $O(1)$. There were $O(D^2)$ queries, each yes with probability D^{-1} followed by $O(D)$ queries each yes with probability $O(D^{-1/2}\ln^{1/2} D)$. Claim 2 then follows from the Martingale Inequality of Section 3.

Remark. Querying only $O(D^2)$ instead of all $O(D^3)$ of the tertiary E' above was a critical savings. This well illustrates the power of the Martingale Inequality. Paul's queries on the tertiary E' do depend on Carole's responses on the secondary E.

We take $\alpha = c\ln^{1/2} D$ with c large so that, say,

$$\Pr[|Z - E[Z]| > cD^{1/2}\ln^{1/2} D] < D^{-20}$$

Note that c does not depend on K. We combine this with the always true bounds on $E[Z]$. Let $BAD(x)$ be the event that either

$$\deg^*(x) > D(1-p^*)^{k-1} + (c+1))D^{1/2}\ln^{1/2} D$$

or

$$\deg^*(x) < D(1-p^*)^{k-1} - [K(1-p^*)^{k-1} + c + 1]D^{1/2}\ln^{1/2} D$$

Then

$$\Pr[BAD(x)] < D^{-20}$$

The event $BAD(x)$ depends only on primary, secondary and tertiary edges and vertices on primary edges. We define a dependency graph on V, making x, y adjacent if there is a link $x = x_0, x_1, x_2, x_3, x_4, x_5, x_6 = y$ (or shorter) with x_i, x_{i+1} sharing an edge. The dependency graph has degree $O(D^6)$ and $BAD(x)$ is mutually independent of all $BAD(y)$ with y not adjacent to x. As $O(D^6)D^{-20} < \frac{1}{4}$ the conditions of the Lovász Local Lemma (cf., e.g., [2]) apply and

$$\Pr[\wedge\overline{BAD(x)}] > \prod[1 - 2\Pr[BAD(x)]] > e^{-3ND^{-20}}.$$

This probability is much larger (!) than the large deviation probabilities (1),(2) for $|W|$ and $|V^*|$. Thus with positive probability the random bite satisfies

$$\neg BAD(x) \text{ for all } x \in V$$

$$|W| = O(ND^{-1/2}\ln^{1/2} D)$$

$$||V^*| - N(1-p^*)| < ND^{-1}$$

Hence there *exists* a bite with these properties. Of course, we only really need $\neg BAD(x)$ for $x \in V^*$.

Lets check that this bite satisfies Theorem 2.1. We set

$$D^* = D(1-p^*)^{k-1} + (c+1)D^{1/2}\ln^{1/2} D$$

$$N^* = |V^*| = N(1-p^*)(1 + O(D^{-1}))$$

so that both

$$N^* = N(1-e^{-k})(1 + O(D^{-0.4}))$$

and

$$\frac{D^*}{D} = \left(\frac{N^*}{N}\right)^{k-1}(1 + O(D^{-0.4}))$$

with room to spare. Now we may take

$$\Delta^* = [K(1-p^*)^{k-1} + 2c + 2]D^{1/2}\ln^{1/2} D$$

As $D^* \sim D(1-p^*)^{k-1}$ we have

$$\frac{\Delta^*}{(D^*)^{1/2}\ln^{1/2} D^*} \sim [K(1-p^*)^{k-1}+2c+2](1-p^*)^{-(k-1)/2}$$

Having chosen c and noting $p^* = e^{-k}+o(1)$ we pick K to be a constant (dependent only on k) so that

$$[K(1-e^{-k})^{k-1}+2c+2](1-e^{-k})^{-(k-1)/2} < K$$

and the proof of Theorem 2.1 is completed. □

Now we prove Theorem 1.1. Begin with $(V,H) = (V_0,H_0)$ with N, D. Let K be the constant of Theorem 2.1. The conditions of Theorem 2.1 certainly apply as (V,H) is precisely D-regular. Apply it repeatedly giving M_i, W_i and (V_{i+1},H_{i+1}) until reaching (V_ω,H_ω) with $D_\omega < D_{min}$ so that Theorem 2.1 no longer applies. We have

$$\frac{D_{i+1}}{D_i} = \left(\frac{N_{i+1}}{N_i}\right)^{k-1}(1+O(D_i^{-0.4}))$$

As the D_i are dropping geometrically the product of the $1+O(D_i^{-0.4})$ terms is bounded from below and above by positive constants so that multiplying over $0 \le i < \omega$

$$\frac{D_\omega}{D} = \Theta\left((\frac{N_\omega}{N})^{k-1}\right)$$

and since we continue until $D_\omega = \Theta(1)$ the process stops with

$$|V_\omega| = N_\omega = \Theta(ND^{-1/(k-1)})$$

The "wasted" vertices have

$$|W_i| = O(N_i D_i^{-1/2}\ln^{1/2} D_i)$$

Changing from i to $i+1$ N is multiplied by roughly $\theta = 1-e^{-k}$ and D by roughly θ^{k-1} while $\ln D$ remains asymptotically the same. Thus the bound on W_i is multiplied by $\theta^{1-(k-1)/2}$.

Now (finally!) we separate the $k>3$ and $k=3$ cases. Suppose $k>3$. Then $1-(k-1)/2<0$ and $\theta<1$ so the bounds on $|W_i|$ increase geometrically and hence $\sum_{i<\omega}|W_i|$ is bounded by a constant times the bound on $|W_{\omega-1}|$. But this is at most the total number $N_{\omega-1}$ of vertices at this penultimate stage which is at most a constant times the number of vertices N_ω at the end. So

$$\sum_{i<\omega}|W_i| = O(N_\omega) = O(ND^{-1/(k-1)})$$

giving the desired bound on $|V-\bigcup P|$. Now suppose $k=3$. We get (V_i,H_i) with $|V_\omega| = O(ND^{-1/2})$. But now

$$|W_i| = O\left(\frac{N_i}{D_i^{1/2}}\ln^{1/2} D_i\right) = O(\frac{N}{D^{1/2}}\ln^{1/2} D_i)$$

as $N_iD_i^{-1/2}$ remains near-constant. The D_i drop geometrically. Working backwards from ω, $\ln D_{\omega-i}$ grows linearly in i up to $i=\omega$. Then

$$\sum \ln^{1/2} D_i = \Theta(\sum_{j<\omega} j^{1/2}) = \Theta(\omega^{3/2})$$

As the D_i drop geometrically from D there are $\omega = \Theta(\ln D)$ steps so the total wastage is

$$O(ND^{-1/2}\ln^{3/2} D).$$

This dominates $|V_\omega|$ and so gives the bound for $|V - \bigcup P|$. This completes the proof of Theorem 1.1. $\square$

3 Martingale Inequalities

Martingales have become in recent years a very powerful tool for the Probabilistic Method. A general discussion is given in [2]. The junior author's breakthrough [9] as well as several previous papers (see, e.g., [8]) illustrated how to make use of small probabilities. Still, both [2] and [9] deal with an essentially static case in which the order of the martingale exposures has no relevance. Here we give a quite general inequality which also precisely suits our purposes. A similar inequality appears in Lemma 3.4 of [8]. Our presentation is entirely self-contained, though a rereading of [2] might help with motivation.

We assume our underlying probability space is generated by a finite set of mutually independent Yes/No choices, indexed by $i \in I$. In our case the choices are of the forms $E \in X$ and $w \in W$. We are given a random variable Y on this space, in our cases $|V^*|$ and $\deg^*(x)$. Let p_i denote the probability that choice i is Yes. Let c_i be such that changing choice i (keeping all else the same) can change Y by at most c_i. We call c_i the *effect* of i. Let C be an upper bound on all c_i. In our cases $C = O(1)$, what we called Limited Effect. We call $p_i(1-p_i)c_i^2$ the *variance* of choice i.

Now consider a solitaire game in which Paul finds the value of Y by making queries of an always truthful oracle Carole. The queries are always of a choice $i \in I$. Paul's choice of query can depend on Carole's previous responses. A strategy for Paul can then naturally be represented in a decision tree form. A "line of questioning" is a path from the root to a leaf of this tree, a sequence of questions and responses that determine Y. The total variance of a line of questioning is the sum of the variances of the queries in it.

Martingale Inequality. For all $\epsilon > 0$ there exists $\delta > 0$ so that the following holds. Suppose Paul has a strategy for finding Y such that every line of questioning has total variance at most σ^2. Then

$$\Pr[|Y - E[Y]| > \alpha\sigma] \leq 2e^{-\frac{\alpha^2}{2(1+\epsilon)}} \tag{4}$$

for all positive α with $\alpha C < \sigma(1+\epsilon)\delta$.

We apply this inequality in this paper only when $C = O(1)$ and $\alpha = o(\sigma)$ so that the side condition on α does hold and the upper bound on the tail distribution can be written as $2\exp[-\Omega(\alpha^2)]$. When a specific (though not best possible) bound is desired we note that $\epsilon = \delta = 1$ meet these conditions. The term variance and letter σ are deliberately suggestive and indeed it can be shown that σ^2 is an upper bound on the variance of Y.

For simplicity we replace Y by $Y - E[Y]$ so that we shall henceforth assume $E[Y] = 0$. By symmetry we shall bound only the upper tail of Y. We set, with foresight, $\lambda = \alpha/[\sigma(1+\epsilon)]$. Our side assumption gives that $C\lambda < \delta$. We will show

$$E[e^{\lambda Y}] \leq e^{(1+\epsilon)\lambda^2\sigma^2/2} \tag{5}$$

From this the Martingale Inequality follows by the Markov bound

$$\Pr[Y > \alpha\sigma] < e^{-\lambda\alpha\sigma}E[e^{\lambda Y}] < e^{-\alpha^2/2(1+\epsilon)}$$

by our (optimal) choice of λ.

We first claim that for all $\epsilon > 0$ there exists $\delta > 0$ so that for $0 \leq p \leq 1$ and $|a| \leq \delta$

$$pe^{(1-p)a} + (1-p)e^{-pa} \leq e^{(1+\epsilon)p(1-p)a^2/2} \tag{6}$$

Take the Taylor Series in a of the left hand side. The constant term is 1, the linear term 0, the coefficient of a^2 is $\frac{1}{2}p(1-p)$ and for $j \geq 3$ the coefficient of a^j is $\frac{1}{j!}p(1-p)[p^{j-2}+(1-p)^{j-2}] \leq \frac{1}{j!}p(1-p)$. Pick δ so that $|a| \leq \delta$ implies

$$\sum_{j=2}^{\infty} \frac{a^j}{j!} < \frac{a^2}{2}(1+\epsilon)$$

and note that in particular this holds for $\epsilon = \delta = 1$. Then

$$pe^{(1-p)a} + (1-p)e^{-pa} \leq 1 + p(1-p)\frac{a^2}{2}(1+\epsilon)$$

and (6) follows from the inequality $1 + x \leq e^x$.

Using this δ we show (5) by induction on the depth M of the decision tree. For $M = 0$ Y is constant and (5) is immediate. Otherwise, let $p, c, v = p(1-p)c^2$ denote the probability, effect and variance respectively of Paul's first query. Let μ_y, μ_n denote the conditional expectations of Y if Carole's response is Yes or No respectively. Then $0 = E[Y]$ can be split into

$$0 = p\mu_y + (1-p)\mu_n$$

The difference $\mu_y - \mu_n$ is the expected *change* in Y when all other choices are made independent with their respective probabilities and the root choice is changed from Yes to No. As this always changes Y by at most c

$$|\mu_y - \mu_n| \leq c$$

Thus we may parametrize

$$\mu_y = (1-p)b \text{ and } \mu_n = -pb$$

with $|b| \leq c$. From (6)

$$pe^{\lambda\mu_y} + (1-p)e^{\lambda\mu_n} \leq e^{(1+\epsilon)p(1-p)b^2\lambda^2/2} \leq e^{(1+\epsilon)v\lambda^2/2}$$

Let A_y denote the expectation of $e^{\lambda(Y-\mu_y)}$ conditional on Carole's first response being Yes and let A_n denote the analogous quantity for No. Given Carole's first response Paul has a decision tree (one of the two main subtrees) that determines Y with total variation at most $\sigma^2 - v$ and the tree has depth at most $M-1$. So by induction $A_y, A_n \leq A^-$ where we set

$$A^- = e^{(1+\epsilon)\lambda^2(\sigma^2-v)/2}$$

Now we split

$$E[e^{\lambda Y}] = pe^{\lambda\mu_y}A_y + (1-p)e^{\lambda\mu_n}A_n \leq [pe^{\lambda\mu_y} + (1-p)e^{\lambda\mu_n}]A^- \leq e^{(1+\epsilon)\lambda^2(v+(\sigma^2-v))/2}$$

completing the proof of (5) and hence of the Martingale Inequality.

We remark that this formal inductive proof somewhat masks the martingale. A martingale $E[Y] = Y_0, \ldots, Y_M = Y$ can be defined with Y_i the conditional expectation of Y after the first i queries and responses. The Martingale Inequality can be thought of as bounding the tail of Y by that of a normal distribution of greater or equal variance. For very large distances from the mean, large α, this bound fails.

4 Concluding remarks and open problems

A *Partial Steiner System* $S(t,k,n)$ is a k-uniform hypergraph on n vertices so that every set of t vertices is contained in at most one edge. Such a system is called a (full) *Steiner System* $S(t,k,n)$ if every subset of t vertices is contained in precisely one edge. Thus, in particular, a Steiner Triple System on n vertices is an $S(2,3,n)$. Simple counting shows that the number of edges in any partial Steiner System $S(t,k,n)$ cannot exceed

$$\frac{\binom{n}{t}}{\binom{k}{t}},$$

and equality holds if and only if the partial system is in fact a full Steiner System.

In his original paper [12], Rödl developed the nibble approach in order to prove the conjecture of Erdös and Hanani [4], which asserts that for every fixed $t < k$ there exists a partial $S(t,k,n)$ with at least

$$(1-o(1))\frac{\binom{n}{t}}{\binom{k}{t}}$$

edges, where the $o(1)$-term tends to 0 as n tends to infinity. For the case $t = k-1$, an easy corollary of Theorem 1.1 is the following result, which provides an effective estimate for the $o(1)$-term.

Corollary 4.1 *For every $k > 3$ there exists a partial $S(k-1, k, n)$ with at most $O(n^{k-1-\frac{1}{k-1}})$ uncovered $(k-1)$-tuples.*

Proof. Apply Theorem 1.1 to the k-uniform hypergraph whose vertices are all $(k-1)$-tuples of a set of size n, and whose edges are all collections of k $(k-1)$-tuples contained in a k-set. This is a simple hypergraph with $N = \binom{n}{k-1}$ vertices, which is $D = (n-k)$-regular, and the result thus follows from Theorem 1.1. □

The above estimate, however, can be improved by an explicit, well known construction, showing that there is a partial system in which the number of uncovered $(k-1)$-tuples is at most $\frac{k-1}{n}\binom{n}{k-1}$ ($= O(n^{k-2})$). Indeed, simply let the partial system be the family of all k-subsets of $\{0, 1, 2, \ldots, n-1\}$ the sum of whose elements is r modulo n, where r is chosen so that the number of these k-subsets is at least $\binom{n}{k}/n$. A simple calculation implies the above claim.

Another simple application of Theorem 1.1 is the following companion to Corollary 1.2.

Corollary 4.2 *For any fixed $k > 3$, any (full) Steiner System $S(2, k, n)$ contains a matching covering all vertices but at most $O(n^{1-1/(k-1)})$.*

The proof of Theorem 1.1 as well as previous results including [12], [14], suggest the following *random greedy* algorithm for constructing a large matching in a D-regular, simple, k-uniform hypergraph on N vertices (where $k \geq 3$ is fixed and D, N are large). Choose a random order of all the edges of the hypergraph, and starting with the empty matching, scan the edges one by one according to this random order, adding each one in its turn to the matching iff it does not intersect any of the edges already in the matching. We conjecture that the expected number of vertices not covered by the matching this algorithm produces is $O(ND^{-1/(k-1)}(\ln N)^{O(1)})$. This remains open. See [14], [13] for some related results. When applied to the hypergraph described in the proof of Corollary 4.1, the above algorithm is the following procedure for constructing a large partial $S(k-1, k, n)$. Let $K_1, K_2, \ldots, K_m$ be a random order of all the $m = \binom{n}{k}$ k-subsets of $\{1, 2, \ldots, n\}$. Starting with the empty system, scan the k-tuples one by one, and pick each K_i in its turn to be a member of the system iff it does not cover any $(k-1)$-subset which is already covered by one of the k-tuples picked so far. Simulation results discussed in [6] supply strong evidence that the expected number of uncovered $(k-1)$-tuples in the partial system this algorithm produces is $n^{k-1-1/(k-1)+o(1)}$, but at the moment we cannot prove that this is the case even for $k = 3$.

Finally, it would be interesting to decide if the assertion of Theorem 1.1 is tight. It is not difficult to give examples of D-regular, k-uniform hypergraphs on N vertices in which any matching misses at least $\Omega(N/D)$ vertices, but this is far from the $O(N/D^{1/(k-1)})$ upper bound proved in the theorem.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, *A dense infinite Sidon sequence*, Europ. J. Combinatorics 2 (1981), 1-11.

[2] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, 1992.

[3] A. E. Brouwer, *On the size of a maximum transversal in a Steiner Triple System*, Can. J. Math. 33 (1981), 1202-1204.

[4] P. Erdös and H. Hanani, *On a limit theorem in combinatorial analysis*, Publ. Math. Debrecen 10 (1963), 10-13.

[5] P. Frankl and V. Rödl, *Near perfect coverings in graphs and hypergraphs*, Europ. J. Combinatorics 6 (1985), 317-326.

[6] D. M. Gordon, G. Kuperberg, O. Patashnik and J. Spencer, *Asymptotically Optimal Covering Designs*, Preprint, 1995.

[7] D. A. Grable, *More-than-nearly perfect packings and partial designs*, Preprint, 1995.

[8] J. Kahn, *Asymptotically good list-colorings*, J. Combinatorial Theory, Ser. A 73 (1996), 1–59.

[9] J.H. Kim, *The Ramsey Number $R(3,t)$ has order of magnitude $t^2/\ln t$*, Random Structures & Algorithms 7 (1995), 173-207.

[10] C. C. Lindner and K. T. Phelps, *A note on partial parallel classes in Steiner systems*, Discrete Math. 24 (1978), 109-112.

[11] N. Pippenger and J. H. Spencer, *Asymptotic behaviour of the chromatic index for hypergraphs*, J. Combinatorial Theory, Ser. A 51 (1989), 24-42.

[12] V. Rödl, *On a packing and covering problem*, Europ. J. Combinatorics 6 (1985), 69-78.

[13] V. Rödl and L. Thoma, *Asymptotic packing and the random greedy algorithm*, Random Structures & Algorithms (to appear)

[14] J. Spencer, *Asymptotic packing via a branching process*, Random Structures & Algorithms 7 (1995), 167-172

[15] S. P. Wang, *On self orthogonal Latin squares and partial transversals of Latin squares*, Ph. D. thesis, Ohio State Univ., Columbus, Ohio, 1978.

[16] D. E. Woolbright, *On the size of partial parallel classes in Steiner systems*, in: *Topics in Steiner Systems* (C. C. Lindner and A. Rosa, eds.), Annals of Discrete Math. 7 (1980), 203-211.