COMBINATORICA 7 (2) (1987) 151—160

THE SMALLEST n-UNIFORM HYPERGRAPH WITH POSITIVE DISCREPANCY

N. ALON, D. J. KLEITMAN, C. POMERANCE, M. SAKS and P. SEYMOUR

Received 15 March 1986

A two-coloring of the vertices X of the hypergraph $H=(X, \mathcal{E})$ by red and blue has *discrepancy* d if d is the largest difference between the number of red and blue points in any edge. A two-coloring is an equipartition of H if it has discrepancy 0, i.e., every edge is exactly half red and half blue. Let $f(n)$ be the fewest number of edges in an n-uniform hypergraph (all edges have size n) having positive discrepancy. Erdős and Sós asked: is $f(n)$ unbounded? We answer this question in the affirmative and show that there exist constants c_1 and c_2 such that

$$\frac{c_1 \log(\operatorname{snd}(n/2))}{\log\log(\operatorname{snd}(n/2))} \leqq f(n) \leqq c_2 \frac{\log^3(\operatorname{snd}(n/2))}{\log\log(\operatorname{snd}(n/2))}$$

where $\operatorname{snd}(x)$ is the least positive integer that does not divide x.

1. Introduction and Main Results

A number of recent papers have been concerned with the problem of two-coloring the vertices of a hypergraph $H=(X, \mathcal{E})$ by red and blue so that for every edge E of H the number of red points in E is roughly equal to the number of blue points. The *discrepancy* of a two-coloring is the maximum difference between the number of red points and blue points in any edge. The *discrepancy* of a hypergraph is the minimum discrepancy of any two-coloring. There have been several results relating discrepancy to other parameters of a hypergraph (number of vertices, maximum degree) and computing discrepancy for special classes of hypergraphs (see, e.g., [2], [3], [7], [9,] [10]).

The focus of this paper is hypergraphs of discrepancy zero, i.e. those that admit a two-coloring of the vertex set such that every edge is divided exactly in half. Such a coloring is called an *equi-partition* of H. For instance, the hypergraphs with $X=\{1, 2, ..., 2n\}$ and $\mathcal{E}$ the set of all intervals of even length is equi-partitioned by coloring the odd numbers red and the even numbers blue. We are interested in the function $f(n)$, defined to be the fewest number of edges in an n-uniform hypergraph (all edges have size n) that admits no equi-partition. Trivially for n odd, $f(n)=1$. Note also that $f(n) \leqq n+1$ for any n since a hypergraph with $|X|=n+1$ and $\mathcal{E}$ consisting of all n element subsets of X cannot be equi-partitioned.

The first two authors were supported in part by NSF under grant DMS 8406100; the third one was supported in part by NSF under grant DCR 8421341.

Much stronger upper bounds on $f(n)$ for large classes of n are easily obtained by construction. For instance, if n is twice an odd number then $f(n) \leqq 3$; more generally, the following simple constructive result shows that $f(n)$ is small whenever some small number fails to divide n.

Proposition 1.1. *For n even, $f(n) \leqq 1 + \operatorname{snd}(n/2)$, where $\operatorname{snd}(x)$ is the smallest positive integer that does not divide x.*

Proof. Let k be the smallest non-divisor of $n/2$ and let $n = ak+r$ where $0 \leqq r < k$. Let $X_1, X_2, \ldots, X_{k+1}$ be disjoint sets such that $|X_1| = |X_2| = \ldots = |X_r| = a+1$ and $|X_{r+1}| = |X_{r+2}| = \ldots = |X_{k+1}| = a$. Let $X = X_1 \cup X_2 \cup \ldots \cup X_{k+1} \cup \{z\}$ where z is not a member of any X_i. Define an n-uniform hypergraph $H = (X, \mathscr{E})$ with edges $E_1, E_2, \ldots, E_{k+1}$ where $E_i = X - X_i$ if $1 \leqq i \leqq r$ and $E_i = X - X_i - z$ if $r < i \leqq k+1$. Suppose $X = R \cup B$ is an equi-partition of H; we derive a contradiction. Assume without loss of generality that $z \in R$. Then $|E_j \cap B| = n/2$ for all j and so $|X_j \cap B| = |X \cap B| - |E_j \cap B| = |X \cap B| - n/2$ is the same for all j. Then

$$n/2 = |E_j \cap B| = \sum_{i \neq j} |X_i \cap B| = k(|X \cap B| - n/2)$$

contradicting the hypothesis that k does not divide $n/2$. ∎

Lower bounds on $f(n)$ are not so easy to obtain. Indeed, this difficulty led Erdős and Sós [4] to pose the question: Is $f(n)$ unbounded? From Proposition 1.1, $f(n)$ is bounded on the sequence $n_1, n_2, \ldots$ if there exists a number k that divides none of the n_i. We prove the converse.

Theorem 1.2. *If $n_1, n_2, \ldots$ is a sequence of integers such that every integer k divides at least one of them then $\{f(n_i)\}$ is unbounded.*

In Section 2 we present the first of two proofs of Theorem 1.2. The proof is based on a limit argument and has the undesirable (yet intriguing) feature that it gives no information about the growth rate of $f(n)$ beyond the fact that it is unbounded. The remainder of the paper is devoted to the derivation of upper and lower bounds on $f(n)$. All of the bounds we obtain can be expressed in terms of the quantity $\operatorname{snd}(n/2)$. As stated above, $f(n) = 1$ if n is odd and $f(n) = 3$ if $n \equiv 2 \bmod 4$. Our result is:

Theorem 1.3. *There exist constants c_1 and c_2 such that for $n \equiv 0 \bmod 4$,*

$$c_1 \frac{\log \operatorname{snd}\left(\frac{n}{2}\right)}{\log\log \operatorname{snd}\left(\frac{n}{2}\right)} \leqq f(n) \leqq c_2 \frac{\log^3 \operatorname{snd}\left(\frac{n}{2}\right)}{\log\log \operatorname{snd}\left(\frac{n}{2}\right)}.$$

Observe that the lower bound in Theorem 1.3 immediately implies Theorem 1.2. Note also that the upper bound is a substantial improvement over Proposition 1.1.

The prime number theorem immediately gives $\operatorname{snd}(n/2) \leqq (1+o(1)) \log n$. For infinitely many n this result is best possible. (If n is twice the least common multiple of the numbers less than k then $\operatorname{snd}(n/2) > k = (1+o(1)) \log n$.) From this we can conclude.

Theorem 1.4. *There exist constants* $c_1, c_2 > 0$ *such that for all* $n \geqq 10$

$$f(n) \leqq c_2 \frac{(\log\log n)^3}{\log\log\log n}$$

and for infinitely many values of n

$$f(n) \geqq c_1 \frac{\log\log n}{\log\log\log n}.$$

The lower bound for $f(n)$ of Theorem 1.3 is obtained by considering the relationship of f to another function g defined as follows. A two-coloring $X = R \dot{\cup} B$ of an n-uniform hypergraph H is *uniform* if $|R \cap E|$ is the same for every edge in $\mathcal{E}$ (and is neither 0 nor n). In particular an equi-partition is a uniform coloring. A hypergraph that admits a uniform coloring is *reducible*, and otherwise it is *irreducible.* Let $g(n)$ be the fewest number of edges in an n-uniform hypergraph that is irreducible. Clearly $g(n) \geqq f(n)$. Let

$$\hat{g}(n) = \min_{m \geqq n} g(m).$$

If $g(n)$ is monotone then $\hat{g}(n) = g(n)$; we do not know whether this is the case. The main result of Section 3 is

Theorem 1.5. $f(n) \geqq \hat{g}\,(\text{snd}\,(n/2))$.

The function g has been studied extensively (in the literature the results are typically discussed in terms of the dual hypergraph; see [5] for a survey). The following bound was proved by Huckemann, Jurkat and Shapley (cf. [5]; see also [1] for an alternate proof).

Theorem 1.6. *If* $n \geqq (k+1)^{(k+1)/2}$ *then* $g(n) \geqq k$, *and so* $\hat{g}(n) \geqq c \log n/\log\log n$ *for some constant* c.

The lower bound of Theorem 1.3 is an immediate corollary of Theorems 1.5 and 1.6.

In Section 4, the upper bound of Theorem 1.3 is proved by a number theoretic argument. Section 5 presents some open questions.

2. Proof of Theorem 1.2.

Without loss of generality we can assume that for every integer k, k divides n_j for all $j \geqq k$. To see this replace the sequence $n_1, n_2, \ldots$ by the sequence whose k^{th} term is the first term of $\{n_i | i \geqq 1\}$ that is divisible by the least common multiple (lcm) of 1, 2, ..., k. Clearly, if $f(n)$ is unbounded on this sequence, it is unbounded on the original sequence. So let $n_1, n_2, \ldots$ be a sequence of integers such that n_k is divisible by all integers less than or equal to k and suppose $f(n)$ is bounded for all n_j. Let q be a bound. Then for each j, there exists an n_j-uniform hypergraph H^j with q edges that is not equi-partitionable. We will derive a contradiction by showing that for some j, H^j has an equi-partition.

Let $E_1^j, E_2^j, \ldots, E_q^j$ be the edges of H^j and for $I \subseteq \{1, 2, \ldots, q\}$ define A_I^j to be the set

$$\bigcap_{i \in I} E_i^j - \bigcup_{i \notin I} E_i,$$

that is, A_I^j is the set of vertices belonging to exactly the edges $\{E_i^j | i \in I\}$. Note that the sets A_I^j for $I \subseteq \{1, \ldots, q\}$ partition the vertex set of H. Defining the vector $\mathbf{a}^j = (a_I^j | I \subseteq \{1, 2, \ldots, q\})$ by setting $a_I^j = |A_I^j|$ we have that for each i,

$$n_j = |E_i^j| = \sum_{I | i \in I} a_I^j.$$

Now, a two-coloring $B \dot{\cup} R$ of the vertices of H^j is essentially defined by the numbers $b_I^j = |B \cap A_I^j|$. Using this correspondence we have that H^j has an equi-partition if and only if there exists an integer vector $\mathbf{b}^j = (b_I^j | I \subseteq \{1, \ldots, q\})$ such that

(2.1) $$0 \leqq b_I^j \leqq a_I^j$$

and for each $i \in \{1, \ldots, q\}$

(2.2) $$\sum_{I | i \in I} b_I^j = n_j/2.$$

We now prove that for some j there exists a vector $\mathbf{b}^j$ as above, so that H^j has an equi-partition. For each j, let $\mathbf{v}^j = \mathbf{a}^j / n_j$. Each vector $\mathbf{v}^j$ satisfies

$$\sum_{I | i \in I} v_I^j = 1$$

for $i = 1, 2, \ldots, q$. By the Bolzano—Weierstrass theorem ([8], pg. 35), there is an infinite subsequence $j_1, j_2, \ldots$ of integers such that the sequence of vectors $\mathbf{v}^{j_1}, \mathbf{v}^{j_2}, \ldots$ converges to some vector $\mathbf{v}^*$, which also satisfies $\sum_{I | i \in I} \mathbf{v}_I^* = 1$ for $i = 1, 2, \ldots, q$. Let $\mathbf{w}$ be a rational vector satisfying $(1/2)\mathbf{v}_I^* \leqq w_I \leqq (3/4) v_I^*$ for each I. Then the system

$$\sum_{I | i \in I} x_I = \frac{1}{2}, \quad i = 1, 2, \ldots, q$$

$$O \leqq x_I \leqq w_I, \quad I \subseteq \{1, 2, \ldots, q\}$$

has a solution, namely $\mathbf{x} = \mathbf{v}^*/2$. Since all of the inequalities and equalities of the system are given by rational coefficients, there must be a rational solution vector $\mathbf{r}$. Let k be the smallest integer such that $k\mathbf{r}$ is an integer vector. Now since $\mathbf{v}^{j_1}, \mathbf{v}^{j_2}, \ldots$ converges to $\mathbf{v}^*$ there is an index h such that $\mathbf{v}^{j_h} \geqq (3/4)\mathbf{v}^* \geqq \mathbf{w}$, and such that $j_h \geqq k$. Now the fact that $k | n_j$ for $j \geqq k$ implies that the vector $\mathbf{b}^{j_h} = n_{j_h} \mathbf{r}$ is integral, and $\mathbf{v}^{j_h} \geqq \mathbf{w}$ and the choice of $\mathbf{r}$ imply that (2.1) and (2.2) are satisfied and thus H^{j_h} has an equi-partition, a contradiction, establishing Theorem 1.2. ∎

3. Proof of Theorem 1.5.

Let $H=(X, \mathcal{E})$ be a hypergraph. For $Y \subseteq X$ let H_Y be the hypergraph with vertex set Y and edge set $\mathcal{E}_Y = \{E \cap Y | E \in \mathcal{E}\}$. Then we have

Lemma 3.1. *Let $H=\{X, \mathcal{E}\}$ be an n-uniform hypergraph and let k be an integer such that $|\mathcal{E}| < \hat{g}(k+1)$. Then X can be partitioned into sets $X_1, X_2, \dots, X_t$ such that H_{X_i} is n_i-uniform with $n_i \leqq k$ with every i.*

Proof. By induction on n. If $n \leqq k$ the result is trivial. If $n > k$ then since $|\mathcal{E}| < \hat{g}(k+1) \leqq g(n)$, H is reducible and so X can be partitioned into sets Y and Z so that H_Y and H_Z are each uniform with edge sizes less than n. Applying the induction hypothesis to H_Y and H_Z proves the lemma. ∎

Now let $k = \mathrm{snd}(n/2)-1$ and let $H=(X, \mathcal{E})$ be an n-uniform hypergraph with fewer than $\hat{g}(k+1)$ edges. We want to show that H has an equi-partition. By Lemma 3.1, X can be partitioned into sets $X_1, \dots, X_t$ such that H_{X_i} is n_i-uniform with $n_i < \mathrm{snd}(n/2)$. Note that $n = \sum_{i=1}^{t} n_i$. It suffices to find a set $I \subseteq \{1, \dots, t\}$ such that $\sum_{i \in I} n_i = n/2$, since then we get an equi-partition by letting $R = \bigcup_{i \in I} X_i$ and $B = \bigcup_{i \notin I} X_i$. Let a_j be the number of n_i's that equal j. Then $\sum_{j=1}^{k} j a_j = n$ and the existence of an equi-partition follows from:

Lemma 3.2. *Let n, k be positive integers such that $1, 2, \dots, k$ divide $n/2$. Let $a_1, a_2, \dots, a_k$ be nonnegative integers such that $\sum i a_i = n$. Then there exist integers $b_1, b_2, \dots, b_k$ with $0 \leqq b_j \leqq a_j$ such that*

$$\sum_{j=1}^{k} j b_j = n/2.$$

Proof. Let $A(n, k)$ be the set of all sequences satisfying the hypothesis and order the sequences lexicographically, i.e. $(a_1, \dots, a_k) > (a_1', \dots, a_k')$ if $a_j > a_j'$ where j is the smallest index in which the two sequences differ. Suppose the theorem is false and let $(a_1, \dots, a_k)$ be the lexicographically least counterexample. Then

$$a_i < n/2i \quad \text{for all } i \tag{3.1}$$

since if $a_j \geqq n/2j$ then $b_j = n/2j$, $b_i = 0$ for $i \neq j$ would satisfy the conclusion of the theorem. Furthermore

$$a_j \geqq k \quad \text{for at most one index } j. \tag{3.2}$$

For suppose, to the contrary, that $a_j, a_h \geqq k$ where $j \geqq h$. Define $(a_1', \dots, a_k')$ by $a_h' = a_h - j$, $a_j' = a_j + h$ and $a_i' = a_i$ for $i \neq h, j$. By the lexicographic minimality of $(a_1, \dots, a_k)$, there exist integers $(b_1', b_2', \dots, b_k')$ with $0 \leqq b_i' \leqq a_i'$ and $\sum i b_i' = n/2$. If $b_j' \leqq a_j$ then $b_i' \leqq a_i$ for all i a contradiction. On the other hand, if $b_j' > a_j$ then the sequence $(b_1, \dots, b_k)$ given by $b_j = b_j' - h$, $b_h = b_h' + j$ and $b_i = b_i'$ for $i \neq h, j$ satisfies the conclusion of the lemma, again a contradiction.

By (3.1) and (3.2) we have

$$n=\sum_{i=1}^{k} ia_i < n/2+(k-1)\sum_{i=1}^{k} i,$$

so

$$n < k^3-k. \tag{3.3}$$

However, $n \geqq 2lcm\,(1, 2, ..., k)$, so we have

$$\frac{1}{2}(k^3-k) > lcm\ (1, 2, ..., k). \tag{3.4}$$

This implies that $k=2, 3, 4$ or 6. (For $lcm(1, 2, ..., k)$ is the product of all maximal prime powers less than or equal to k. The largest power of p less than or equal to k is at least $(k+1)/p$ and thus

$$lcm\,(1, 2, ..., k) \geqq \frac{(k+1)^8}{2\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13\cdot 17\cdot 19}$$

which with (3.4) implies that $k \leqq 21$. Checking all values of k up to 21 yields $k= =2, 3, 4,$ or 6.)

For $k=2$ and $k=3$ the theorem is trivial. For $k=4$, (3.3) implies that $n<60$ so n is 24 or 48. For $k=6$, (3.3) implies that $n<210$ so $n=120$. In these cases, simple ad hoc arguments (left to the reader) show that the lemma holds. Thus $(a_1, ..., a_n)$ is not a counterexample and the lemma is true. ∎

4. An upper bound for $f(n)$

Here we prove the upper bound on $f(n)$ given by Theorem 1.3. We need some additional definitions. Let $\mathcal{M}$ denote the set of all matrices M with entries in $\{0, 1\}$ such that the equation $M\mathbf{x}=\mathbf{e}$ has exactly one nonnegative solution. (Here $\mathbf{e}$ is the vector with all entries equal to 1.) This unique solution is denoted $\mathbf{x}^M$. Let $d(M)$ be the least integer such that $d(M)\mathbf{x}^M$ is integral and let $\mathbf{y}^M=d(M)\mathbf{x}^M$. For each positive integer n, let $t(n)$ be the least r such that there exists a matrix $M\in\mathcal{M}$ with r rows such that $d(M)=n$. (For instance, the $(n+1)$ by $(n+1)$ matrix with 0's on the diagonal and 1's off the diagonal has $d(M)=n$, so $t(n)\leqq n+1$). Our upper bound on $f(n)$ is an immediate consequence of the following three results.

Theorem 4.1. *Let n be a natural number and m be an integer such that $\lfloor n/m \rfloor$ (the greatest integer less than or equal to n/m) is odd. Then $f(n)\leqq t(m)$.*

Lemma 4.2. *For any positive integer n, there exists an integer $m\leqq[\text{snd}\,(n/2)]^2$ such that $\lfloor n/m \rfloor$ is odd.*

Theorem 4.3. $t(m)=O(\log^3 m/\log\log m)$.

Proof of Theorem 4.1. Let $M\in\mathcal{M}$ be a matrix with $t(m)$ rows such that $d(M)=m$. Let c be the number of columns of M. We use M to construct an n-uniform hypergraph with $t(m)$ edges having no equi-partition.

Let $n=am+r$ where $0\leqq r<m$ and $a=\lfloor n/m\rfloor$ is odd. Let $Y_1, Y_2, \ldots, Y_c$, Z be disjoint sets with $|Y_j|=amx_j^M$ and $|Z|=r$. Define a hypergraph $H=(X, \mathcal{E})$ where $X=Y_1\cup Y_2\cup\ldots\cup Y_c\cup Z$ and $\mathcal{E}$ has edges $E_1, E_2, \ldots, E_{t(m)}$ given by

$$E_i = \Big(\bigcup_{j|M_{ij}=1} Y_j\Big)\cup Z.$$

For each i,

$$|E_i| = \sum_{j|M_{ij}=1} amx_j^M+r = am\sum_{j=1}^{c} M_{ij}x_j^m+r = am+r = n,$$

so H is n-uniform. We claim that H has no equi-partition.

Suppose the two-coloring $X=R\dot{\cup}B$ is an equi-partition of H. Let $Y= =Y_1\cup Y_2\cup\ldots\cup Y_c$. For each edge, $|E_i\cap R|=n/2$ so $|E_i\cap R\cap Y|=n/2-|Z\cap R|$. For $1\leqq j\leqq c$, let $z_j=|Y_j\cap R|/(n/2-|Z\cap R|)$. Then

$$1 = |E_i\cap R\cap Y|/(n/2-|Z\cap R|) = \sum_{j|M_{ij}=1} z_j = \sum_{j=1}^{c} M_{ij}z_j.$$

Thus $M\mathbf{z}=\mathbf{e}$ and since $M\in\mathcal{M}$, $\mathbf{z}=\mathbf{x}^M$. Now $|Y_j\cap R|=z_j(n/2-|Z\cap R|)$ is an integer for each j and m is the least integer such that mz_j is integral, so $(n/2-|Z\cap R|)/m$ must be an integer. Symmetrically, $(n/2-|Z\cap B|)/m$ is also an integer. Their difference $(|Z\cap R|-|Z\cap B|)/m$ is an integer and since $0\leqq|Z|<m$, we have $|Z\cap R|= =|Z\cap B|=r/2$. Thus $(n/2-|Z\cap R|)/m=(n-r)/2m=am/2m=a/2$ is an integer, contradicting the fact that a is odd. Therefore H has no equi-partition. ∎

Proof of Lemma 4.2. Let $s=\operatorname{snd}(n/2)$. If s is a power of 2 then letting $m=s$ we have $\lfloor n/m\rfloor=n/m$ is odd. If s is not a power of 2, let 2^j be the smallest power of 2 that exceeds s. Then $n/2^j$ is an integer and $n/(2^j s)=q+(a/s)$ where q is an integer and a is a positive integer less than s. Let i be the smallest integer such that $a2^i>s$. Clearly $i\geqq 1$ and $a2^i/s<2$. Taking $m=2^{j-i}s$ we obtain $m\leqq s^2$ and

$$\lfloor n/m\rfloor = \lfloor n/2^{j-i}s\rfloor = \left\lfloor 2^i q+\frac{a2^i}{s}\right\rfloor = 2^i q+1. \quad \blacksquare$$

Proof of Theorem 4.3. The upper bound on $t(m)$ is obtained by construction. Let $q_1, q_2, \ldots, q_k$ be positive integers. Let $M(q_1, q_2, \ldots, q_k)$ be the $q_1+q_2+\ldots+q_k+k$ by $q_1+q_2+\ldots+q_k+k$ matrix with k diagonal blocks, the j^{th} block being a q_j+1 by q_j+1 identity matrix, and all off-block entries equal to 1. A routine computation shows that $M(q_1, \ldots, q_k)\mathbf{x}^M=\mathbf{e}$ has a unique nonnegative solution and

$$d\big(M(q_1, \ldots, q_k)\big) = lcm\,(q_1, \ldots, q_k)\left(\sum_{i=1}^{k}\frac{1}{q_i}+k-1\right).$$

Hence if m is any positive integer and $q_1, q_2, \ldots, q_k$ are positive integers such that

$$m = lcm(q_1, \ldots, q_k)\left(\sum_{i=1}^{k}\frac{1}{q_i}+k-1\right)$$

then $t(m)\leqq q_1+q_2+\ldots+q_k+k$. Thus Theorem 4.3 follows from the following number theoretic result.

Lemma 4.4. *For every natural number m, there exist natural numbers $q_1, q_2, \ldots, q_k$ such that*

$$m = lcm\,(q_1, \ldots, q_k)\left(k-1+\sum_{i=1}^{k}\frac{1}{q_i}\right)$$

and $k+\sum_{i=1}^{k} q_i = O(\log^3 m/\log\log m)$.

Proof. Let p_i denote the i^{th} prime, and L_i denote the product of the first i primes. Say we let u of the q's equal 1 and v_i of the q's equal p_i for $i=1, \ldots, h$. Then

$$k = u+\sum_{i=1}^{h} v_i. \tag{4.1}$$

We want to choose $h, u, v_1, \ldots, v_h$ with all of the v_i non-zero so that

$$m = L_h\left(2u-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \tag{4.2}$$

and so that

$$2u+\sum_{i=1}^{h} v_i(p_i+1) = O(\log^3 m/\log\log m). \tag{4.3}$$

In the analysis below, we will need the following facts which are elementary consequences of the prime number theorem [I].

Lemma 4.5. *If* $L_h = C$ *then*

(i) $h = (1+o(1))\log C/\log\log C$;

(ii) $p_h = (1+o(1))\log C$;

(iii) $\sum_{i=1}^{h} p_i = \left(\frac{1}{2}+o(1)\right)\log^2 C/\log\log C$;

(iv) $\sum_{i=1}^{h} p_i^2 = \left(\frac{1}{3}+o(1)\right)\log^3 C/\log\log C$. ∎

Continuing with the proof of lemma 4.4, we observe that to satisfy (4.2), it is enough to find $v_1, \ldots, v_h$ such that

$$m \geqq L_h\left(-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \tag{4.4}$$

and

$$m \equiv L_h\left(-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \mod 2L_h \tag{4.5}$$

since then u can be chosen to make (4.2) hold. Since $2L_h = 4p_2p_3\cdots p_h$, (4.5) holds if and only if the following set of congruences hold:

$$m \equiv L_h\left(-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \mod p_j \quad \text{for} \quad 2 \leqq j \leqq h \tag{4.6}$$

and

$$m \equiv L_h\left(-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \mod 4. \tag{4.7}$$

These congruences can be simplified to

$$m \equiv v_j L_h/p_j \mod p_j \quad \text{for} \quad 2 \leqq j \leqq h \tag{4.8}$$

$$m \equiv (3v_1-2)L_h/2 \mod 4. \tag{4.9}$$

For each $2 \leqq j \leqq h$ there is a unique number v_j between 1 and p_j that satisfies (4.8) and a unique number v_1 between 1 and 4 satisying (4.9). Now for v_i chosen in these ranges

$$L_h\left(-1+\sum_{i=1}^{h} v_i+\sum_{i=1}^{h}\frac{v_i}{p_i}\right) \leqq L_h\left(5+\sum_{i=2}^{h} p_i+h-1\right). \tag{4.10}$$

To satisfy (4.4), it is sufficient to choose h so that the right hand side of (4.10) is less than or equal to m. Assume m is large and choose h to be the largest integer for which $L_h < m \log\log m/\log^2 m$. Then from Lemma 4.5, $h = \log m/\log\log m(1+o(1))$ and

$$\sum_{i=1}^{h} p_i = \left(\frac{1}{2}+o(1)\right)\log^2 m/\log\log m.$$

Thus the right hand side of (4.10) is at most $(m/2)+o(m)$, so that (4.4) and hence (4.2) holds. Finally

$$\begin{aligned} 2u+\sum_{i=1}^{h} v_i(p_i+1) &\leqq \frac{m}{L_h}+12+\sum_{i=2}^{h}(p_i^2+p_i) = \\ &= \frac{mp_{h+1}}{L_{h+1}}+6+\sum_{i=1}^{h}(p_i^2+p_i) \leqq \\ &\leqq \frac{mp_{h+1}}{m\log\log m/\log^2 m}+6+\sum_{i=1}^{h}(p_i^2+p_i) = \\ &= O(\log^3 m/\log\log m) \end{aligned}$$

by Lemma 4.5, which is (4.3). ▮

5. Open questions

The first obvious question is to resolve the disparity between the upper and lower bounds of Theorem 1.3. It is not possible to substantially improve the lower bound by improving the lower bound on $g(n)$ because there is a known upper bound on $g(n)$ of $\log n$. On the other hand, it seems likely that the upper bound on $f(n)$ can be improved by improving the upper bound on $t(m)$ through better constructions. In fact, it seems quite reasonable to expect that $t(n)$ behaves much the same as $g(n)$, which would imply that the true behavior of $f(n)$ is close to the lower bound.

The appearance of snd $(n/2)$ in both the upper and lower bounds suggests that $f(n)$ is a function only of snd $(n/2)$ and this would be interesting to know. If so, is $f(n)$ an increasing function of snd $(n/2)$? A weaker but still interesting result would be to show that f satisfies $f(a+b) \geqq \min(f(a), f(b))$.

Another question we would like to see resolved: Is $g(n)$ a monotone function of n?

References

[1] N. ALON and K. BERMAN, Regular hypergraphs, Gordon's lemma, Steinitz' lemma and invariant theory, *to appear in J. Combinatorial Theory, Ser. A.*

[2] J. BECK, Balanced 2-colorings of finite sets in the square I, *Combinatorica*, **1** (1981), 327—335.

[3] J. BECK and T. FIALA, "Integer-making" theorems, *Discrete Applied Math*, 3 (1981), 1—8.

[4] P. ERDŐS and V. T. SÓS, *private communication.*

[5] J. E. GRAVER, A survey of the maximum depth problem for indecomposable exact covers, *in: Infinite and Finite Sets, Col. Math. János Bolyai,* (1973) 731—743.

[6] A. E. INGHAM, *The Distribution of Prime Numbers,* Cambridge Tracts in Mathematics, no. 30 Cambridge University Press, 1937.

[7] L. LOVÁSZ, J. SPENCER and K. VESZTERGOMBI, Discrepancy of set-systems and matrices, *preprint.*

[8] W. RUDIN, *Principles of Mathematical Analysis.* McGraw Hill, 1964.

[9] V. T. SÓS, Irregularities of partitions: Ramsey theory, uniform distribution, *in Surveys in Combinatorics,* London Math. Soc. Lecture Notes Series **82** (1983), 201—246.

[10] J. SPENCER, Six standard deviations suffice, *Trans. Amer. Math. Soc.*, **289** (1985), 679—706.

N. Alon

Tel Aviv University
Ramat Aviv, Tel Aviv 69978
Israel

C. Pomerance, M. Saks, P. Seymour

Bell Communications Research
Morristown, NJ 07960
USA

D. J. Kleitman

Massachusetts Institute of Technology
Cambridge, MA 02139
USA