# Linear versus Hereditary Discrepancy *

Tom Bohman † Ron Holzman ‡

March 24, 2002

**Abstract**

Lovász, Spencer and Vesztergombi proved that the linear discrepancy of a hypergraph $\mathcal{H}$ is bounded above by the hereditary discrepancy of $\mathcal{H}$, and conjectured a sharper bound that involves the number of vertices in $\mathcal{H}$. In this paper we give a short proof of this conjecture for hypergraphs of hereditary discrepancy 1. For hypergraphs of higher hereditary discrepancy we give some partial results and propose a sharpening of the conjecture.

## 1 Introduction

The concept of hypergraph discrepancy was introduced and studied by Beck, Sós, Spencer and others in the last two decades [2] [3] [4] [13] [14]. This concept provides a unified combinatorial framework for a number of problems arising in geometry and number theory and has found applications in many areas (see [4, page 1442]).

The *discrepancy* of a hypergraph $\mathcal{H} \subseteq 2^{[n]}$ is defined by

$$\operatorname{disc}(\mathcal{H}) = \min_f \max_{X \in \mathcal{H}} \left| \sum_{x \in X} f(x) \right|,$$

where $f$ ranges over all two-colorings $f : [n] \to \{-1, 1\}$ of the ground set of $\mathcal{H}$. In words, the goal of hypergraph discrepancy is to find a two-coloring of the vertex set such that the colors are as balanced as possible in each edge. Note that the discrepancy of a hypergraph can be small for somewhat arbitrary reasons (e.g. if $\mathcal{H}$ is an $r$-regular hypergraph and we attach a set of $r$ vertices to every edge in $\mathcal{H}$ then the resulting hypergraph has discrepancy 0 while the discrepancy of $\mathcal{H}$ itself can be arbitrarily large). The following two variants of hypergraph discrepancy are better measures of the 'intrinsic balance' of a hypergraph as

*A proof of the conjecture of Lovász, Spencer and Vesztergombi for hypergraphs of hereditary discrepancy 1 has also been independently obtained by B. Doerr [6].

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, USA. Supported in part by NSF grant DMS-0100400. E-mail: tbohman@moser.math.cmu.edu

‡Department of Mathematics, Technion–Israel Institute of Technology, Haifa, Israel. Research supported by the Technion V.P.R. Fund—M. and M.L. Bank Mathematics Research Fund and by the Fund for the Promotion of Research at the Technion. E-mail: holzman@tx.technion.ac.il

they are given by the worst-cases from large collections of 'balancing problems' that depend on $\mathcal{H}$. The *hereditary discrepancy* of $\mathcal{H}$ is defined as

$$\mathrm{herdisc}(\mathcal{H}) = \max_{Y \subseteq [n]} \mathrm{disc}\left(\mathcal{H}|_Y\right),$$

where $\mathcal{H}|_Y$ denotes the restriction of $\mathcal{H}$ to $Y$, i.e. the hypergraph with ground set $Y$ and edge set $\{X \cap Y : X \in \mathcal{H}\}$. The *linear discrepancy* of $\mathcal{H}$ is defined by

$$\mathrm{lindisc}(\mathcal{H}) = \max_{p_1,\ldots,p_n \in [0,1]} \min_{\epsilon_1,\ldots,\epsilon_n \in \{0,1\}} \max_{X \in \mathcal{H}} \left| \sum_{i \in X} \epsilon_i - p_i \right|.$$

The task for linear discrepancy is to round the real numbers $p_1, \ldots, p_n$ up or down in such a way as to minimize the total error on an edge of $\mathcal{H}$.

An investigation of the relationship between linear and hereditary discrepancy was undertaken by Lovász, Spencer and Vesztergombi [9] [15]. They proved that

$$\mathrm{lindisc}(\mathcal{H}) < \mathrm{herdisc}(\mathcal{H}).$$

We note in passing that a key lemma in Baranyai's proof of the existence of a factorization of the complete uniform hypergraph [1] can be viewed as a special case of this statement. Lovász, Spencer and Vesztergombi went on to ask: What is the maximal $c_n$ so that if $\mathcal{H}$ is a hypergraph on a ground set of size $n$ then

$$\mathrm{lindisc}(\mathcal{H}) \leq (1 - c_n)\mathrm{herdisc}(\mathcal{H})?$$

Since the hypergraph on $[n]$ with edges $\{1\}, \ldots, \{n\}, \{1, \ldots, n\}$ has hereditary discrepancy 1 and linear discrepancy at least $\frac{n}{n+1}$ (to see the latter, consider the assignment $p_i = \frac{1}{n+1}$ for all $i \in [n]$) we have $c_n \leq \frac{1}{n+1}$.

**Conjecture 1 (Lovász, Spencer, Vesztergombi).** *If $\mathcal{H}$ is a hypergraph on ground set $[n]$ then*

$$\mathrm{lindisc}(\mathcal{H}) \leq \left(1 - \frac{1}{n+1}\right) \mathrm{herdisc}(\mathcal{H}).$$

Conjecture 1 was proved to hold for two special types of hypergraphs, both of hereditary discrepancy 1. For interval hypergraphs (the vertex-set is $[n]$ and the edge-set consists of the integer intervals $[i, j]$ for $1 \leq i \leq j \leq n$) Spencer [15] gave a short argument (a 'gem' shown to him by Lovász). For hypergraphs with edge-set consisting of initial segments in either of two given orderings of $[n]$, Knuth gave a complicated proof [8]; another proof was given independently by J. Ossowski. Knuth dubbed this the 'two-way rounding' problem.

Here we prove that Conjecture 1 holds for all hypergraphs of hereditary discrepancy 1; in fact, we prove a stronger matrix-version of the conjecture for this special case. The linear discrepancy of an $m \times n$ real matrix $A$ is defined as

$$\mathrm{lindisc}(A) = \max_{p \in [0,1]^n} \min_{\epsilon \in \{0,1\}^n} \|Ap - A\epsilon\|_\infty.$$

So, for example, the linear discrepancy of a hypergraph $\mathcal{H}$ equals the linear discrepancy of the incidence matrix of $\mathcal{H}$. Now, a matrix $A$ is *totally unimodular* if each subdeterminant of $A$ is $0, 1$ or $-1$ (for a discussion of totally unimodular matrices and their relevance in integer programming see [11]).

**Theorem 2.** *If $A$ is a totally unimodular $m \times n$ matrix then*

$$\operatorname{lindisc}(A) \leq 1 - \frac{1}{n+1}.$$

Since the incidence matrix of a hypergraph with hereditary discrepancy 1 is totally unimodular (this is a result of Ghouila-Houri [7]), it follows from Theorem 2 that if $\mathcal{H}$ is a hypergraph on $[n]$ of hereditary discrepancy 1 then $\operatorname{lindisc}(\mathcal{H}) \leq 1 - \frac{1}{n+1}$.

The remainder of this paper is organized as follows. The proof of Theorem 2 is given in Section 2. Section 3 contains a corollary of Theorem 2 for hypergraphs of hereditary discrepancy greater than 1. We conclude the paper in Section 4 by proposing a stronger version of the conjecture.

## 2 Proof of Theorem 2

Let $A$ be a totally unimodular $m \times n$ matrix and let $p \in [0,1]^n$. We have to show that there exists $\epsilon \in \{0,1\}^n$ such that $\|Ap - A\epsilon\|_\infty$ is at most $\frac{n}{n+1}$.

Consider the polytope

$$Q = \{x \in [0,1]^n : \lfloor Ap \rfloor \leq Ax \leq \lceil Ap \rceil\}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the component-wise floor and ceiling functions, respectively. Since $A$ is totally unimodular, the extreme points of $Q$ are integral. By Carathéodory's Theorem, there exist extreme points $x_1, \ldots, x_{n+1}$ of $Q$ and nonnegative reals $\lambda_1, \ldots, \lambda_{n+1}$ such that

$$p = \lambda_1 x_1 + \cdots + \lambda_{n+1} x_{n+1},$$

and $\sum_{j=1}^{n+1} \lambda_j = 1$.

Assume without loss of generality that $\lambda_1 \geq \frac{1}{n+1}$. We claim that we can take $\epsilon = x_1$. To see this, first note that we have

$$Ap = \lambda_1 Ax_1 + \cdots + \lambda_{n+1} Ax_{n+1}. \tag{1}$$

For $i = 1, \ldots, m$ let $y_i$ be the $i^{\text{th}}$ coordinate of $Ap$, let $z_i$ be the $i^{\text{th}}$ coordinate of $Ax_1$, and let $W_i$ be the set of indices $j \in \{2, \ldots, n+1\}$ for which the $i^{\text{th}}$ coordinate of $Ax_j$ does not equal $z_i$. It follows from (1) and the fact that the $i^{\text{th}}$ coordinate of each $Ax_j$ is either $\lfloor y_i \rfloor$ or $\lceil y_i \rceil$ that we have

$$|y_i - z_i| = \sum_{j \in W_i} \lambda_j \leq 1 - \frac{1}{n+1}.$$

## 3 Extension to Larger Hereditary Discrepancies

Seymour proved that any totally unimodular matrix (and therefore any hypergraph of hereditary discrepancy 1) can be 'built' out of matrices coming from a relatively simple collection [12]. Unfortunately, we have no such decomposition theorem for hypergraphs of hereditary discrepancy $d$ where $d$ is fixed and larger than 1. Furthermore, there appears to be little discussion of constructions for rich classes of such hypergraphs in the literature.

One natural way to construct hypergraphs of hereditary discrepancy $d$ is to take 'powers' of hypergraphs of hereditary discrepancy 1: For a hypergraph $\mathcal{H}$, we denote by $\mathcal{H}^{(d)}$ the hypergraph having the same ground set as $\mathcal{H}$, and whose edges are the disjoint unions of at most $d$ edges from $\mathcal{H}$. Clearly, we have

$$\begin{aligned} &\operatorname{herdisc}(\mathcal{H}^{(d)}) \leq d \cdot \operatorname{herdisc}(\mathcal{H}), \\ \text{and} \qquad &\operatorname{lindisc}(\mathcal{H}^{(d)}) \leq d \cdot \operatorname{lindisc}(\mathcal{H}). \end{aligned}$$

So, it follows from Theorem 2 that

$$\operatorname{herdisc}(\mathcal{H}) = 1 \qquad \Rightarrow \qquad \operatorname{lindisc}\left(\mathcal{H}^{(d)}\right) \leq \left(1 - \frac{1}{n+1}\right) d.$$

Thus, we have the following corollary of Theorem 2:

**Corollary 3.** *If $\mathcal{G}$ is a hypergraph on $[n]$ of hereditary discrepancy $d$ and there exists a hypergraph $\mathcal{H}$ of hereditary discrepancy 1 such that $\mathcal{G} \subseteq \mathcal{H}^{(d)}$ then we have*

$$\operatorname{lindisc}(\mathcal{G}) \leq \left(1 - \frac{1}{n+1}\right) \operatorname{herdisc}(\mathcal{G}).$$

While Corollary 3 suffices to verify Conjecture 1 for many natural hypergraphs having hereditary discrepancy greater than 1, Corollary 3 does *not* imply Conjecture 1. For example, there exists a hypergraph $\mathcal{G}$ of hereditary discrepancy 2 that is not contained in $\mathcal{H}^{(2)}$ for *any* hypergraph $\mathcal{H}$ of hereditary discrepancy 1.

It seems to the authors that all natural constructions for rich collections of hypergraphs with hereditary discrepancy $d$ are based on hypergraphs of hereditary discrepancy 1 in a way that makes it easy to verify Conjecture 1 for the given collection as a consequence of Theorem 2. The 'powers' of hypergraphs of hereditary discrepancy 1 discussed above are just one example of this phenomenon. A construction that does not have this property (i.e. a construction that produces a rich collection of hypergraphs of hereditary discrepancy $d$ and is independent of the hypergraphs of hereditary discrepancy 1 in that Conjecture 1 cannot be easily verified for the collection using Theorem 2) might give some insight into Conjecture 1 and would be interesting in its own right.

## 4 A Stronger Conjecture

Note that the simple extremal example for Conjecture 1 (i.e. the hypergraph having vertex set $\{1, \ldots, n\}$ and edges $\{1\}, \{2\}, \ldots, \{n\}, \{1, 2, \ldots, n\}$) is a hypergraph of hereditary discrepancy 1. Furthermore, there appears to be no extremal example having larger hereditary discrepancy. We believe that Conjecture 1 can be sharpened by including a dependence on the hereditary discrepancy.

Let $\mathcal{H}_n$ be the interval hypergraph on $[n]$; that is, the hypergraph on vertex-set $[n]$ whose edge-set consists of the integer intervals $[i, j]$ for $1 \leq i \leq j \leq n$.

**Theorem 4.** *For $d = 1, \ldots, \lfloor \frac{n+2}{2} \rfloor$ we have*

$$\operatorname{lindisc}\left(\mathcal{H}_n^{(d)}\right) = \left(1 - \frac{d}{n+1}\right) d.$$

The proof of Theorem 4 is deferred to the end of this section. Now, the hypergraph $\mathcal{H}_n^{(d)}$ is known to have the maximum number of edges among hypergraphs on $n$ vertices having hereditary discrepancy at most $d$ [10], and so, in a sense, it is the hardest instance of the rounding problem within this family of hypergraphs.

**Conjecture 5.** *If $\mathcal{H}$ is a hypergraph on vertex set $[n]$ of hereditary discrepancy $d$ then we have*

$$\operatorname{lindisc}(\mathcal{H}) \leq \left(1 - \frac{d}{n+1}\right) d.$$

The best known upper bound on the linear discrepancy of a hypergraph $\mathcal{H}$ on vertex set $[n]$ with hereditary discrepancy $d$ follows from two results. The first is the fact, mentioned above, that the maximum number of edges in such a hypergraph is the number of edges in $\mathcal{H}_n^{(d)}$, which we denote $f(n,d)$. The second is the bound $\operatorname{lindisc}(\mathcal{H}) \leq \left(1 - \frac{1}{2m}\right) d$, where $m$ is the number of edges in $\mathcal{H}$, which follows from a modification of a rounding argument given by Spencer in [15] (see [5]). Combining these observations, we see that the linear discrepancy of a hypergraph on vertex set $[n]$ with hereditary discrepancy $d$ is at most $\left(1 - \frac{1}{2f(n,d)}\right) d$. In particular, if we write this bound in the form $\operatorname{lindisc}(\mathcal{H}) \leq (1 - c_{n,d})\, d$ then $c_{n,d}$ is 1 over a polynomial in $n$ of degree $2d$.

In connection with Theorem 4, we make the following remark, the proof of which is excluded for the sake of brevity.

**Remark 6.** *Given $p_1, p_2, \ldots, p_n \in [0,1]$ there exist $\epsilon_1, \epsilon_2, \ldots, \epsilon_n \in \{0,1\}$ such that*

$$\left|\sum_{i=a}^{b} p_i - \epsilon_i\right| \leq 1 - \frac{1}{n+1} \qquad \text{for } 1 \leq a \leq b \leq n$$

*and*

$$\left|\sum_{i=a}^{b} p_i - \epsilon_i + \sum_{i=c}^{d} p_i - \epsilon_i\right| \leq 2\left(1 - \frac{2}{n+1}\right) \qquad \text{for } 1 \leq a \leq b < c \leq d \leq n.$$

In words, Remark 6 says that it is possible to round real numbers assigned to $[n]$ such that the error over any interval is at most $1 - \frac{1}{n+1}$ and the error over any union of two disjoint intervals is at most $2\left(1 - \frac{2}{n+1}\right)$. In other words, it is possible to achieve the linear discrepancies of $\mathcal{H}_n$ and $\mathcal{H}_n^{(2)}$ *simultaneously.* This observation leads naturally to the question: Given weights $p_1, p_2, \ldots, p_n \in [0,1]$, does there exist a rounding such that the error over the disjoint union of $d$ intervals is at most $d\left(1 - \frac{d}{n+1}\right)$ for $d = 1, \ldots, \lfloor \frac{n+2}{2} \rfloor$?

We now turn to the proof of Theorem 4. We note that the $d = 1$ case of Theorem 4 was proved by Lovász (this is the 'gem' mentioned by Spencer) and that the main idea in Lovász's proof is central to the proof of Theorem 4 that we give here.

*Proof of Theorem 4.* We first show that any collection of weights $p_1, \dots, p_n \in [0,1]$ on the vertices can be rounded to integers so that the error on every edge is at most $(1-\frac{d}{n+1})d$.

We begin by associating the collection of weights with a collection of intervals on the unit circle. We use the following notational convention for circular intervals: if $a > b$ then $[a,b] = [a,1) \cup [0,b]$. We denote by $|I|$ the length of the circular interval $I$. Set $a_0 = 0$ and for $i = 1, \dots, n$ set

$$a_i = \sum_{j=1}^{i} p_j \pmod 1.$$

The points $a_0, \dots, a_n$ partition the unit circle naturally into a collection of intervals which we enumerate in cyclic order as $J_0, \dots, J_n$ (n.b. some of these intervals may be trivial). For notational convenience, the subscripts of these intervals are taken modulo $n+1$.

**Claim.** *There exists $k$ such that*

$$d\,|J_k| + \sum_{i=1}^{d-1} (d-i)\left(|J_{k-i}| + |J_{k+i}|\right) \ge \frac{d^2}{n+1}.$$

*Proof.* Assume for the sake of contradiction that no such $k$ exists. Summing over the resulting $n+1$ strict inequalities we have

$$d^2 = (n+1)\frac{d^2}{n+1} > \sum_{i=0}^{n} |J_i| \left(d + \sum_{j=1}^{d-1} 2j\right) = d^2 \sum_{i=0}^{n} |J_i| = d^2.$$

□

We are now ready to define the rounding $\epsilon_1, \epsilon_2, \dots, \epsilon_n \in \{0,1\}$. Let $J_k$ be the interval given by the Claim. For $i = 1, \dots, n$ we set $I_i = [a_{i-1}, a_i]$ (note that each $I_i$ is the union of consecutive $J_j$'s ) and set $\epsilon_i = 1$ if and only if $J_k \subseteq I_i$.

Note that the absolute value of the rounding error over the integer interval $[i,j]$, where $1 \le i \le j \le n$, equals the length of the circular arc between $a_{i-1}$ and $a_j$ that does not contain $J_k$. Now, every edge of $\mathcal{H}_n^{(d)}$ is the union of at most $d$ such intervals $[i_1, j_1], [i_2, j_2], \dots$, with all indices $i_1 - 1, j_1, i_2 - 1, j_2, \dots$ distinct. Therefore, the total error over such an edge is at most

$$d - d|J_k| - (d-1)\left(|J_{k+1}| + |J_{k-1}|\right) - \cdots - \left(|J_{k+d-1}| + |J_{k-d+1}|\right)$$

(note that here we use $d \le \lfloor \frac{n+2}{2} \rfloor$). By the choice of $k$, this bound on the error is at most $(1-\frac{d}{n+1})d$.

We have shown that the linear discrepancy of $\mathcal{H}_n^{(d)}$ is at most $\left(1-\frac{d}{n+1}\right)d$. It remains to exhibit an assignment of weights $p_1, \dots, p_n \in [0,1]$ that cannot be rounded to integers $\epsilon_1, \dots, \epsilon_n \in \{0,1\}$ with the rounding error over every edge strictly less than $(1-\frac{d}{n+1})d$. Consider the assignment $p_i = \frac{d}{n+1}$ for $i = 1, \dots, n$. Since any $d$-element subset of $[n]$ is an edge of $\mathcal{H}_n^{(d)}$, a rounding that assigns $d$ or more 1's will produce an edge whose error is $(1-\frac{d}{n+1})d$. On the other hand, if fewer than $d$ 1's are assigned then the set $\{i \in [n] : \epsilon_i = 0\}$ is an edge of $\mathcal{H}_n^{(d)}$ whose error is at least $(1-\frac{d}{n+1})d$. □

**Acknowledgement:** The authors thank Joel Spencer for introducing us to this fascinating topic and Jiří Matoušek for bringing the work of Benjamin Doerr to our attention.

## References

[1] Z. Baranyai, On the factorization of the complete uniform hypergraph, in *Infinite and Finite Sets* (A. Hajnal, T. Rado and V. T. Sós, editors), North-Holland, 1975, pp. 91 -108.

[2] J. Beck, Roth's estimate of the discrepancy of integer sequences is nearly sharp, *Combinatorica* **1** (1981), 319-325.

[3] J. Beck, Balanced two-colorings of finite sets in the square I, *Combinatorica* **1** (1981), 327-335.

[4] J. Beck, V. T. Sós, Discrepancy theory, in *Handbook of Combinatorics, Volume II* (R.L. Graham, M. Grötschel and L. Lovász, editors), North-Holland, 1995, pp. 1405-1446.

[5] B. Doerr, Linear and hereditary discrepancy, *Combinatorics, Probability and Computing*, **9** (2000), 349-354.

[6] B. Doerr, Linear discrepancy of totally unimodular matrices, *Combinatorica*, to appear.

[7] A. Ghouila-Houri, Caractérisation des matrices totalement unimodulaires, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)* **254** (1962), 1192-1194.

[8] D. E. Knuth, Two-way rounding, *SIAM Journal on Discrete Math* **8** (1995), 281-290.

[9] L. Lovász, J. Spencer, K. Vesztergombi, Discrepancy of set systems and matrices, *European Journal of Combinatorics* **7** (1986), 151-160.

[10] L. Lovász, K. Vesztergombi, Extremal problems for discrepancy, in *Irregularities of Partitions (Fertőd, 1986)*, Springer, 1989, pp. 107-113.

[11] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, 1986.

[12] P. D. Seymour, Decomposition of regular matroids, *Journal of Combinatorial Theory (B)* **28** (1980), 305-359.

[13] V. T. Sós, Irregularities of partitions: Ramsey theory, uniform distribution, in *Surveys in Combinatorics, London Mathematical Society Lecture Notes Series 82*, 1983, pp. 201-246.

[14] J. Spencer, Six standard deviations suffice, *Transactions of the American Mathematical Society* **289** (1985), 679-706.

[15] J. Spencer, *Ten Lectures on the Probabilistic Method, 2nd edition*, SIAM, 1994.