

A Trace Bound for the Hereditary Discrepancy*

B. Chazelle¹ and A. Lvov²

¹Department of Computer Science, Princeton University,
 Princeton, NJ 08544, USA
 chazelle@cs.princeton.edu

and

NEC Research Institute, 4 Independence Way,
 Princeton, NJ 08540, USA

²Program in Applied and Computational Mathematics,
 Princeton University, Princeton, NJ 08544, USA
 lvov@math.princeton.edu

Abstract. Let A be the incidence matrix of a set system with m sets and n points, $m \leq n$, and let $t = \text{tr } M$, where $M = A^T A$. Finally, let $\sigma = \text{tr } M^2$ be the sum of squares of the elements of M . We prove that the hereditary discrepancy of the set system is at least $\frac{1}{4} c^{n\sigma/t^2} \sqrt{t/n}$, with $c = \frac{1}{324}$. This general *trace bound* allows us to resolve discrepancy-type questions for which spectral methods had previously failed. Also, by using this result in conjunction with the *spectral lemma* for linear circuits, we derive new complexity bounds for range searching.

- We show that the (red–blue) discrepancy of the set system formed by n points and n lines in the plane is $\Omega(n^{1/6})$ in the worst case and always¹ $\tilde{O}(n^{1/6})$.
- We give a simple *explicit* construction of n points and n halfplanes with hereditary discrepancy $\tilde{\Omega}(n^{1/4})$.
- We show that in any dimension $d = \Omega(\log n / \log \log n)$, there is a set system of n points and n axis-parallel boxes in \mathbf{R}^d with discrepancy $n^{\Omega(1/\log \log n)}$.
- Applying these discrepancy results together with a new variation of the spectral lemma, we derive a lower bound of $\Omega(n \log n)$ on the arithmetic complexity of off-line range searching for points and lines (for nonmonotone circuits). We also prove a lower bound of $\Omega(n \log n / \log \log n)$ on the complexity of orthogonal range searching in any dimension $\Omega(\log n / \log \log n)$.

* A preliminary version of this paper appeared in *Proc. 16th Annual ACM Symp. Comput. Geom.* (2000), pp. 64–69. This work was supported in part by NSF Grant CCR-96-23768, ARO Grant DAAH04-96-1-0181, and NEC Research Institute.

¹ We use the notation $\tilde{O}(m)$ and $\tilde{\Omega}(m)$ as shorthand for $O(m \log^c m)$ and $\Omega(m / \log^c m)$, respectively, for some constant $c > 0$.

1. Introduction

The red–blue discrepancy of a set system is the minimum value of $\|Ax\|_\infty$ over all $x \in \{-1, 1\}^n$, where A is the m -by- n incidence matrix of the set system. In a geometric context, the columns correspond to points in \mathbf{R}^d and the rows correspond to the characteristic vectors of certain regions of space. The discrepancy is not necessarily a robust notion: It might be very small but shoots up as soon as we remove a few points. To avoid this type of pathology, Lovász et al. [10] introduced the hereditary discrepancy, which they defined as the maximum discrepancy of any set system derived from A by removing any number of columns. It is important to observe that in geometric set systems this is just like removing points. Geometric set systems are defined by collections of points and regions: For example, one might consider the family of set systems formed by n points and m axis-parallel boxes. To remove a point produces another (smaller) geometric set system of the same type, and so in that case the hereditary discrepancy can be used to prove bounds (both lower and upper) on the standard discrepancy. It is simply a more convenient tool; not a restrictive one. Lovász et al. proved the classical bound, stated here for the case $m = n$,

$$\text{herdisc}(A) \geq \frac{1}{2} |\det A|^{1/n}.$$

A similar lower bound by Lovász and Vesztergombi [11] relates the hereditary discrepancy to the volume of the convex hull of the $2n$ row vectors of A and $-A$.

Lower bounds of this type suffer from two weaknesses: one is that the determinant can be very small for the “wrong” reasons. In other words, the discrepancy could be very high even though $\det A$ is very small or even null. The other problem is that to derive a decent estimate on the determinant might be very difficult. Harmonic analysis is the standard vehicle for estimating the singular values of A , and, hence, $\det A$. This works well when the set systems consist of full-dimensional regions with smooth parameterizations (e.g., points in balls, boxes, or simplices) [1], [7], [12], [13]. The approach seems to fail, however, when the regions are curves or surfaces.

Our strategy to overcome this difficulty is to relate the discrepancy directly to combinatorial (as opposed to spectral) features of the set system. In particular, we focus on the traces of $M = A^T A$ and M^2 , and prove the following, general *trace bound*:

Lemma 1.1. *If A is an m -by- n incidence matrix ($m \leq n$) and $M = A^T A$, then*

$$\text{herdisc}(A) \geq \frac{1}{4} c^{n \text{tr } M^2 / \text{tr}^2 M} \sqrt{\text{tr } M / n},$$

where $c = \frac{1}{324}$.

Note that $\text{tr } A^T A = \text{tr } A A^T$ and $\text{tr}(A^T A)^2 = \text{tr}(A A^T)^2$, and so the trace bound is self-dual for square matrices. In particular, we can always replace $M = A^T A$ by $M = A A^T$ if this makes calculations easier. What makes this lower bound particularly useful is that the traces of M and M^2 have natural interpretations:

- Algebraically, $\text{tr } M$ is the sum of the eigenvalues. Combinatorially, it is the number of ones in A . Geometrically, it is the count of point/region incidences.

- Algebraically, $\text{tr } M^2$ is the sum of the squares of the eigenvalues. Combinatorially, it is the number of rectangles of ones in A or, equivalently, the number of closed paths (simple and nonsimple) of length 4 in the bipartite graph corresponding to A . Geometrically, it counts the pairs of points in all the pairwise intersections of regions.

These characterizations give us the combinatorial tools needed to resolve a few open questions: some have to do with discrepancy theory proper, others with the complexity of off-line range searching in a nonmonotone model (the standard model of linear circuits with bounded coefficients). A problem instance is specified by a set system of n points p_i , each one associated with a real number x_i (its weight) and m regions. The output consists of the sums of the weights in each region. This is equivalent to multiplying the incidence matrix A by the weight vector $x = (x_1, \dots, x_n)^T$. The circuit takes x as input and produces Ax as output. It consists of a number of gates with two inputs and one output, where each gate computes a linear combination of its inputs. The circuit depends on A , but the coefficients at the gates are bounded by some absolute constant. The complexity of a circuit is the total number of gates in it. Perhaps the best known linear circuit is the Fast Fourier Transform, which is of complexity $O(n \log n)$.

The *spectral lemma* [3], [5] states that the complexity of a bounded-coefficient linear circuit for computing $x \mapsto Ax$ is at least $\Omega(\max_k k \log \lambda_k)$, where λ_k is the k th largest eigenvalue of $A^T A$. By appealing to Lemma 1.1 we are able to bypass eigenvalues altogether to give a lower bound based solely on traces.²

Lemma 1.2. *The circuit complexity of $x \mapsto Ax$, where A is an m -by- n ($m \leq n$) incidence matrix is*

$$\Omega_\varepsilon(n \log(\text{tr } M/n - \varepsilon \sqrt{\text{tr } M^2/n})),$$

where $M = A^T A$ and $\varepsilon > 0$ is an arbitrarily small constant.

This paper follows the principle of the discrepancy method [5], which is to derive complexity bounds from discrepancy results. We specialize the two lemmas above to two classical set systems that had resisted previous investigations: points in lines and points in higher-dimensional boxes. We show that the (red–blue) discrepancy of the set system formed by n points and n lines in the plane is $\Omega(n^{1/6})$. Using the partial-coloring method of discrepancy theory we give a nearly matching upper bound of $O(n^{1/6} \log^{2/3} n)$. We also prove that the corresponding off-line range searching problem has complexity $\Omega(n \log n)$.

The exponent of $\frac{1}{6}$ for points-in-lines in the real plane stands in sharp contrast with that of $\frac{1}{4}$ associated with points-in-lines in the finite projective plane or points-in-halfplanes in \mathbf{R}^2 . Halfplane range searching does not seem harder than line range searching (think of Hopcroft’s problem), and the best algorithms for both run in time $\tilde{O}(n^{4/3})$, so it is intriguing that their discrepancies should differ by so much.

² The spectral lemma allows the presence of help gates, which are discussed in detail in [3]. So does Lemma 1.2, but we ignore this point for the purpose of this paper.

Of incidental interest, our techniques allow a very elementary derivation of an $\tilde{\Omega}(n^{1/4})$ bound for points and halfplanes. The precise bound we get is $\Omega(n/\log n)^{1/4}$, which is not quite the optimal bound of $\Omega(n^{1/4})$ established in [4]. There are two reasons why our bound is interesting nevertheless: (i) it is purely combinatorial and, in particular, it requires no harmonic analysis; (ii) the construction is *explicit* whereas the previous one was existential.

Also, we provide an explicit construction of a set system with n points and n axis-parallel boxes in any dimension $\Omega(\log n/\log \log n)$, whose hereditary discrepancy is $n^{\Omega(1/\log \log n)}$. Of course, this implies the same lower bound on the standard discrepancy, but the construction is no longer explicit. This quasi-linear discrepancy stands in sharp contrast with the polylogarithmic bound for the fixed-dimensional case. In a companion paper [6], we have shown, using a completely different construction, that the discrepancy becomes $n^{\Omega(1)}$ in any dimension $\Omega(\log n)$.

We establish a lower bound of $\Omega(n \log n)$ on the complexity of off-line line range searching. The only previous result was $\Omega(n^{4/3})$ for monotone circuits [5]. We also prove a lower bound of $\Omega(n \log n/\log \log n)$ on the complexity of orthogonal range searching in any dimension $\Omega(\log n/\log \log n)$. By Lemma 1.2 and the results of [6], the bound is $\Omega(n \log n)$ in any dimension $\Omega(\log n)$. Note that in fixed dimension, the best known lower bound is $\Omega(n \log \log n)$ [2].

2. The Trace Bound

The linear discrepancy of an m -by- n 0/1 matrix A is defined as

$$\text{lindisc}(A) = \max_{c \in [-1, 1]^m} \min_{x \in \{-1, 1\}^n} \|Ax - c\|_\infty.$$

The main advantage of the linear discrepancy over the hereditary kind is that the continuous parameter c brings linear algebra into the picture. Lovász et al. [10] proved that

$$\text{herdisc}(A) \geq \frac{1}{2} \text{lindisc}(A). \tag{1}$$

By definition, for any $c \in [-1, 1]^m$, there exists some $x \in \{-1, 1\}^n$ such that $Ac = Ax + y$, where $y \in [-\text{lindisc}(A), \text{lindisc}(A)]^m$. The image of the cube $[-1, 1]^n$ under the transformation given by A is a polyhedron in \mathbf{R}^m whose vertices belong to $A\{-1, 1\}^n$. We conclude:

Lemma 2.1. *The polyhedron $A[-1, 1]^n$ is covered by the $\leq 2^n$ m -dimensional cubes of side length $2 \text{lindisc}(A)$ centered at the vertices of $A[-1, 1]^n$.*

The trace bound (Lemma 1.1) does not involve eigenvalues explicitly but its proof does. We assume from now on that $m \leq n$. Let $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ be the eigenvalues of $A^T A$; ties are broken arbitrarily.

Lemma 2.2. *For any $1 \leq k \leq n$,*

$$\text{lindisc}(A) \geq 18^{-n/k} \sqrt{\lambda_k}.$$

To prove this lemma we begin with a technical result about projections of boxes.

Lemma 2.3. *Let U be an n -dimensional unit cube and let $\text{proj}(U)$ be its orthogonal projection on some k -dimensional subspace, $1 \leq k \leq n$. Then the k -dimensional volume of $\text{proj}(U)$ satisfies*

$$3^{-n} \leq \text{vol}(\text{proj}(U)) \leq 3^n.$$

Proof. We begin with the upper bound. The number of k -dimensional faces of an n -dimensional cube is $\binom{n}{k} \cdot 2^{n-k}$ or the k th term of the expansion of $(2 + 1)^n$. Thus the number of k -faces is $\leq 3^n$:

$$\text{vol}(\text{proj}(U)) \leq \sum_{\text{all } k\text{-faces } F} \text{vol}(\text{proj}(F)) \leq 3^n.$$

For the lower bound, we consider the projection space Q and its orthogonal complement Q^\perp . Then $W = \text{proj}_Q(U) \times \text{proj}_{Q^\perp}(U)$ is an n -dimensional polyhedron that contains U . It follows that

$$\text{vol}(W) = \text{vol}(\text{proj}_Q(U)) \cdot \text{vol}(\text{proj}_{Q^\perp}(U)) \geq 1.$$

However, as we just showed, $\text{vol}(\text{proj}_{Q^\perp}(U)) \leq 3^n$, therefore

$$\text{vol}(\text{proj}_Q(U)) \geq 3^{-n}. \quad \square$$

Proof of Lemma 2.2. The singular-value decomposition of a matrix allows us to express A as UDV^T , where U (resp. V) is the orthogonal matrix whose columns are the eigenvectors of AA^T (resp. $A^T A$) and D is the m -by- n diagonal matrix whose only nonzero entries are $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots$ (the singular values of A). Let L be the subspace spanned by the k eigenvectors of $A^T A$ corresponding to $\lambda_1, \dots, \lambda_k$. Consider the orthogonal projection of $[-1, 1]^n$ on L . By Lemma 2.3, projecting a cube from \mathbf{R}^n to \mathbf{R}^k creates a volume distortion factor which lies (conservatively) between 3^{-n} and 3^n . It follows that

$$\text{vol}(A \text{proj}_L[-1, 1]^n) = \sqrt{\lambda_1 \cdots \lambda_k} \text{vol}(\text{proj}_L[-1, 1]^n) \geq 2^k 3^{-n} \lambda_k^{k/2}. \quad (2)$$

Given any $x \in L$ and $y \in L^\perp$, $A^T A x$ lies in L and so $(Ax)^T (Ay) = (A^T A x)^T y = 0$. In fact, not only are AL and $A(L^\perp)$ orthogonal, but they span all $A\mathbf{R}^n$ and therefore $(AL)^\perp = A(L^\perp)$. It easily follows that

$$A \text{proj}_L[-1, 1]^n = \text{proj}_{AL} A[-1, 1]^n$$

and by (2)

$$\text{vol}(\text{proj}_{AL} A[-1, 1]^n) \geq 2^k 3^{-n} \lambda_k^{k/2}.$$

By Lemma 2.1, the projection of $A[-1, 1]^n$ on AL is covered by the projections of 2^n cubes of side length $2 \text{lindisc}(A)$. Accounting for the dilation factor of 3^n , we find that

$$\text{vol}(\text{proj}_{AL} A[-1, 1]^n) \leq 2^n 3^m (2 \text{lindisc}(A))^k,$$

which, together with $m \leq n$, completes the proof. □

We prove the trace bound by a probabilistic argument. Let λ be the random variable λ_i , where i is chosen at random uniformly in $\{1, \dots, n\}$.

Lemma 2.4. *If $n/k \geq (2 \mathbf{var} \lambda)/(\mathbf{E}\lambda)^2 + 1$, then $\lambda_k \geq \mathbf{E}\lambda/4$.*

Proof. We use a second-moment argument based on a one-way Chebyshev inequality.³ For $x \geq 0$, let G_x be the event, $\lambda \geq \mathbf{E}\lambda - x$, and let p be its probability:

$$\begin{aligned} 0 &= \mathbf{E}[\lambda \mid G_x] p + \mathbf{E}[\lambda \mid \bar{G}_x] (1 - p) - \mathbf{E}\lambda \\ &= (\mathbf{E}[\lambda \mid G_x] - \mathbf{E}\lambda) p + (\mathbf{E}[\lambda \mid \bar{G}_x] - \mathbf{E}\lambda) (1 - p) \\ &\leq (\mathbf{E}[\lambda \mid G_x] - \mathbf{E}\lambda) p - x(1 - p) \end{aligned}$$

and, hence,

$$\mathbf{E}[\lambda \mid G_x] \geq \mathbf{E}\lambda + x(1/p - 1).$$

Consider the random variable $\mathbf{E}[\lambda \mid Y]$, where Y is G_x with probability p and \bar{G}_x with probability $1 - p$. The conditional variance $\mathbf{var}[\lambda \mid Y]$, defined as the variance of $\mathbf{E}[\lambda \mid Y]$, can never exceed the (unconditional) variance (see pages 195–196 of [9]), and so

$$\begin{aligned} \mathbf{var} \lambda &\geq \mathbf{var} \mathbf{E}[\lambda \mid Y] = \mathbf{E} (\mathbf{E}[\lambda \mid Y] - \mathbf{E}\lambda)^2 \\ &\geq (\mathbf{E}[\lambda \mid G_x] - \mathbf{E}\lambda)^2 p + (\mathbf{E}[\lambda \mid \bar{G}_x] - \mathbf{E}\lambda)^2 (1 - p) \\ &\geq x^2(1/p - 1)^2 p + x^2(1 - p) \geq x^2(1/p - 1). \end{aligned}$$

It follows that

$$p \geq \frac{1}{1 + x^{-2} \mathbf{var} \lambda}. \quad (3)$$

Setting $x = 3\mathbf{E}\lambda/4$, we find that the number of eigenvalues greater than or equal to $\mathbf{E}\lambda/4$ is at least

$$\frac{n}{1 + (16/9\mathbf{E}^2\lambda) \mathbf{var} \lambda},$$

which completes the proof. \square

We can now prove Lemma 1.1. If we write $M = A^T A$, $t = \text{tr} M$, and $\sigma = \text{tr} M^2$, then $\mathbf{E}\lambda = t/n$ and $\mathbf{var} \lambda = \sigma/n - t^2/n^2$. Therefore, by setting n/k to about $2(\mathbf{var} \lambda/\mathbf{E}^2\lambda + 1) = 2n\sigma/t^2$, we derive⁴ from (1) and Lemmas 2.2 and 2.4:

$$\begin{aligned} \text{herdisc}(A) &\geq \frac{1}{2} \text{lindisc}(A) \geq \frac{1}{2} \cdot 18^{-n/k} \sqrt{\lambda_k} \\ &\geq \frac{1}{4} \cdot 18^{-2n\sigma/t^2} \sqrt{t/n}, \end{aligned}$$

which proves Lemma 1.1. \square

³ Interestingly, this cannot be derived directly from Chebyshev's inequality.

⁴ Note that by Cauchy–Schwarz we easily see that $2n\sigma/t^2 \geq 2$, so k remains within a valid range.

We now prove Lemma 1.2. By the *spectral lemma* [3], the circuit complexity of $x \mapsto Ax$ is at least $\Omega(k \log \lambda_k)$, where λ_k is the k th largest eigenvalue of $A^T A$. This result holds under the assumption that gates compute functions of the form $(u, v) \mapsto \alpha u + \beta v$, with $|\alpha|, |\beta| = O(1)$. Set $x = \varepsilon \sqrt{\text{tr } M^2/n}$ in (3). Then

$$\begin{aligned} & \text{Prob}\{\lambda \geq \text{tr } M/n - \varepsilon \sqrt{\text{tr } M^2/n}\} \\ & \geq \frac{1}{1 + n\varepsilon^{-2} \text{var } \lambda / \text{tr } M^2} \geq \frac{1}{1 + \varepsilon^{-2}}, \end{aligned}$$

which is independent of n . Setting $k = \lceil n/(1 + \varepsilon^{-2}) \rceil$, the spectral lemma yields an asymptotic circuit complexity lower bound of

$$(n/(1 + \varepsilon^{-2})) \log(\text{tr } M/n - \varepsilon \sqrt{\text{tr } M^2/n}),$$

which proves Lemma 1.2. □

Remark. The spectral lemma allows the use of help gates, which take two arguments and outputs any function of them. In the presence of h help gates, Lemma 1.2 still holds if we replace $n \log(\dots)$ by $(n - ch) \log(\dots)$, for some constant $c > 0$ depending on ε .

3. The Discrepancy for Lines

We consider the discrepancy of an n -by- n set system formed by points and lines in \mathbf{R}^2 . We prove a tight bound of $n^{1/6}$ (up to a polylogarithmic factor). This is a surprising result because it differs so noticeably from the $\Theta(n^{1/4})$ bound for points and halfplanes.

Theorem 3.1. *For any $n > 0$, there exists an n -by- n set system of points and lines in the plane with discrepancy $\Omega(n^{1/6})$. This is optimal up to a factor of $(\log n)^{2/3}$.*

Proof. We begin with the lower bound, whose one-paragraph proof illustrates the surprising power of the trace bound. A classical construction of Erdős (see [14]), whose description we omit here, produces an n -by- n set system of points and lines whose incidence matrix A has $\Theta(n^{1/3})$ ones per row (i.e., points per line) and $\Theta(n^{1/3})$ ones per column. Each diagonal element of $M = A^T A$ is $\Theta(n^{1/3})$, so the trace of M is $\Omega(n^{4/3})$. The squares of the elements along the diagonal sum up to $O(n^{5/3})$. Each nondiagonal element is 0/1 (at most one of the n lines passes through two distinct points), and the number of ones among them coincides with the number of pairs of points on any of the n lines, i.e., $O(n^{5/3})$. It follows that the trace of M^2 is $O(n^{5/3})$. By Lemma 1.1, it follows that

$$\text{herdisc}(A) \geq \frac{1}{4} c^{n^{1+5/3}/(n^{4/3})^2} \sqrt{n^{4/3}/n} = \Omega(n^{1/6}),$$

hence the lower bound.

We now establish an upper bound of $O(n^{1/6})(\log n)^{2/3}$ on the discrepancy of any n -by- n set system of points and lines. Our argument builds on two known facts. One is the Szemerédi–Trotter bound of $O(m^{2/3}n^{2/3} + m + n)$ on the number of incidences

between n points and m lines in the plane [14]–[16]. The other is Beck’s partial coloring lemma of discrepancy theory [1], [5], [12]: Let (X, \mathcal{R}_0) and (X, \mathcal{R}_1) be two set systems, and assume that

$$\prod_{R \in \mathcal{R}_0} (|R| + 1) \leq 2^{(n-1)/5},$$

where $n = |X|$, and that $|R| \leq r$ for any $R \in \mathcal{R}_1$. Then there exists a partial coloring $\chi: X \mapsto \{-1, 0, 1\}$, such that χ is nonzero over at least one-tenth of X , $\chi(R) = 0$ for each $R \in \mathcal{R}_0$, and $|\chi(R)| \leq \sqrt{2r \ln(4|\mathcal{R}_1|)}$, for each $R \in \mathcal{R}_1$; of course, $\chi(R)$ is defined as $\sum_{x \in R} \chi(x)$.

Consider a set system formed by n points and $m \geq n$ lines. We partition it into two subsystems: \mathcal{R}_0 includes the subsets with at least $r = c(n \log n)^{1/3}$ incidences, for some constant $c > 0$ large enough, while \mathcal{R}_1 consists of the others. Clearly, assuming that n is large enough,

$$r|\mathcal{R}_0| = O(n^{2/3}|\mathcal{R}_0|^{2/3} + n + |\mathcal{R}_0|),$$

and so $|\mathcal{R}_0| \leq n/6 \log n$ and, indeed,

$$\prod_{R \in \mathcal{R}_0} (|R| + 1) \leq (n + 1)^{n/6 \log n} < 2^{(n-1)/5}.$$

This implies the existence of a partial coloring with at least $n/10$ nonzeros and discrepancy at most proportional to $n^{1/6}(\log m)^{2/3}$. Next, we extract the points corresponding to the color 0, and we iterate through this partial coloring process. The successive discrepancies decrease geometrically, and adding them up increases the bound above by only a constant factor. □

4. The Discrepancy for Halfplanes

By way of illustration, here is another interesting application of the trace bound. We describe a set system of n points and n halfplanes with discrepancy $\Omega(n/\log n)^{1/4}$. The optimal bound [4] is $\Omega(n^{1/4})$, but the proof is purely existential. Our simple construction below illustrates purely by combinatorial means why $\frac{1}{4}$ is the right exponent.

We use strips instead of halfplanes. This is fine for lower bound purposes (up to a constant factor), since a strip is the difference between two halfplanes. Consider a $\lfloor \sqrt{n/\log n} \rfloor \times \lfloor \sqrt{n \log n} \rfloor$ square grid. The set of points consists of the centers of all the unit squares, so their number is $\Theta(n)$. Let $S_{u,v}$ denote the strip intersecting the top and bottom sides of the grid in the horizontal intervals $[u, u + 1]$ and $[u + v \log n, u + v \log n + 1]$, respectively. The set of strips consists of all the $S_{u,v}$ ’s, for integral u, v , that intersect the grid’s boundary only in the top and bottom sides: their number is clearly $\Theta(n)$. For any fixed v_0 , the strips S_{u,v_0} are disjoint while, for any two distinct v_0 and v_1 , the intersection of any two S_{u,v_0} and S_{u',v_1} can be covered with a parallelogram of x -extent 1 and y -extent $O(n^{1/2})/(|v_0 - v_1| \log^{3/2} n)$ and thus contains $O(n^{1/2})/(|v_0 - v_1| \log^{3/2} n)$ points. Also, given $v_1 < v_0$, the strip S_{u,v_0} cannot intersect more than $O(v_0 - v_1) \log n$ strips S_{u',v_1} inside the grid.

We can now bound the traces $t = \text{tr } M$ and $\sigma = \text{tr } M^2$ for this set system. For convenience, we consider $M = AA^T$ instead of $M = A^T A$. Each strip contains $\Theta(n/\log n)^{1/2}$

points, so the total number of incidences is $t = \Omega(n^{3/2})/\sqrt{\log n}$. Let L_{u_0, v_0} be the sum of the squares of the number of points in $S_{u_0, v_0} \cap S_{u, v}$ over all (u, v) such that $v \leq v_0$. Clearly, $\sigma \leq 2 \sum_{u_0, v_0} L_{u_0, v_0}$. Let P denote the set of points:

$$\begin{aligned} L_{u_0, v_0} &= \sum_{(u, v): v \leq v_0} |P \cap S_{u_0, v_0} \cap S_{u, v}|^2 \\ &= O(n/\log n) + 2 \sum_{v < v_0} O(\sqrt{n}/((v_0 - v) \log^{3/2} n))^2 (v_0 - v) \log n \\ &= O(n/\log n) + O(n/\log^2 n) \sum_{i=1}^{n^{1/2}} (1/i) = O(n/\log n). \end{aligned}$$

It follows that $\sigma = O(n^2/\log n)$. By Lemma 1.1,

$$\text{herdisc}(A) \geq \frac{1}{4} c^{n\sigma/t^2} \sqrt{t/n} = \Omega(n/\log n)^{1/4},$$

as claimed.

5. The Discrepancy for Axis-Parallel Boxes

Next, we consider the case of point systems formed by points and axis-parallel boxes in nonconstant dimension.

Theorem 5.1. *In any dimension $d = \Omega(\log n/\log \log n)$ there is a set system of n points and n axis-parallel boxes in \mathbf{R}^d with discrepancy $n^{\Omega(1/\log \log n)}$.*

Proof. It suffices to prove that, for any large enough n , there exists an n -by- n set system of points and axis-parallel boxes in $\mathbf{R}^{O(\log n/\log \log n)}$ with discrepancy $n^{\Omega(1/\log \log n)}$. (To see why, set $n' = n^\varepsilon$, for a small enough constant $\varepsilon > 0$ and pad the set system in the obvious way.)

As usual, it suffices to prove the theorem relative to the hereditary discrepancy; indeed, by moving the irrelevant points outside of the boxes, the hereditary discrepancy becomes the actual (standard) discrepancy. Given an integer p , define $s = \lfloor \sqrt{p}/\log p \rfloor$, $d = \lfloor s/5 \rfloor$, and $k = \lfloor d/2 \log p \rfloor$. To establish the theorem, we may restrict ourselves to n -by- n set systems where n is of the form p^s , for a large prime p . Indeed, given an arbitrary (large) value of n , we can apply the theorem for $n' = p^s$, where p is defined as the largest prime not exceeding $\log^2 n$. The resulting n' -by- n' set system is made n -by- n trivially by adding $n - n'$ empty boxes and far away points. This addition does not change the discrepancy. Furthermore, by the prime distribution law [8], it easily follows that $\sqrt{n} \leq n' \leq n$; therefore, the bounds on the dimension and the discrepancy claimed by the theorem (for n') remain true if we change n' into n .

So, from now on, we assume that $n = p^s$, with s, d, k as specified above, and p a large prime. The set of n points is a p -ary Faure set⁵ of p^s points in \mathbf{R}^d . We postpone a

⁵ For a detailed description of Faure sets, see, e.g., pages 51–54 of [12].

review of the properties of such points to the place in the proof where they are needed. To define our set B of n boxes, we consider all grids $\mathcal{G}_{i_1, \dots, i_d}$ in $[0, 1]^d$ consisting of cells of the size

$$\frac{1}{p^{i_1}} \times \frac{1}{p^{i_2}} \times \dots \times \frac{1}{p^{i_d}}, \quad i_1 + i_2 + \dots + i_d = s - k.$$

Note that the total number of such grids is $\binom{s-k+d-1}{d-1}$.

Each cell σ of $\mathcal{G}_{i_1, \dots, i_d}$ is a box that contains exactly p^k points. Let B be a set consisting of any $n = p^s$ of these boxes chosen from (any) p^k grids (we prove later that there are at least p^s such boxes). A few words of motivation for the choice of parameters d, k :

- The trace t of $M = A^T A$ is equal to np^k , so k should be as large as possible. However,
- as k grows, $d = d(k)$ should grow to ensure that our set system contains at least p^s boxes. At the same time,
- d should be kept small enough so as to bound by a constant the exponent $n\sigma/t^2$ in Lemma 1.1 (recall that $\sigma = \text{tr } M^2$).

We now prove that the total number of boxes is at least p^s , and so B is well defined. The total number of boxes in all the grids is equal to # grids $\times p^{s-k}$, so it is enough to show that the number of grids is at least p^k . As we observed earlier, this number is

$$\binom{s-k+d-1}{d-1} > \left(\frac{s}{2(d-1)}\right)^{d-1} > 2^d \geq 2^{\lfloor d/2 \log p \rfloor \log p} = p^k.$$

Next, we prove that the hereditary discrepancy of the set system is $n^{\Omega(1/\log \log n)}$, as desired. The number of points in each box is p^k and so the trace t satisfies

$$\sqrt{t/n} = p^{k/2} = p^{(1/2)\lfloor (1/(2 \log p)) \lfloor s/5 \rfloor \rfloor} \geq p^{s/21 \log p} = 2^{\Omega(s)} = n^{\Omega(1/\log \log n)}. \quad (4)$$

On the other hand, the trace of M^2 satisfies

$$\sigma \leq p^s \sum \{(\# \text{ inters})^2 \mid \text{relative to a fixed box and all boxes in } B\}.$$

Let boxes b_1 and b_2 belong to grids $\mathcal{G}_{i_1, \dots, i_d}$ and $\mathcal{G}_{j_1, \dots, j_d}$, respectively. Then the number of points in their intersection is at most $\max\{1, p^{k-\text{dist}/2}\}$, where dist is the L^1 -distance between the vectors (i_1, \dots, i_d) and (j_1, \dots, j_d) . So,

$$\sigma \leq p^s \left(\underbrace{p^{2k}}_{\text{self-inters}} + \underbrace{p^{2k-1}d^2 + p^{2k-2}d^4 + \dots}_{\text{inters. with grids at dist } \leq 2k} + \underbrace{p^k \cdot (\# \text{ of grids})}_{\text{inters. with other grids}} \right).$$

The first and the third items are equal to p^{2k} and the second item forms a convergent geometric series (because $d^2/p < \frac{1}{2}$), so

$$\sigma \leq O(p^{2k+s}) = O(t^2/n).$$

The lower bound follows from (4) and Lemma 1.1. □

In view of Lemma 1.2, the proof above immediately implies:

Theorem 5.2. *Off-line orthogonal range searching in \mathbf{R}^d has complexity $\Omega(n \log n / \log \log n)$ for any dimension $d = \Omega(\log n / \log \log n)$.*

Acknowledgments

We wish to thank Jirka Matoušek for reading a draft of this paper and making several useful comments and suggestions. We also thank the referees for helpful remarks.

References

1. Beck, J., Chen, W. W. L., *Irregularities of Distribution*, Cambridge Tracts in Mathematics, 89, Cambridge University Press, Cambridge, 1987.
2. Chazelle, B., Lower bounds for off-line range searching, *Discrete Comput. Geom.*, **17** (1997), 53–65.
3. Chazelle, B., A spectral approach to lower bounds with applications to geometric searching, *SIAM J. Comput.*, **27** (1998), 545–556.
4. Chazelle, B., Discrepancy bounds for geometric set systems with square incidence matrices, in *Advances in Discrete and Computational Geometry*, B. Chazelle, J. E. Goodman and R. Pollack, eds., Contemporary Mathematics, 223, American Mathematical Society, Providence, RI, 1999, pp. 103–107.
5. Chazelle, B., *The Discrepancy Method: Randomness and Complexity*, Cambridge University Press, Cambridge, 2000.
6. Chazelle, B., Lvov, A., A note on the discrepancy of boxes in higher dimension, *Discrete Comput. Geom.*, **25** (2001), 519–524.
7. Drmota, M., Tichy, R. F., *Sequences, Discrepancies and Applications*, Lecture Notes in Mathematics, 1651, Springer-Verlag, Berlin, 1997.
8. Huxley, M., On the difference between consecutive primes, *Invent. Math.*, **15** (1972), 164–170.
9. Kalbfleisch, J. G., *Probability and Statistical Inference*, Vol. 1, Springer-Verlag, New York, 1995.
10. Lovász, L., Spencer, J., Vesztergombi, K., Discrepancy of set systems and matrices, *European J. Combin.*, **7** (1986), 151–160.
11. Lovász, L., Vesztergombi, K., Extremal problems for discrepancy, in *Irregularities of Partitions*, G. Halász, V. T. Sós, eds., Academic Press, New York, 1989, pp. 107–113.
12. Matoušek, J., *Geometric Discrepancy: An Illustrated Guide*, Algorithms and Combinatorics, 18, Springer-Verlag, Berlin, 1999.
13. Montgomery, H. L., *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, CBMS Regional Conference Series in Mathematics, 84, American Mathematical Society, Providence, RI, 1994.
14. Pach, J., Agarwal, P. K., *Combinatorial Geometry*, Wiley, New York, 1995.
15. Sharir, M., Agarwal, P. K., *Davenport–Schinzel Sequences and Their Geometric Applications*, Cambridge University Press, Cambridge, 1995.
16. Szeremédi, E., Trotter, W., Extremal problems in discrete geometry, *Combinatorica*, **3** (1983), 381–392.

Received April 5, 2000, and in revised form October 17, 2000. Online publication June 22, 2001.