and linearity of expectation gives

$$\mathrm{E}\,[X] = \sum_{A\in\mathcal{A}} \mathrm{E}\,[X_A] < |\mathcal{A}|(1/n) = 1\,.$$

Thus for some χ we must have $X = 0$. This means $\mathrm{disc}(\mathcal{A},\chi) \le \alpha$ and therefore $\mathrm{disc}(\mathcal{A}) \le \alpha$. ■

13.2 SIX STANDARD DEVIATIONS SUFFICE

When $\mathcal{A}$ has both n sets and n points Theorem 13.1.1 gives

$$disc(\mathcal{A}) = O(\sqrt{n\ln(n)}) \tag{13.1}$$

This was improved by the second author in Spencer (1985a).

Theorem 13.2.1 *Let $\mathcal{A}$ be a family of n subsets of an n element set Ω. Then*

$$disc(\mathcal{A}) \le 6\sqrt{n}$$

With $\chi : \Omega \to \{-1,+1\}$ random, $A \in \mathcal{A}$, $\chi(A)$ has zero mean and standard deviation at most $\sqrt{n}$. If $|\chi(A)| > 6\sqrt{n}$ then $\chi(A)$ is at least six standard deviations off the mean. The probability of this occurring is very small but a fixed positive constant and the number of sets A is going to infinity. In fact, a random χ almost always will *not* work. The specific constant 6 (actually 5.32) was the result of detailed calculations that could certainly be further improved and will not concern us here. Rather we show Theorem 13.2.1 with some constant K replacing 6. The initial argument (found in earlier editions of this work) did not yield an efficient algorithm for finding the desired coloring χ. Indeed, for many years the second author conjectured that no such algorithm would exist. Bansal (2010) gave the first algorithmic argument for Theorem 13.2.1. Here we follow the approach of Lovett and Meka (2012). Their argument is a virtual cornucopia of modern probabilistic methods, we give the basic ideas and leave many of the details to the exercises. We begin by generalizing the problem to vectors.

Theorem 13.2.2 *Let $\vec{r_i} \in R^n$, $1 \le i \le n$ with all $|\vec{r_i}|_\infty \le 1$. Let $\vec{z} = (z_1,\dots,z_n)$ with all $z_j \in [-1,+1]$. Then there exists $\vec{x} = (x_1,\dots,x_n)$ with all $x_j \in \{-1,+1\}$ such that*

$$|\vec{r_i}\cdot(\vec{x}-\vec{z})| \le K\sqrt{n} \tag{13.2}$$

for all $1 \le i \le n$. Here K is an absolute constant.

When $\mathcal{A}$ is a family of n subsets $A_1,\dots,A_n$ of $\Omega = \{1,\dots,n\}$ consider the $n\times n$ incidence matrix A, $a_{ij} = 1$ if $j \in A_i$, else $a_{ij} = 0$. Let $\vec{r_i}$ be the i-th row of A and set $\vec{z} = 0$. The $\vec{x} = (x_1,\dots,x_n)$ given by Theorem 13.2.2 gives the coloring $\chi(j) = x_j$ with the properties of Theorem 13.2.1.

During the proof the vector $\vec{x}$ shall *move* inside the cube $[-1,+1]^n$. We refer to this general technique as a *floating colors* method. It will initially have value $\vec{x} = \vec{z}$ so that (13.2) is trivially satisfied. When a coordinate x_i comes close to ± 1 it will be frozen. For definiteness we set

$$\varepsilon = n^{-1} \tag{13.3}$$

and say x_i is *near the border* if $1 - \varepsilon \leq |x_i| \leq 1$. We call such i *frozen*, and all other i *floating*.

We reduce Theorem 13.2.2 to:

Theorem 13.2.3 *Let $\vec{r_i} \in R^n$, $1 \leq i \leq n$ with all $|\vec{r_i}|_\infty \leq 1$. Let $\vec{z} = (z_1, \ldots, z_n)$ with all $z_j \in [-1,+1]$. Then there exists $\vec{x} = (x_1, \ldots, x_n)$ with all x_j near the border such that*

$$|\vec{r_i} \cdot (\vec{x} - \vec{z})| \leq K\sqrt{n} \tag{13.4}$$

for all $1 \leq i \leq n$. Here K is an absolute constant.

With $\vec{x}$ given by Theorem 13.2.3 one can then simply round each x_i to either -1 or $+1$, whichever is closer. The values $\vec{r_i} \cdot (\vec{x} - \vec{z})$ are then changed by at most $n\varepsilon = 1$ which is $o(\sqrt{n})$, thus giving Theorem 13.2.2.

We find $\vec{x}$ in *phases*. Phase t ends when at most $n2^{-t}$ of the x_i are not near the border. As Phase One contains the basic ideas of the argument we state it separately.

Theorem 13.2.4 *Let $\vec{r_i} \in R^n$, $1 \leq i \leq n$ with all $|\vec{r_i}|_\infty \leq 1$. Let $\vec{z} = (z_1, \ldots, z_n)$ with all $z_j \in [-1,+1]$. Then there exists $\vec{x} = (x_1, \ldots, x_n)$ with at least $n/2$ of the x_j near the border such that*

$$|\vec{r_i} \cdot (\vec{x} - \vec{z})| \leq K_1\sqrt{n} \tag{13.5}$$

for all $1 \leq i \leq n$. Here K_1 is an absolute constant.

Set $\vec{u_i} = n^{-1/2}\vec{r_i}$. We will use (in Phase One) only that the Euclidean norm of $\vec{u_i}$ is at most one. We initially set $\vec{x} = \vec{z}$. We move $\vec{x}$ in *steps* until at least half the coordinates are near the border. For $1 \leq i \leq n$ set

$$L_i = \vec{u_i} \cdot (\vec{x} - \vec{z}) \tag{13.6}$$

Call i *dangerous* if $|L_i|$ is one of the $\frac{n}{4}$ largest of the values $|L_s|$, $1 \leq s \leq n$. In case of ties select precisely $\frac{n}{4}$ values i. We emphasize that as $\vec{x}$ moves the $\frac{n}{4}$ dangerous i can and will change.

We define a vector space $V \subset R^n$, which will describe the allowable directions in which $\vec{x}$ may move. V is those $\vec{y} = (y_1, \ldots, y_n)$ satisfying the following linear conditions:

1. If x_i is near the border then $y_i = 0$.
2. $\vec{y} \cdot (\vec{x} - \vec{z}) = 0$.
3. $\vec{y} \cdot \vec{u_i} = 0$ for all dangerous i.

The number of linear conditions is less than $\frac{n}{2}+1+\frac{n}{4}$. Letting d denote the dimension of V, $d \geq \frac{n}{4}$. We let $\vec{y}$ be a standard multidimensional Gaussian on V. That is, let $\vec{b_1}, \ldots, \vec{b_d}$ be an orthonormal basis for V and set

$$\vec{y} = d^{-1/2}[n_1\vec{b_1} + \cdots + n_d\vec{b_d}] \tag{13.7}$$

where the n_j are independent, each with the standard Normal distribution.

We shall use the directionless property of the Gaussian. Let $\vec{a} \in V$. Then $\vec{y} \cdot \vec{a}$ has a Gaussian Distribution with mean 0 and variance $d^{-1}|\vec{a}|^2$. Suppose $\vec{b} \in R^n$. We can decompose $\vec{b} = \vec{a} + \vec{c}$ with $\vec{a} \in V$, $\vec{c} \in V^\perp$. Then $\vec{y} \cdot \vec{b} = \vec{y} \cdot \vec{a}$. Thus $\vec{y} \cdot \vec{b}$ has a Gaussian Distribution with mean 0 and variance at most $d^{-1}|\vec{b}|^2$.

We now move $\vec{x}$ a small distance in direction $\vec{y}$. Set, for definiteness,

$$\delta = n^{-10} \tag{13.8}$$

A single step then consists of resetting

$$\vec{x} \leftarrow \vec{x} + \delta\vec{y} \tag{13.9}$$

While the Lovett-Meka algorithm is discrete, as the δ of (13.8) becomes small one may think of $\vec{x}$ as moving in a controlled Brownian motion, with the vector space V of permissible directions always changing.

A step fails if some $|x_i| > 1$. When x_i is near the border, $y_i = 0$ and so x_i does not change. If x_i is not near the border it would need to change by at least ε in one step. Let $\vec{U_i}$ denote the vector with one in the i-th position, zero elsewhere. In one step the change in x_i is $\delta\vec{y} \cdot \vec{U_i}$ which is Gaussian with mean zero and variance at most $d^{-1}\delta^2$. With the values ε, δ the probability that the change in x_i is more than ε is then exponentially small. There are only n choices of i and we shall see that there are only polynomially many steps. Thus with probability $1 - o(1)$ no step fails. The chi squared distribution (see exercises) $n_1^2 + \ldots + n_d^2$ is tightly concentrated around its mean d. Thus $|\delta\vec{y}|^2$ is at least $(1 - o(1))\delta^2$ throughout Phase One. At each step $|\vec{x} - \vec{z}|^2$ is being increased by this amount. As they both lie in $[-1, +1]^n$, $|\vec{x} - \vec{z}|^2 \leq 4n$. Letting T denote the number of steps in Phase One, we deduce $T \leq (1 + o(1))4n\delta^{-2}$.

Fix $1 \leq i \leq n$. Let $L_i(t)$ denote the value of L_i given in (13.6) after the t-th step, with initial value $L_i(0) = 0$. With $\vec{y}$ the Gaussian selected at the t-th step

$$L_i(t) = L_i(t-1) + \delta\vec{u_i} \cdot \vec{y} \tag{13.10}$$

Hence L_i will change by a Gaussian with variance $\tau^2 \leq \delta^2$. The $L_i(t)$ then form a martingale. We apply the martingale inequality (13.17) in the Exercises. Here

$$\sigma^2 = T\delta^2 \leq (1 + o(1))4n\delta^{-2}d^{-1}\delta^2 \leq (1 + o(1))4n/d \leq 16 + o(1) \tag{13.11}$$

so $\sigma = 4 + o(1)$. Thus

$$\Pr[\max_{0 \leq t \leq T} |L_i(t)| > K(1 + o(1))] < 2e^{-K^2/32}(1 + o(1)) \tag{13.12}$$

Theorem 13.2.4 is shown by selecting $K_1 = K$ such that

$$2e^{-K^2/32} < 0.05 \tag{13.13}$$

Each L_i has probability less than 0.05 of ever becoming bigger than K in absolute value. By linearity of expectation, the expected number of such i is less than $0.05n$. The randomized Phase One is a success if one never has $|x_i| > 1$, it ends in at most $(1 + o(1))4n\delta^{-2}$ steps *and* there are fewer than $0.1n$ values i such that $|L_i|$ ever becomes bigger than K. The last occurs with probability at least 0.5 and so Phase One is a success with probability at least $0.5 - o(1)$.

Suppose we have success. Have we improved the situation over a standard randomized selection of the x_i? It seems that we still have a positive proportion of *outliers* to deal with. But look again! At each step the dangerous i had their L_i unchanged since the move $\vec{v}$, being in V, was orthogonal to $\vec{u_i}$. As less than $0.1n$ of the L_i ever have $|L_i| \geq K$ it must be that whenever an $|L_i|$ becomes at least K it will become, and stay, dangerous and so $|L_i|$ will remain the same throughout the remainder of Phase One. The single move in which $|L_i|$ exceeds K is miniscule so that after it $|L_i|$ is only $K + o(1)$. Therefore at the end of the process *all* of the $|L_i| \leq K + o(1)$, completing the argument.

We outline the remainder of the argument for Theorem 13.2.3. The $\vec{x}$ at the end of Phase $t-1$ becomes the initial $\vec{z}$ of Phase t. (When the number of floating variables reaches $O(n \ln^{-1/2} n)$ we can switch to a more standard random choosing of the x_i. See the Exercises.) In Phase t we begin with $n2^{-t} \leq m \leq n2^{1-t}$ floating variables so that $n \leq m2^t$. Ignore the nonfloating variables so that we consider $\vec{r_i} \in R^m$. As all coefficients lie in $[-1, +1]$ we may bound $|\vec{r_i}|^2 \leq m$. We set $\vec{u_i} = m^{-1/2}\vec{r_i}$ so that $|\vec{u_i}| \leq 1$. We modify Theorem 13.2.4 as follows:

Theorem 13.2.5 *Let $n \leq m2^t$. Let $\vec{r_i} \in R^m$, $1 \leq i \leq n$ with all $|\vec{r_i}|_\infty \leq 1$. Let $\vec{z} = (z_1, \ldots, z_m)$ with all $z_j \in [-1, +1]$. Then there exists $\vec{x} = (x_1, \ldots, x_n)$ with at least $m/2$ of the x_j near the border such that*

$$|\vec{r_i} \cdot (\vec{x} - \vec{z})| \leq K_t\sqrt{m} \tag{13.14}$$

for all $1 \leq i \leq n$. Here K_t is an absolute constant.

We define $L_i = \vec{u_i} \cdot (\vec{x} - \vec{z})$ as in (13.6). Now i is dangerous if $|L_i|$ is one of the $n2^{-t-2} \leq \frac{m}{4}$ largest values. The large deviation bound (13.12) for the L_i is still valid but now, instead of (13.13) we define $K = K_t$ such that

$$2e^{-K^2/2} \leq 0.05 \cdot 2^{-t} \tag{13.15}$$

Now the expected number of i, $1 \leq i \leq n$, for which $|L_i(t)| \geq K$ ever occurs is less than $0.05n2^{-t} \leq 0.05m$. The remainder of the argument is as before.

From (13.15) we may set $K_t = \sqrt{c_1 + c_2 \ln t} = O(\sqrt{\ln t})$. As $m \leq n2^{-t+1}$, in Phase t all $|L_i| \leq K_t^*\sqrt{n}$ with $K_t^* = 2^{(1-t)/2}\sqrt{c_1 + c_2 \ln t} = O(2^{-t/2}\sqrt{\ln t})$.

Finally we glue all the phases together. For each i, using the original definition (13.6) of L_i, the absolute value of the change in L_i in Phase t is at most K_t^*. But

$\sum_{t=1}^{\infty} K_t^*$ converges to some K – basically the $2^{-t/2}$ gain by having fewer variables outweighs the $\sqrt{\ln t}$ loss by having more vectors than variables – and hence at the end of the process all $|L_i| \leq K$.

13.3 LINEAR AND HEREDITARY DISCREPANCY

We now suppose that $\mathcal{A}$ has more points than sets. We write $\mathcal{A} = \{A_1, \ldots, A_n\}$ and $\Omega = \{1, \ldots, m\}$ and assume $m > n$. Note that $\operatorname{disc}(\mathcal{A}) \leq K$ is equivalent to the existence of a set S; namely $S = \{j : \chi(j) = +1\}$, with $|S \cap A|$ within $K/2$ of $|A|/2$ for all $A \in \mathcal{A}$. We define the *linear discrepancy* $\operatorname{lindisc}(\mathcal{A})$ by

$$\operatorname{lindisc}(\mathcal{A}) = \max_{p_1,\ldots,p_m \in [0,1]} \min_{\epsilon_1,\ldots,\epsilon_m \in \{0,1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} (\epsilon_i - p_i) \right| .$$

The upper bound $\operatorname{lindisc}(\mathcal{A}) \leq K$ means that given any $p_1, \ldots, p_m$ there is a "simultaneous roundoff" $\epsilon_1, \ldots, \epsilon_m$ so that, with $S = \{j : \epsilon_j = 1\}$, $|S \cap A|$ is within K of the weighted sum $\sum_{j \in A} p_j$ for all $A \in \mathcal{A}$. Taking all $p_j = \frac{1}{2}$, the upper bound implies $\operatorname{disc}(\mathcal{A}) \leq 2K$. But $\operatorname{lindisc}(\mathcal{A}) \leq K$ is much stronger. It implies, taking all $p_j = \frac{1}{3}$, the existence of an S with all $|S \cap A|$ within K of $|A|/3$, and much more. Linear discrepancy and its companion hereditary discrepancy defined below have been developed in Lovász, Spencer and Vesztergombi (1986). For $X \subset \Omega$ let $\mathcal{A}|_X$ denote the restriction of $\mathcal{A}$ to X, i.e., the family $\{A \cap X : A \in \mathcal{A}\}$. The next result "reduces" the bounding of $\operatorname{disc}(\mathcal{A})$ when there are more points than sets to the bounding of $\operatorname{lindisc}(\mathcal{A})$ when the points do not outnumber the sets.

Theorem 13.3.1 *Let $\mathcal{A}$ be a family of n sets on m points with $m \geq n$. Suppose that $\operatorname{lindisc}(\mathcal{A}|_X) \leq K$ for every subset X of at most n points. Then $\operatorname{lindisc}(\mathcal{A}) \leq K$.*

Proof. Let $p_1, \ldots, p_m \in [0,1]$ be given. We define a reduction process. Call index j fixed if $p_j \in \{0,1\}$, otherwise call it floating, and let F denote the set of floating indices. If $|F| \leq n$ then halt. Otherwise, let $y_j, j \in F$, be a nonzero solution to the homogeneous system

$$\sum_{j \in A \cap F} y_j = 0, \qquad A \in \mathcal{A} .$$

Such a solution exists since there are more variables ($|F|$) than equations (n) and may be found by standard techniques of linear algebra. Now set

$$\begin{aligned} p'_j &= p_j + \lambda y_j, \quad j \in F, \\ p'_j &= p_j, \qquad\quad j \notin F \end{aligned}$$

where we let λ be the real number of least absolute value so that for some $j \in F$ the value p'_j becomes zero or one. Critically,

$$\textstyle\sum_{j \in A} p'_j = \sum_{j \in A} p_j + \lambda \sum_{j \in A \cap F} y_j = \sum_{j \in A} p_j \qquad (*)$$