Equilateral sets in l_p^n

Noga Alon * Pavel Pudlák †

December 2, 2002

Abstract

We show that for every odd integer $p \geq 1$ there is an absolute positive constant c_p, so that the maximum cardinality of a set of vectors in R^n such that the l_p distance between any pair is precisely 1, is at most $c_p n \log n$. We prove some upper bounds for other l_p norms as well.

1 Introduction

An *equilateral set* (or a simplex) in a metric space, is a set A, so that the distance between any pair of distinct members of A is b, where $b \neq 0$ is a constant. Trivially, the maximum cardinality of such a set in R^n with respect to the (usual) l_2-norm is $n+1$. Somewhat surprisingly, the situation is far more complicated for the other l_p norms. For a finite $p > 1$, the l_p-distance between two points $\vec{a} = (a_1, \ldots a_n)$ and $\vec{b} = (b_1, \ldots, b_n)$ in R^n is $\|\vec{a} - \vec{b}\|_p = (\sum_{k=1}^n |a_i - b_i|^p)^{1/p}$. The l_∞-distance between $\vec{a}$ and $\vec{b}$ is $\|\vec{a} - \vec{b}\|_\infty = \max_{1 \leq k \leq n} |a_i - b_i|$. For all $p \in [1, \infty]$, l_p^n denotes the space R^n with the l_p-distance. Let $e(l_p^n)$ denote the maximum possible cardinality of an equilateral set in l_p^n.

Petty [6] proved that the maximum possible cardinality of an equilateral set in l_p^n is at most 2^n and that equality holds only in l_∞^n, as shown by the set of all 2^n vectors with $0,1$-coordinates. The set of standard basis vectors together with an appropriate multiple of the

*Institute for Advanced Study, Princeton, NJ 08540, USA and Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: nogaa@post.tau.ac.il. Research supported in part by a State of New Jersey grant, by a USA Israeli BSF grant, by a grant from the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

†Mathematical Institute, Academy of Sciences of the Czech Republic, Prague, Czech Republic. Email: pudlak@math.cas.cz. Research supported in part by grant No. A1019901 of the Academy of Sciences of the Czech Republic and by project No. LN00A056 of the Ministry of Education of the Czech Republic. Part of this research was done while the author was visiting IAS, Princeton.

all 1 vector shows that $e(l_p^n) \geq n+1$ for all $1 \leq p < \infty$, and the set of standard basis vectors and their negatives shows that $e(l_1^n) \geq 2n$.

Kusner conjectured that both examples are extremal.

Conjecture 1.1 (Kusner, [3]) *$e(l_1^n) = 2n$.*

Conjecture 1.2 (Kusner, [3]) *For every $1 < p < \infty$, $e(l_p^n) = n+1$.*

The assertion of Conjecture 1.1 is easy for $n \leq 2$, and has been proved for $n = 3$ in [2] and for $n = 4$ in [5]. For large n, the best known upper bound is $2^n - 1$, and the existing techniques supplied no nontrivial upper bound. Our main result here is a significant improvement of the trivial estimate.

Theorem 1.3 *There exists an absolute constant $c > 0$ such that the cardinality of any equilateral set of vectors in l_1^n is at most $cn \log n$, that is: $e(l_1^n) \leq cn \log n$.*

The assertion of Conjecture 1.2 is easy for $p = 2$, is proved in [7] for all real p sufficiently close to 2 by a continuity argument, and has also been proved by Swanepoel (c.f., [7]) for $p = 4$. Galvin (c.f., [7]) has shown that for every even integer $p \geq 2$, $e(l_p^n) \leq 1 + (p-1)n$. For general $p \neq 1$ it is proved in [7] that $e(l_p^n) \leq c(p)n^{(p+1)/(p-1)}$.

The proof of Theorem 1.3 can be extended, with some additional effort, to provide a similar bound for $e(l_p^n)$ for every odd integer p.

Theorem 1.4 *For every odd integer $p \geq 1$ there exists an absolute constant $c_p > 0$ such that the cardinality of any equilateral set of vectors in l_p^n is at most $c_p n \log n$, that is: $e(l_p^n) \leq c_p n \log n$.*

Combining our basic approach here with the technique of [7] we can slightly improve the known estimates for general non-integral p as well and prove the following.

Theorem 1.5 *For every $p \geq 1$ there exists an absolute constant $c = c(p) > 0$ such that $e(l_p^n) \leq cn^{(2p+2)/(2p-1)}$.*

Our proofs combine probabilistic, combinatorial and linear algebra tools with some results from approximation theory. The proof of Theorem 1.3 is presented in Section 2. We then describe, in Section 3, how to extend it and prove Theorem 1.4. Section 4 contains the short proof of Theorem 1.5. The final Section 5 contains some concluding remarks.

2 Equilateral sets in l_1^n

The basic approach

In this section we prove Theorem 1.3. Let A be a set of vectors in l_1^n such that the distance between any two is 1. Let $m = |A|$. Consider the $m \times m$ matrix M indexed by the vectors of A in which $M_{\vec{a},\vec{b}} = \|\vec{a} - \vec{b}\|_1$. By subtracting this matrix from the matrix of 1's we get the identity matrix I that has full rank m. Using the vectors $\vec{a} \in A$ we shall construct an approximation of M, hence also of I. The precision of approximation and the rank of the approximating matrix will depend on m. To get an upper bound on m we shall use the following well-known lemma, whose short proof is included, for the sake of completeness.

Lemma 2.1 *For any real symmetric matrix M,*

$$\operatorname{rank} M \geq \frac{(\sum_i M_{i,i})^2}{\sum_{i,j} M_{i,j}^2}. \tag{1}$$

Proof. Put $r = \operatorname{rank} M$ and let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be all nonzero eigenvalues of M. Then $\sum_{k=1}^r \lambda_k = \operatorname{trace}(M) = \sum_i M_{i,i}$ and $\sum_{k=1}^r \lambda_k^2 = \operatorname{trace}(M^2) = \sum_{i,j} M_{i,j}^2$. Therefore, by Cauchy Schwartz

$$\sum_{i,j} M_{i,j}^2 = \sum_{k=1}^r \lambda_k^2 \geq \frac{(\sum_{k=1}^r \lambda_k)^2}{r} = \frac{(\sum_i M_{i,i})^2}{r},$$

implying the desired result. □

As shown in [1], this lemma can be used for bounding the minimum possible dimension of an Euclidean space in which one can embed an m-simplex with low distortion, and it is therefore not surprising that it can be useful for our purpose here as well.

In our approximation of the identity, the diagonal elements will be close to 1, hence $(\sum_i M_{i,i})^2$ will be close to m^2, while $\sum_{i,j} M_{i,j}^2$ will be $O(m)$. Thus $\operatorname{rank} M$ will be essentially an upper bound on m. Hence to get a bound on the number of vectors we have to get a good upper bound on the rank of M. We shall construct M in such a way that $\operatorname{rank} M$ will be a function $N(n,m)$, thus the bound will be the maximum m such that $N(n,m)$ is greater or equal to the bound of Lemma 2.1. The approximation matrix will be a matrix of scalar products of vectors in a space of dimension $N(n,m)$.[1] It will be constructed by splitting each coordinate of the n coordinates of l_1 into several ones and by defining for each vector $\vec{a} \in A$ a corresponding vector $\vec{a^*}$ in dimension $N(n,m)$. Here are more details.

[1]This is a slight simplification of what we actually will do. In our construction the dimension will depend on the particular set A, so think of $N(n,m)$ as an upper bound on it.

Let us denote by $\delta(x)$ the following function

$$\delta(x) = \begin{matrix} 1 \text{ if } x > 0 \\ \\ 0 \text{ otherwise} \end{matrix}$$

We shall use the following equality

$$|x - y| = |x| + |y| - 2\delta(xy)\min(|x|, |y|). \tag{2}$$

We decompose the matrix of distances in l_1^n as follows

$$\|\vec{a} - \vec{b}\|_1 = \sum_i |a_i| + \sum_i |b_i| - 2\delta(a_i b_i)\sum_i \min(|a_i|, |b_i|).$$

The matrix $(\sum_i |a_i| + \sum_i |b_i|)_{\vec{a},\vec{b}}$ has rank at most 2, hence the essential part is the matrix $2\delta(a_i b_i)(\sum_i \min(|a_i|, |b_i|))_{\vec{a},\vec{b}}$. Our goal can now be stated as follows. We want to assign a vector $\vec{a^*}$ to every vector $\vec{a} \in A$ in order to approximate $\sum_i \delta(a_i b_i)\min(a_i, b_i)$ by the scalar product $(\vec{a^*}, \vec{b^*})$.[2] The aim is to get a good approximation of the identity matrix

$$I = J - \left(\sum_i |a_i| + \sum_i |b_i|\right)_{\vec{a},\vec{b}} + 2\left(\sum_i \delta(a_i b_i)\min(|a_i|, |b_i|)\right)_{\vec{a},\vec{b}},$$

where J is the $m \times m$ matrix of 1's, by the matrix

$$I^* = J - \left(\sum_i |a_i| + \sum_i |b_i|\right)_{\vec{a},\vec{b}} + 2\left((\vec{a^*}, \vec{b^*})\right)_{\vec{a},\vec{b}}.$$

Then we want to use the inequality (1) as stated above, which means that the diagonal elements of I^* should differ from 1 by a constant less than 1, say $2/3$, and the sum of the squares of all (or all off-diagonal) elements should be $O(m)$. This is the same as saying that $I - I^*$ has elements in the range $[1/3, 5/3]$ on the main diagonal and the sum of the squares of all off-diagonal elements is $O(m)$. Taking the right hand parts of the two equalities above this translates to

1. for every $\vec{a} \in A$,

$$\left|\sum_i \min(|a_i|, |a_i|) - (\vec{a^*}, \vec{a^*})\right| = \left|\sum_i |a_i| - \|\vec{a^*}\|_2^2\right| \le \frac{1}{3};$$

2.

$$\sum_{\vec{a} \ne \vec{b}} \left(\sum_i \delta(a_i b_i)\min(|a_i|, |b_i|) - (\vec{a^*}, \vec{b^*})\right)^2 = O(m).$$

[2]To be more precise: we will approximate it by the difference between two such scalar products.

We shall call $I - I^* = (\sum_i \delta(a_i b_i)\min(|a_i|,|b_i|) - (\vec{a^*},\vec{b^*}))_{\vec{a},\vec{b}}$ the *matrix of error terms.* Let us note that the rank of I^* is at most the dimension of the space of the vectors $\vec{a^*}$ plus 2.

Splitting coordinates will be done as follows. Without loss of generality we may assume that A contains the zero vector (by picking any of the vectors and shifting all by it). This zero vector will be a crucial reference point in the process of splitting coordinates. First we shall dispose of the annoying function δ. This is easy, split each coordinate i into two and represent a_i by $a^*_{i,1}, a^*_{i,2}$, where $a^*_{i,1} = a_i$ and $a^*_{i,2} = 0$ if $a_i \geq 0$, and $a^*_{i,1} = 0$ and $a^*_{i,2} = |a_i|$ otherwise. Then

$$\delta(a_i b_i)\min(|a_i|,|b_i|) = \sum_{j=1,2} \min(a^*_{i,j}, b^*_{i,j})$$

Notice that the new vectors have only nonnegative entries, in fact all $\vec{a} \in [0,1]^n$ because of the zero vector. Thus we have reduced our problem to approximating a sum of minima of coordinates of nonnegative vectors.

This was only the first stage of splitting coordinates, we need splitting coordinates into more pieces in order to get a good approximation. As we do not want to overload notation by too many indices, we shall assume that already the vectors in A are nonnegative. Next, for each coordinate $i = 1,\ldots,n$, we divide the interval $[0,1]$ into some intervals

$$[0,u_1],(u_1,u_2],\ldots,(u_{\nu_i},1]$$

(we shall use different partitions for different i's). Each interval will correspond to a new coordinate. Given a vector $\vec{a} \in A$, its "approximation" $\vec{a^*}$ will be defined as follows. The value a_i on the coordinate i will be replaced by ν_i values $a^*_{i,1},\ldots,a^*_{i,\nu_i}$. The basic idea of this transformation is to take t such that $u_{t-1} < a_i \leq u_t$ and define

1. $a^*_{i,j} = \sqrt{u_j - u_{j-1}}$, for $j < t$;
2. $a^*_{i,j} = 0$, for $j > t$;
3. $a^*_{i,t}$ will be a suitable value between 0 and $\sqrt{u_t - u_{t-1}}$.

Let $\vec{b}$ be another vector from A and let us see how the scalar product $(\vec{a^*},\vec{b^*})$ approximates the sum $\sum_i \min(a_i,b_i)$. Since both expressions are sums, we can compare them locally. Fix an i and suppose $a_i = \min(a_i,b_i)$. The corresponding term in $(\vec{a^*},\vec{b^*})$ is

$$\sum_{j=1}^{\nu_i} a^*_{i,j}b^*_{i,j} = (\sqrt{u_1 - 0})^2 + (\sqrt{u_2 - u_1})^2 + \ldots + (\sqrt{u_{t-1} - u_{t-2}})^2 + a^*_{i,t}b^*_{i,t} = u_{t-1} + a^*_{i,t}b^*_{i,t}.$$

Note that $|a^*_{i,t}b^*_{i,t}| \leq |u_t - u_{t-1}|$, thus the error is at most $u_t - u_{t-1}$. We shall reduce the error further by choosing the last nonzero component in a suitable way.

Applying this in the most straightforward way gives a rather poor upper bound (which is still far better than the previously known bounds) $m = O(n^4)$. To get a better bound we use the following four ideas.

1. Recall that $\vec{0} \in A$ and all $\vec{a} \in [0,1]^n$. Thus, in particular, for all $\vec{a} \in A$, $\vec{a} \neq \vec{0}$,

$$\sum_i a_i = 1. \tag{3}$$

Therefore most of the components of $\vec{a}$ will be close to 0. Namely, at most $1/d$ of the a_i's will be $\geq d$, for $0 < d \leq 1$. Since large a_i's do not occur frequently, we do not have to approximate them very precisely. Thus we may reduce the number of intervals by having intervals of size increasing with the distance from 0. However the distance from 0 is not crucial, what we need is that the interval is not hit by many a_i's. Therefore we adjust the length of the interval to the density of values of vectors. The average length of the intervals will be automatically larger further from 0.

2. Let $\vec{a}, \vec{b}, i$ and t be as above and suppose, moreover, that b_i does not fall in the same interval as a_i, namely $u_t < b_i$. Then

$$\sum_{j=1}^{\nu_i} a^*_{i,j} b^*_{i,j} = u_{t-1} + a^*_{i,t}\sqrt{u_t - u_{t-1}}.$$

Hence choosing $a^*_{i,t} = (a_i - u_{t-1})/\sqrt{u_t - u_{t-1}}$ ensures that this scalar product is equal exactly to $a_i = \min(a_i, b_i)$.

3. Another strategy for choosing the value $a^*_{i,t}$ is to choose randomly one of the two values 0 or $\sqrt{u_t - u_{t-1}}$ hoping that the errors will partially cancel out. The right way of doing this *randomized rounding* is to choose randomly and uniformly a threshold $\tau \in [u_{t-1}, u_t]$, one for all $\vec{a} \in A$ with a_i in the interval, and round down, i.e., put $a^*_{i,t} = 0$ if $a_i \leq \tau$, and round up, i.e., put $a^*_{i,t} = \sqrt{u_t - u_{t-1}}$ otherwise.

Let $\vec{a}, \vec{b} \in A$, let i be given and suppose $a_i \leq b_i$. Let X_i denote the error on coordinate i, which is $X_i = |\min(a_i, b_i) - \sum_{j=1}^{\nu_i} a^*_{i,j} b^*_{i,j}|$. Assume, moreover, that both a_i and b_i are in the same interval and randomized rounding is used for both. (The bound below holds also in other cases, but we do not need them.)

Lemma 2.2 *$E(X_i) = 0$ and $E(X_i^2) = Var(X_i) \leq \frac{1}{4}(u_t - u_{t-1})^2$.*

The first claim follows from the observation that because $a_i \leq b_i$ we have always

$$\sum_{j=1}^{\nu_i} a^*_{i,j} b^*_{i,j} = u_{t-1} + a^*_{i,j}\sqrt{u_t - u_{t-1}}.$$

Since $E(a^*_{i,j}) = (a_i - u_{t-1})/\sqrt{u_t - u_{t-1}}$ we get the first part of the lemma. The second part follows from the following:

Fact: Let Y be a random variable taking values in the interval $[a, b]$. Then the variance of Y is at most $(b-a)^2/4$.

Proof: Define $Z = \frac{2Y-a-b}{b-a}$, then $-1 \le Z \le 1$ and hence $E(Z^2) \le 1$. But

$$\frac{4}{(b-a)^2} Var(Y) = Var(Z) \le E(Z^2) \le 1,$$

as needed. □

We need to estimate the square of the sum of errors on all coordinates. If we used randomized rounding everywhere, this would be $(\sum_i X_i)^2$. Because X_i and X_j are independent random variables for $i \ne j$ and $E(X_i) = 0$, $E((\sum_i X_i)^2) = \sum_i E(X_i^2)$ (in other words $Var(\sum_i X_i) = \sum_i Var(X_i)$). Using the linearity of expectation, the expected sum of the squares of error terms for all pairs $\vec{a}, \vec{b}$ would be bounded by the sum of such terms. Hence for some choice of thresholds τ the sum of squares would be at most this.

4. The method of 2 ensures no error when b_i does not fall in the same interval as a_i, but it is worse than the randomized rounding of 3 otherwise. On the other hand, randomized rounding introduces errors always. To get the least possible error using these two methods we would like to combine them so that we get the benefits of both. Surprisingly it is possible. We shall triple the number of coordinates and denote the values $a^*_{i,t,s}$, where $s = 1, 2, 3$. For $s = 1$ we shall define the values as in 2; so these serve to get zero error when a_i and b_i hit different intervals. The second set of coordinates will serve us to remove the effect of the first one when a_i and b_i hit the same interval. So $a^*_{i,t,2} = 0$ if $a_i \notin [u_{t-1}, u_t]$, and $a^*_{i,t,2} = (a_i - u_{t-1})/\sqrt{u_t - u_{t-1}}$ if $a_i \in [u_{t-1}, u_t]$; furthermore the scalar products of these coordinates will be taken with the negative sign (think of subtracting this matrix of scalar products). The last set of coordinates will be used to introduce the randomized rounding of 3 in case a_i and b_i hit the same interval. This will be done in the same manner as in 3 with the difference that we put 0 on *all* coordinates that correspond to intervals not hit by a_i. The resulting effect will be exactly what we wanted: zero error when a_i and b_i hit different intervals, and the error estimated in 3 when they do hit the same interval.

The choice of the intervals

1. In each coordinate divide $[0, 1/n]$ into $\sqrt{m}/\sqrt{n}$ equal parts of size $1/\sqrt{nm}$. The number of such intervals summed over all coordinates is $\sqrt{nm}$, which is $o(m)$, if $n = o(m)$. These intervals will be called *base intervals*. To avoid the problem with nonintegral values, we can assume without loss of generality that n and m are powers of 4.

2. For every coordinate i we divide the rest of the interval as follows. We start from 1 and go down until the length of the interval times the number of hits is at least $cn\log n/m$, where c is a suitable constant that we will determine at the end of proof, or until we reach $1/\sqrt[3]{2}$, whichever comes first. Then we continue in the same way, i.e., until the product is at least $cn\log n/m$ or we reach some $2^{-t/3}$, for some $t = 1,\ldots,3\log n$. In the first case we call the interval *regular*, in the second case *singular*.

Let $(i_1,(u_1,v_1]),\ldots,(i_K,(u_K,v_K])$ be an arbitrary enumeration of pairs of the form a coordinate i and a non-base interval chosen on this coordinate, such that regular intervals come first. Denote by k the number of regular intervals, let $\ell_j = v_j - u_j$ be the length of the j-th interval and let h_j be the number of times it is hit, i.e., h_j is the number of vectors $\vec{a}\in A$ such that $a_{i_j}\in(u_j,v_j]$. Thus we have

$$\ell_j h_j \geq \frac{cn\log n}{m}, \tag{4}$$

for regular intervals, and

$$\ell_j h_j \leq \frac{2cn\log n}{m}, \tag{5}$$

for all intervals. The number of singular intervals over all coordinates is bounded by $3n\log n$.

Computation

The main point is to upper bound the number of regular intervals. To this end we derive several inequalities. Denote by t_j the number t such that $(u_j,v_j]\subseteq[2^{-t/3},2^{-(t-1)/3}]$. First we prove

$$\sum_j 2^{-t_j/3}h_j \leq m. \tag{6}$$

This is a consequence of (3):

$$\sum_j 2^{-t_j/3}h_j \leq \sum_j u_j h_j \leq \sum_j \sum_{\vec{a};\ a_{i_j}\in(u_j,v_j]} a_{i_j} \leq \sum_{i,\vec{a}} a_i \leq m.$$

Next we prove

$$\sum_{j=1}^{k} \frac{1}{2^{-t_j/3}h_j} \leq \frac{3m}{c}. \tag{7}$$

Since intervals are disjoint (except for endpoints), by considering the volume we get

$$\sum_{j;\ t_j=t} \ell_j \leq 2^{-t/3}n.$$

By (4), we get

$$\ell_j \geq \frac{cn\log n}{h_j m},$$

for regular intervals ($j \le k$). Hence

$$\sum_{j;\ t_j=t} \frac{cn\log n}{h_j m} \le 2^{-t/3} n,$$

where the sum is only over regular intervals. Thus, again summing only over regular intervals,

$$\sum_{j;\ t_j=t} \frac{1}{2^{-t/3} h_j} \le \frac{m}{c\log n}.$$

Summing over $t = 1, \dots, 3\log n$, we get (7).

To estimate k, the number of regular intervals in all coordinates, we use the Cauchy-Schwartz inequality for $x_j^2 = 2^{-t_j/3} h_j$ with (6) and (7).

$$k^2 = \left(\sum_{j=1}^{k} \frac{1}{x_j} x_j\right)^2 \le \left(\sum_{j=1}^{k} \frac{1}{x_j^2}\right)\left(\sum_{j=1}^{k} x_j^2\right) \le \frac{3m^2}{c}.$$

Hence

$$k \le \sqrt{\tfrac{3}{c}} \cdot m. \tag{8}$$

We shall estimate the error terms on diagonal elements (i.e., we shall estimate the diagonal elements of the matrix $I - I^*$). Let $\vec{a} \in A$. Let i be a coordinate and let $a_i \in (u, v]$ or $\in [0, v]$. The error on this coordinate is at most the length of the interval (this is achieved when a_i is one endpoint and the rounding gives the other).

First suppose that $(u, v]$ or $[0, v]$ is a base interval. Its length is $1/\sqrt{nm}$. Thus the error that all coordinates in which a_i falls into a base interval can make is at most $\sqrt{n/m}$, which is $o(1)$ if $n = o(m)$.

Now suppose $(u, v]$ is a regular or singular interval. Then $(u, v] \subseteq [2^{-t/3}, 2^{-(t-1)/3}]$, for some t, hence $v \le \sqrt[3]{2}u$. Thus the length of the interval is

$$v - u \le \sqrt[3]{2}u - u = (\sqrt[3]{2} - 1)u \le (\sqrt[3]{2} - 1)a_i.$$

Hence the error over all coordinates where $a_i \ge 1/n$ is

$$\le \sum_i (\sqrt[3]{2} - 1)a_i = (\sqrt[3]{2} - 1).$$

Thus for n sufficiently large the error is at most $1/3$.

Now we estimate the sum of squares of all off-diagonal elements of I^*, which is also the sum of squares of all off-diagonal errors. Let $\vec{a}, \vec{b} \in A$ be two different vectors. Recall that the expected value of the square of the error for such a pair is bounded by $1/4$ of the sum of the squares of the lengths of intervals that are hit by both vectors. We can choose the rounding thresholds so that we get at most the expected value, hence it suffices to estimate the sum of the squares of such intervals.

The sum of the squares of the lengths of base intervals that are hit by such a pair $\vec{a}, \vec{b}$ is at most

$$n \cdot \left(\frac{1}{\sqrt{nm}}\right)^2 = \frac{1}{m}$$

Thus the base intervals contribute totally at most $\frac{1}{4}m$. To estimate the contribution of non-base intervals, observe that each such interval contributes the square of its length times the number of ordered pairs of different vectors that hit the interval. So this contribution is

$$\sum_{j=1}^{K}(h_j^2 - h_j)\ell_j^2 \le \sum_{j=1}^{K}(h_j\ell_j)^2 \le K\left(\frac{2cn\log n}{m}\right)^2$$

Hence, for a suitable choice of rounding thresholds the sum of the squares of all off-diagonal elements is at most

$$\tfrac{1}{4}m + K\left(\frac{2cn\log n}{m}\right)^2 \tag{9}$$

Now we are ready to apply Lemma 2.1. Recall that the matrix of scalar products is multiplied by 2, hence our estimates of the diagonal elements and the sum of the squares of the off-diagonal elements must be multiplied by 2. Thus for diagonal elements we get a lower bound 1/3 and an upper bound 5/3 (assuming $n = o(m)$ and m sufficiently large). For the sum of squares of off-diagonal elements the upper bound is twice the expression (9). On the other hand, I^* is the sum of a matrix of rank 2 and a matrix of scalar products (more precisely a sum of such matrices). Hence I^* has rank 2 plus at most the number of coordinates of vectors that we use in these scalar products, which is the number of intervals times 6, since we have doubled the number of coordinates in order to get rid of negative signs and then tripled again to combine our two methods in an optimal way. Substituting in Lemma 2.1, we get the following inequality:

$$\frac{(\frac{1}{3}m)^2}{m(\frac{5}{3})^2 + \frac{1}{2}m + 2K\left(\frac{2cn\log n}{m}\right)^2} \le 6K + 2.$$

Recall that we have proved (see (8))

$$K \le k + 3n\log n \le \sqrt{\tfrac{3}{c}} \cdot m + 3n\log n,$$

hence we get

$$\frac{(\frac{1}{3}m)^2}{m(\frac{5}{3})^2 + \frac{1}{2}m + 2\left(\sqrt{\frac{3}{c}} \cdot m + 3n\log n\right)\left(\frac{2cn\log n}{m}\right)^2} \le 6\left(\sqrt{\tfrac{3}{c}} \cdot m + 3n\log n\right) + 2.$$

Let us write it in a more conspicuous form

$$\frac{(\frac{1}{3}m)^2}{m(\frac{5}{3})^2 + \frac{1}{2}m + 2\left(\sqrt{\frac{3}{c}} \cdot m + 3n\log n\right)\left(\frac{2cn\log n}{m}\right)^2} - 6\sqrt{\frac{3}{c}} \cdot m - 2 \le 18n\log n.$$

Then it is clear that if we take c sufficiently big, it implies a bound $m = O(n\log n)$. This completes the proof of Theorem 1.3. □

3 Extending the bound to all odd p

In this section we extend our $O(n\log n)$ bound to all odd integers $p > 1$ by proving Theorem 1.4. Since the proof is similar to that of Theorem 1.3, we only sketch it, highlighting the differences. First we shall explain the idea for $p = 3$. Recall that we have reduced approximating l_1 norm to approximating a sum of minima. We shall do the same for l_3 using the following formula

$$|x-y|^3 = |x^3| + |y^3| - 2\delta(xy)\min(|x^3|,|y^3|) - 3|x|xy - 3x|y|y + 6\delta(xy)\min(|x^2y|,|xy^2|),$$

(see below for a derivation of the general formula for all odd integers). Note, first, that the contribution of all terms with no minima to the rank of the approximating matrix is $O(n)$, as these can be expressed precisely by inner products of vectors of length $O(n)$. Again we shall assume that the zero vector is present in the set. When approximating $\sum_i \min(a_i, b_i)$ we only used that the distance from the zero vector was 1. We used the fact that other distances are 1 only in the final step, where we estimated the rank of the approximating matrix. So to approximate the matrix of $\sum_i \min(|a_i^3|, |b_i^3|)$ we only need the following equality

$$\sum_i |a_i^3| = 1. \tag{10}$$

Again, to get rid of the function δ, we split coordinates so that different ones are used for positive and negative values of a_i, so we can think of just having only positive values of a_i's. Then we choose intervals in the same way as we did for a_i, but now we do it for a_i^3. The only difference will be that we need to get a better precision for approximating error terms on the diagonal, since the matrices of minima are with total weight 8 (instead of 2 in the case of $p = 1$). Thus instead of $1, 2^{-1/3}, 2^{-2/3}\ldots$ we need a finer division. Therefore we replace $2^{-t/3}$ by $2^{-\gamma t}$ for $\gamma > 0$ such that $2^\gamma - 1 < 1/8$ (eg., $\gamma = 1/8$ will do; for p larger we will need an appropriately smaller constant). The bound on the resulting error of approximating $\sum_i \min(a_i^3, b_i^3)$ will be the same up to a multiplicative constant.

Next we have to approximate $\sum_i \min(a_i^2 b_i, a_i b_i^2)$. For this we shall use the same choice of intervals, so the number of coordinates to which we split the original n coordinates will be the same. (Thus, the total bound on the number of coordinates will be only twice bigger than it was in case of the l_1 norm.) But the values that we shall use will be different. Let $\vec{a} \in A$ be given, and let i be a coordinate. Let t be such that $u_{t-1} < a_i^3 \leq u_t$. Then we define

1. $a_{i,j}^* = a_i\sqrt{\sqrt[3]{u_j} - \sqrt[3]{u_{j-1}}}$, for $j < t$;
2. $a_{i,j}^* = 0$, for $j > t$;
3. $a_{i,t}^*$ will be 0 or $a_i\sqrt{\sqrt[3]{u_t} - \sqrt[3]{u_{t-1}}}$ in case of randomized rounding, and
$$a_i\frac{a_i - \sqrt[3]{u_j}}{\sqrt{\sqrt[3]{u_j} - \sqrt[3]{u_{j-1}}}}$$
otherwise.

We can view it also as if we scaled up the intervals from $[u, v]$ to $[\sqrt[3]{u}, \sqrt[3]{v}]$ and multiplied the original values by a_i. Clearly, if a_i^3 and b_i^3 are in different intervals, we get the value $\min(a_i^2 b_i, a_i b_i^2)$ precisely. If $u_{t-1} < a_i^3, b_i^3 \leq u_t$, the error is at most

$$a_i b_i(\sqrt[3]{u_t} - \sqrt[3]{u_{t-1}}) = a_i b_i\sqrt[3]{u_t} - a_i b_i\sqrt[3]{u_{t-1}} \leq u_t - u_{t-1}.$$

This guarantees that the error on diagonal elements will be bounded by the same term as for cubes, namely, it will be less than $2/3$ for n sufficiently big.

In randomized rounding there is a small difference between what we do for cubes and what we do for terms $a_i^2 b_i$. When doing randomized rounding for cubes, we take randomly and uniformly a threshold τ between u and v and round up if $\tau \leq a_i^3$. When doing it for terms $a_i^2 b_i$ we take randomly and uniformly τ between $\sqrt[3]{u}$ and $\sqrt[3]{v}$ and round up if $\tau \leq a_i$. The effect will be again zero expected error and the variance at most $1/4$ of

$$(a_i b_i(\sqrt[3]{u_t} - \sqrt[3]{u_{t-1}}))^2 \leq (u_t - u_{t-1})^2,$$

the square of the interval. Hence the estimate of the error will be again at most constant times bigger than in the case of l_1 norm.

When applying Lemma 2.1, it is clear that the constant time increase in the number of coordinates and in the errors can be compensated by a larger constant in the estimate for m. Thus we get an $O(n \log n)$ upper bound for the size of equilateral sets in l_3.

Turning now to a general odd positive integer, let q be an arbitrary natural number. We notice that

$$\begin{aligned}(x-y)^{2q+1} &= \sum_{r=0}^{q}(-1)^r\binom{2q+1}{r}(x^{2q+1-r}y^r - x^r y^{2q+1-r}) \\ &= \sum_{r=0}^{q}(-1)^r\binom{2q+1}{r}(xy)^r(x^{2q+1-2r} - y^{2q+1-2r}).\end{aligned}$$

Since for odd t, $|x^t - y^t|$ is $x^t - y^t$ if $x \geq y$ and it is $y^t - x^t$ otherwise, we get

$$|x-y|^{2q+1} = \sum_{r=0}^{q} (-1)^r \binom{2q+1}{r} (xy)^r |x^{2q+1-2r} - y^{2q+1-2r}|.$$

Using our formula (2) for the absolute value of a difference, we get

$$|x-y|^{2q+1} = \sum_{r=0}^{q} (-1)^r \binom{2q+1}{r} (xy)^r \left(|x^{2q+1-2r}| + |y^{2q+1-2r}| - 2\delta(xy) \min(|x^{2q+1-2r}|, |y^{2q+1-2r}|) \right).$$

Since $\delta(xy)(xy)^r = \delta(xy)|xy|^r$, we get finally $|x-y|^{2q+1} =$

$$\sum_{r=0}^{q} (-1)^r \binom{2q+1}{r} \left(|x^{2q+1-2r}| x^r y^r + x^r y^r |y^{2q+1-2r}| - 2\delta(xy) \min(|x^{2q+1-r} y^r|, |x^r y^{2q+1-r}|) \right).$$

So again we only need to approximate the minima. As in the case $p = 3$, the number of coordinates and the error will be only constantly larger than it was estimated in the proof for $p = 1$. Hence the bound $O(n \log n)$ follows along the same lines.

4 A general upper bound

In this section we prove Theorem 1.5. The proof is short, and combines the argument of Smyth in [7] with Lemma 2.1 here. We need the following result in approximation theory, derived in [7] as a simple consequence of the classical result of Jackson [4].

Lemma 4.1 *For every fixed real $p \geq 1$ and for every $d \geq \lceil p \rceil$ there is a polynomial g of degree at most d such that*

$$|g(t) - |t|^p| \leq B(p)/d^p$$

for all $t \in [-1, 1]$, where $B(p) = (\lceil p \rceil^p (1 + \pi^2/2)^{\lceil p \rceil} (p)_{\lceil p \rceil - 1}) / \lceil p \rceil !$.

Fix $p > 1$, and let

$$A = \{\vec{a}^{(i)} = (a_1^{(i)}, \ldots, a_n^{(i)}) : \ 1 \leq i \leq m \ \}$$

be a set of m vectors in R^n so that the l_p-distance between any pair of vectors in A is 1.

Let $g(t)$ be a polynomial of degree d satisfying $|g(t) - |t|^p| \leq \frac{1}{nm^{1/2}}$ for all $t \in [-1, 1]$, where $d \leq c(p) n^{1/p} m^{1/(2p)}$. Such a polynomial exists by Lemma 4.1.

Define an m by m matrix matrix $M = (M_{i,j})$ by

$$M_{i,j} = 1 - \|\vec{a}^{(i)} - \vec{a}^{(j)}\|_p^p = 1 - \sum_{k=1}^{n} |a_k^{(i)} - a_k^{(j)}|^p.$$

Obviously this matrix is simply the m by m identity matrix. Define an approximation M^* of M by

$$M^*_{i,j} = 1 - \sum_{k=1}^{n} g(a_k^{(i)} - a_k^{(j)}).$$

By the properties of the polynomial g, $|M_{i,j} - M^*_{i,j}| \le 1/\sqrt{m}$ for all i, j. It thus follows from Lemma 2.1 that the rank of M^* satisfies

$$\mathrm{rank}(M^*) \ge \frac{(m(1-1/\sqrt{m}))^2}{m(1+1/\sqrt{m})^2 + m(m-1)/m} \ge \frac{m}{2}(1-o(1)). \tag{11}$$

On the other hand, for each i, the row number i of M^* is the vector whose j-th coordinate is $g_i(a_1^{(j)}, \ldots, a_n^{(j)})$, where g_i is the polynomial

$$g_i(x_1, \ldots, x_n) = 1 - \sum_{k=1}^{n} g(a_k^{(i)} - x_k).$$

Since all these polynomials lie in the linear span of the members of

$$U = \left\{1,\ \textstyle\sum_{k=1}^{n} x_i^d,\ x_1\ , x_1^2,\ \ldots\ , x_1^{d-1},\ x_2,\ x_2^2,\ \ldots\ , x_2^{d-1},\ \ldots\ , x_n,\ x_n^2,\ \ldots\ , x_n^{d-1}\right\},$$

it follows that the dimension of the row space of M^* is at most $2 + (d-1)n \le dn$. Indeed, each row of M^* is a linear combination of the $2 + (d-1)n$ vectors

$$\left\{\ \left(f(a_1^{(1)}, a_2^{(1)}, \ldots, a_n^{(1)}), (f(a_1^{(2)}, a_2^{(2)}, \ldots, a_n^{(2)}), \ldots, (f(a_1^{(m)}, a_2^{(m)}, \ldots, a_n^{(m)})\right)\ :\ f \in U\ \right\}.$$

Combining this with (11) and the fact that $d \le c(p) n^{1/p} m^{1/(2p)}$ we conclude that

$$\frac{m}{2}(1-o(1)) \le c(p) n^{1/p} m^{1/(2p)} n,$$

implying that $m \le c'(p) n^{(2p+2)/(2p-1)}$. This completes the proof of Theorem 1.5. $\square$

5 Concluding remarks and open problems

- The proof of Theorem 1.5 works for $p = 1$ as well, and provides a short proof for the fact that $e(l_1^n) \le cn^4$ for some absolute constant c. The more complicated proof of Theorem 1.3 supplies a better estimate.

- Both conjectures of Kusner mentioned in the introduction remain open, and although the estimates we prove here may possibly be improved using similar techniques, it seems unlikely that the methods described here will suffice to prove the precise statements of the conjectures.

- Our methods here can be used to estimate the maximum possible cardinality of a set of vectors A in R^n, such that for every two distinct members $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ of A, the sum $\sum_i \phi(|a_i - b_i|)$ is a nonzero constant, where ϕ is a sufficiently smooth function satisfying $\phi(0) = 0$. The method of Section 4 works better for functions ϕ that have low degree polynomial approximations, and the quality of these approximations near 0 is the main factor effecting the bound obtained.

Acknowledgment: We would like to thank Cliff Smyth and Vojta Rödl for helpful discussions.

References

[1] N. Alon, Problems and results in extremal combinatorics, Part I, Discrete Math., to appear.

[2] H. J. Bandelt, V. Chepoi and M. Laurent, Embedding into rectilinear spaces, *Disc. and Comp. Geom.* **19** (1998), 595–604.

[3] R. Guy, editor, Unsolved Problems: An Olla-Podrida of Open Problems, Often Oddly Posed, *Amer. Math. Monthly* **90** (1983), 196–200.

[4] D. Jackson, The Theory of Approximation, American Mathematical Society, New York, 1930.

[5] J. Koolen, M. Laurent and A. Schrijver, Equilateral dimension of the rectilinear space, *Designs, Codes and Crypt.* **21** (2000), 149–164.

[6] C. M. Petty, Equilateral sets in Minkowski spaces, *Proc. Amer. Math. Soc.* **29** (1971), 369–374.

[7] C. Smyth, Equilateral or 1-distance sets and Kusner's conjecture, submitted.