SIAM J. DISCRETE MATH.
Vol. 24, No. 3, pp. 910–920


# DOMINATING SUBSETS UNDER PROJECTIONS*

ROM PINCHASI† AND ALLAN PINKUS†

**Abstract.** Let $d$ be a fixed integer, and let $W$ be any $d$-dimensional linear subspace of $\mathbb{R}^n$. There then exists a subset $I$ of the $n$ coordinates $\{1, 2, \ldots, n\}$ of $\mathbb{R}^n$ of cardinality at least $(\frac{1}{2} - o(1))n$ such that for every vector $w = (w_1, \ldots, w_n) \in W$ we have $\sum_{i\in I} |w_i| \le \sum_{i\notin I} |w_i|$. Equivalently, let $P$ be any multiset of $n$ arbitrary vectors in $\mathbb{R}^d$. Then there exists a subset $S$ of $P$ of size at least $(\frac{1}{2} - o(1))n$ such that for every vector $u \in \mathbb{R}^d$ we have $\sum_{x\in S} |\langle x, u\rangle| \le \sum_{x\in P\setminus S} |\langle x, u\rangle|$. A continuous analogue of the former result is also considered.

**Key words.** dominating subsets, discrepancy, $L^1$-approximation

**AMS subject classifications.** 41A45, 52C99, 41A50

**DOI.** 10.1137/090781164

## 1. Introduction.

Approximation in the $L^1$-norm from linear subspaces has various distinct properties. One such feature is concerned with the identification of large sets of functions for which the zero function is always a best approximant from the linear subspace. Let us explain this property in the particular case of $\mathbb{R}^n$.

Let $W$ be a $d$-dimensional subspace of $\mathbb{R}^n$ equipped with the $\ell_1^n$-norm $\|x\| = \sum |x_i|$. For each $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have that $w^* = (w_1^*, \ldots, w_n^*) \in W$ is a best $\ell_1^n$-approximant to $x$ from $W$ if and only if

$$\left| \sum_{\{i: x_i - w_i^* \neq 0\}} [\operatorname{sgn}(x_i - w_i^*)] w_i \right| \le \sum_{\{i: x_i - w_i^* = 0\}} |w_i|$$

for all $w \in W$; see, for example, [Pi89]. Thus the 0 vector is a best $\ell_1^n$-approximant to $x$ from $W$ if and only if

$$\left| \sum_{\{i: x_i \neq 0\}} [\operatorname{sgn} x_i] w_i \right| \le \sum_{\{i: x_i = 0\}} |w_i|$$

for all $w \in W$.

Let $I \subseteq \{1, \ldots, n\}$, and let $\operatorname{supp} x = \{i : x_i \neq 0\}$. It easily follows from the above characterization of a best approximant that the 0 vector is a best $\ell_1^n$-approximant to $x$ from $W$ for all $x \in \mathbb{R}^n$ with $\operatorname{supp} x \subseteq I$ if and only if

$$\sum_{i\in I} |w_i| \le \sum_{i\notin I} |w_i| \qquad (1)$$

for all $w \in W$. It is therefore natural to ask if such subsets $I$ exist and, if so, how large they can be. (In the subject of sparse representations one asks for conditions on $W$ ensuring that the above holds for all $I \subseteq \{1, \ldots, n\}$ of a given size.)

*Received by the editors December 28, 2009; accepted for publication (in revised form) June 3, 2010; published electronically August 10, 2010.
http://www.siam.org/journals/sidma/24-3/78116.html

†Mathematics Department, Technion—Israel Institute of Technology, Haifa 32000, Israel (room@math.technion.ac.il, pinkus@tx.technion.ac.il).

While there are $d$-dimensional linear subspaces $W$ for which there exist $I$ with $|I| = n-d$ satisfying the above (and this is the largest $|I|$ can possibly be), our interest is in determining lower bounds on $|I|$ that are valid for every $d$-dimensional subspace $W$ of $\mathbb{R}^n$. That is, given $d$ and $n$, we want to find the largest possible $k$ such that for every $d$-dimensional subspace $W$ of $\mathbb{R}^n$ there exists an $I \subseteq \{1,\dots,n\}$ satisfying $|I| \geq k$ such that (1) holds. A simple lower bound is the following.

PROPOSITION 1. *Let $W$ be any $d$-dimensional linear subspace of $\mathbb{R}^n$. Then there exists an $I \subseteq \{1,\dots,n\}$ with $|I| \geq [n/(d+1)]$ satisfying*

$$\sum_{i\in I} |w_i| \leq \sum_{i\notin I} |w_i|$$

*for all $w \in W$.*

*Proof.* Let $r = [n/(d+1)]$. We divide the $n$ indices $\{1,\dots,n\}$ into $r$ groups of $d+1$ distinct indices and discard whatever might be left over. Since $W$ is a subspace of dimension $d$, for each given group of indices $J = \{j_1,\dots,j_{d+1}\}$ there exists a nonzero $a \in \mathbb{R}^n$ of unit $\ell_\infty^n$ norm such that $a_i = 0$ for all $i \notin J$, and $\langle a, w\rangle = 0$ for all $w \in W$. Let $j_p$ be such that $1 = |a_{j_p}| \geq |a_{j_k}|$, $k = 1,\dots,d+1$. Thus we have

$$|w_{j_p}| = \left|\sum_{k=1,k\neq p}^{d+1} a_{j_k} w_{j_k}\right| \leq \sum_{k=1,k\neq p}^{d+1} |w_{j_k}|$$

for all $w \in W$. Take $I$ to be the $r$ distinct such $j_p$. ☐

The main result in this paper is the following improvement of this bound.

THEOREM 1. *Let $d$ be fixed, and let $W$ be any $d$-dimensional linear subspace of $\mathbb{R}^n$. Then there exists a subset $I$ of the $n$ coordinates $\{1,2,\dots,n\}$ of $\mathbb{R}^n$ of cardinality $(\frac{1}{2} - o(1))n$ such that for every vector $w = (w_1,\dots,w_n) \in W$ we have*

$$\sum_{i\in I} |w_i| \leq \sum_{i\notin I} |w_i|.$$

Notice that in Theorem 1 the cardinality of $I$ cannot exceed $n/2$ for all $W$. To see this take any $W$ that contains the 1-dimensional subspace spanned by the vector, all of whose entries are 1. If $n/d$ is an odd integer, then $|I|$ cannot exceed $n/2 - d/2$ for all $W$. To see this set $k = n/d$, and let $W$ be the $d$-dimensional linear subspace of $\mathbb{R}^n$ spanned by $w^1,\dots,w^d$, where for each $i = 1,\dots,d$ we let $w^i$ be the vector with 1's in coordinates $k(i-1)+1,\dots,ki$, and 0 elsewhere.

Theorem 1 is "equivalent" to its following dual form, which we prove in section 2.

THEOREM 2. *Let $P$ be a multiset of $n$ arbitrary vectors in $\mathbb{R}^d$. There exists a subset $S$ of $P$ of size at least $(\frac{1}{2} - o(1))n$ such that for every vector $u \in \mathbb{R}^d$ we have*

$$\sum_{x\in S} |\langle x,u\rangle| \leq \sum_{x\in P\setminus S} |\langle x,u\rangle|.$$

To see how Theorem 1 follows from Theorem 2 let $W$ be any $d$-dimensional linear subspace of $\mathbb{R}^n$. Take a basis $w^1 = (w^1_1,\dots,w^1_n),\dots,w^d = (w^d_1,\dots,w^d_n)$ for $W$, and consider the set $P$ of $n$ vectors in $\mathbb{R}^d$:

$$P = \{x^i = (w^1_i, w^2_i,\dots,w^d_i) : i = 1,\dots,n\}.$$

Let $S$ be the subset of $P$ guaranteed by Theorem 2, and take $I \subset \{1, \ldots, n\}$ such that $S = \{x^i : i \in I\}$. We claim that for every $w = (w_1, \ldots, w_n) \in W$ we have

$$\sum_{i \in I} |w_i| \leq \sum_{i \notin I} |w_i|.$$

Indeed, there exists a $u = (u_1, \ldots, u_d)$ such that $w = u_1 w^1 + \cdots + u_d w^d$. Therefore, reading this equality coordinatewise, for every $i = 1, \ldots, n$ we have $w_i = \langle x^i, u \rangle$. Thus

$$\sum_{i \in I} |w_i| = \sum_{x^i \in S} |\langle x^i, u \rangle| \leq \sum_{x^i \in P \setminus S} |\langle x^i, u \rangle| = \sum_{i \notin I} |w_i|.$$

By an analogous argument, Theorem 2 follows from Theorem 1.

**2. Proof of Theorem 2.** We begin with the following simple lemma.

LEMMA 1. *Let $P$ be a finite set of vectors in $\mathbb{R}^d$. Then one can find $d$ distinct vectors $x^1, \ldots, x^d$ in $P$ with the property that for every $u \in \mathbb{R}^d$ and every $x \in P$ we have*

$$|\langle x, u \rangle| \leq \sum_{i=1}^{d} |\langle x^i, u \rangle|.$$

*Proof.* Without loss of generality let us assume that the vectors in $P$ span $\mathbb{R}^d$. We take $x^1, \ldots, x^d$ to be $d$ distinct vectors in $P$ for which $|\det(x^1, \ldots, x^d)|$ is maximized over all choices of $d$ vectors in $P$. The maximality of $|\det(x^1, \ldots, x^d)|$ implies that $P$ is contained in the parallelepiped

$$\{c_1 x^1 + \cdots + c_d x^d : |c_i| \leq 1 \text{ for every } i = 1, \ldots, d\}.$$

Indeed, assume to the contrary that $x \in P$ and $x = c_1 x^1 + \cdots + c_d x^d$ such that, say, $|c_1| > 1$. Then

$$|\det(x, x^2, \ldots, x^d)| = |c_1||\det(x^1, x^2, \ldots, x^d)| > |\det(x^1, x^2, \ldots, x^d)|,$$

contradicting the maximality of $|\det(x^1, \ldots, x^d)|$.

Now, for any vector $u \in \mathbb{R}^d$ and $x \in P$, we have $x = c_1 x^1 + \cdots + c_d x^d$, where $|c_1|, \ldots, |c_d| \leq 1$, from which we conclude that

$$|\langle x, u \rangle| = \left| \sum_{i=1}^{d} c_i \langle x^i, u \rangle \right| \leq \sum_{i=1}^{d} |c_i||\langle x^i, u \rangle| \leq \sum_{i=1}^{d} |\langle x^i, u \rangle|. \qquad \square$$

Note that Lemma 1 is an improved form of what is in the proof of Proposition 1.

*Proof of Theorem* 2. We recall the notion of geometric discrepancy of a family $\mathcal{F}$ of subsets of $\mathbb{R}^d$ with respect to a set of points $P$. We define $\operatorname{disc} \mathcal{F}$ to be the minimum number $t$ such that there exists a coloring of the points of $P$ by two colors red and blue in such a way that for every $F \in \mathcal{F}$ the difference between the number of blue points of $P$ in $F$ and the number of red points of $P$ in $F$ is at most $t$.

We take $\mathcal{F}$ to be the set of all linear half-spaces in $\mathbb{R}^d$, and $P$ is our multiset of $n$ points (vectors) in $\mathbb{R}^d$. Recall that $\mathcal{F}$ has a bounded VC dimension (in fact, $d+1$) that depends only upon $d$. It follows that the $\operatorname{disc} \mathcal{F}$ with respect to $P$ is sublinear; for example, $\operatorname{disc} \mathcal{F} = O(\sqrt{n \log n})$ follows from a random coloring of the points of $P$ (see, for example, [Sp85, Ma99, Ch00]). In fact, in our case, the tight bound is

$\operatorname{disc}\mathcal{F} = \theta(n^{\frac{1}{2}-\frac{1}{2d}})$ [Al90, Ma95], but we will use only the fact that the bound is sublinear.

Let $t$ denote the maximum discrepancy $\operatorname{disc}\mathcal{F}$ of $\mathcal{F}$ with respect to any set of at most $n$ points in $\mathbb{R}^d$, and recall that $t = o(n)$. We now define a subset $C$ of $P$ in the following manner. We will apply Lemma 1 $2t$ times. We start with $C = \emptyset$. We then apply Lemma 1 to obtain $d$ vectors $x^1, \ldots, x^d$ in $P$ satisfying

$$|\langle x, u\rangle| \leq \sum_{i=1}^{d} |\langle x^i, u\rangle|$$

for every $u \in \mathbb{R}^d$ and every $x \in P$. We move these $d$ vectors to $C$, thus decreasing the size of $P$ by $d$. From the resulting subset of $P$ we apply Lemma 1 again to obtain $d$ new vectors from this subset that we move to $C$ and so on, altogether $2t$ times. After we are done we have a set $C$ of cardinality $2dt$ and a remaining subset $P'$ of $P$ of cardinality $n - 2td$. Lemma 1 guarantees that for every $x \in P'$ we have

$$2t|\langle x, u\rangle| \leq \sum_{x \in C} |\langle x, u\rangle| \qquad (2)$$

for all $u \in \mathbb{R}^d$.

Next we find a coloring of the points of $P'$ by two colors red and blue that realizes a geometric discrepancy of at most $t$ with respect to $P'$ and the family $\mathcal{F}$. We denote the set of blue points of $P'$ by $B$ and the set of red points of $P'$ by $A$. Thus $P$ is the disjoint union of $A$, $B$, and $C$. We will show that the set $S = B$ has the desired property (as does $S = A$). First observe that $|P \setminus S| - |S| = |C| + |A| - |B| \leq 2dt + t = o(n)$ and therefore $|S| = (\frac{1}{2} - o(1))n$.

It remains to show that for every vector $u \in \mathbb{R}^d$ we have

$$\sum_{x \in S} |\langle x, u\rangle| \leq \sum_{x \in P\setminus S} |\langle x, u\rangle|.$$

We will use the following technical lemma.

LEMMA 2. *Let $a_1, \ldots, a_k$ and $b_1, \ldots, b_m$ be two sequences of nonnegative numbers all smaller than some number $M$. Assume that for every nonnegative number $L$ we have $\big||\{i : a_i \geq L\}| - |\{j : b_j \geq L\}|\big| \leq r$, where $r$ is some constant integer. Then $|\sum_{i=1}^k a_i - \sum_{j=1}^m b_j| \leq rM$.*

*Proof.* Without loss of generality we may assume that the two sequences of numbers are disjoint. Arrange the elements $a_1, \ldots, a_k, b_1, \ldots, b_m$ in a new decreasing sequence $c_1, \ldots, c_{k+m}$ of size $k + m$. For each $i = 1, \ldots, k + m$ define $s_i = 1$ if $c_i$ is one of $a_1, \ldots, a_k$, and let $s_i = -1$ if $c_i$ is one of $b_1, \ldots, b_m$. We wish to bound $|\sum_{i=1}^{k+m} s_i c_i|$. The assumption of the lemma is equivalent to $|\sum_{i=1}^{\ell} s_i| \leq r$ for every $1 \leq \ell \leq k + m$.

For each $\ell = 1, \ldots, k + m$ let $S_\ell = \sum_{i=1}^{\ell} s_i$. Then we have

$$\begin{aligned}\left|\sum_{i=1}^{k+m} s_i c_i\right| &= \left|S_1 c_1 + \sum_{i=2}^{k+m} (S_i - S_{i-1}) c_i\right| \\ &= |S_1(c_1 - c_2) + S_2(c_2 - c_3) + S_3(c_3 - c_4) + \cdots + S_{n-1}(c_{n-1} - c_n) + S_n c_n| \\ &\leq r(|c_1 - c_2| + |c_2 - c_3| + \cdots + |c_{n-1} - c_n|) + rc_n \\ &= r(c_1 - c_n) + rc_n = rc_1 \leq rM. \qquad \square\end{aligned}$$

Returning to the proof of Theorem 2, let $u$ be any vector in $\mathbb{R}^d$. We first claim that if $M$ denotes the maximum of $|\langle x,u\rangle|$ over all $x \in P'$, then

$$\left|\sum_{x\in A}|\langle x,u\rangle| - \sum_{x\in B}|\langle x,u\rangle|\right| \le 2tM.$$

To prove this let $H$ be the hyperplane orthogonal to $u$. Let $A^+ \subset A$ be the set of all vectors $x \in A$ such that $\langle x,u\rangle \ge 0$ and let $A^- = A \setminus A^+$. Similarly, let $B^+ \subset B$ be the set of all vectors $x \in B$ such that $\langle x,u\rangle \ge 0$ and let $B^- = B \setminus B^+$. The claim will then follow if we can show that $\left|\sum_{x\in A^+}|\langle x,u\rangle| - \sum_{x\in B^+}|\langle x,u\rangle|\right| \le tM$ and $\left|\sum_{x\in A^-}|\langle x,u\rangle| - \sum_{x\in B^-}|\langle x,u\rangle|\right| \le tM$. As the proof of both inequalities is similar, we show only the first. Let $x^1,\dots,x^k$ denote the vectors in $A^+$ and $y^1,\dots,y^m$ the vectors in $B^+$. For every $i=1,\dots,k$, set $a_i = \langle x^i,u\rangle$, and for every $i=1,\dots,m$, set $b_i = \langle y^i,u\rangle$.

The sequences $a_1,\dots,a_k$ and $b_1,\dots,b_m$ satisfy the conditions in Lemma 2 where we take $r=t$ in the lemma. Indeed let $L$ be any nonnegative number. The set of vectors $x \in A^+ \cup B^+$ such that $\langle x,u\rangle \ge L$ is exactly the set of points of $P'$ inside a half-space bounded by a hyperplane parallel to $H$, i.e., the intersection of $P'$ with some set $F \in \mathcal{F}$. By construction we know that the difference between the number of points of $A$ and $B$ in $F$ is at most $t$. We conclude from Lemma 2 that

$$\left|\sum_{x\in A^+}|\langle x,u\rangle| - \sum_{x\in B^+}|\langle x,u\rangle|\right| = \left|\sum_{x\in A^+}\langle x,u\rangle - \sum_{x\in B^+}\langle x,u\rangle\right| \le tM,$$

as required.

In order to complete the proof of Theorem 2, observe that by (2) we have

$$2tM \le \sum_{x\in C}|\langle x,u\rangle|. \tag{3}$$

It now follows from the above result and from (3) that

$$\sum_{x\in B}|\langle x,u\rangle| \le \sum_{x\in A}|\langle x,u\rangle| + \sum_{x\in C}|\langle x,u\rangle|$$

or, in other words,

$$\sum_{x\in S}|\langle x,u\rangle| \le \sum_{x\in P\setminus S}|\langle x,u\rangle| \qquad \square$$

**3. Some consequences.** There are several interesting ways to generalize Theorem 2. We present here a few of them. These generalizations either follow from Theorem 2 or can be obtained by easy modifications of its proof.

We start with the case in which instead of "dominating" the whole set of points we would like to "dominate" it up to a constant factor and hope that the size of the dominating set will change accordingly.

Theorem 3. *Let $0 < c < 1$ be a constant, and let $P$ be a set of $n$ arbitrary vectors in $\mathbb{R}^d$. Then there exists a subset $S$ of $P$ of size at least $(c - o(1))n$ such that for every vector $u \in \mathbb{R}^d$ we have*

$$\sum_{x\in S}|\langle x,u\rangle| \le \frac{c}{1-c}\sum_{x\in P\setminus S}|\langle x,u\rangle|$$

*or, equivalently,*

$$\sum_{x\in S}|\langle x,u\rangle| \le c\sum_{x\in P}|\langle x,u\rangle|.$$

Theorem 2 is a special case of this with $c = \frac{1}{2}$. The proof of Theorem 3 is, up to a modification, essentially identical to that of Theorem 2. This outlines this modification.

Put $g(n) = 10dn^{2/3}(\log n)^{1/3}$. This will serve as the $o(n)$ in Theorem 3. Apply Lemma 1 $2g(n)$ times, and obtain a subset $C$ of $P$ of size $2dg(n)$ and another subset $P' = P \setminus C$. The set $C$ has the property that for every $u \in \mathbb{R}^d$ and every $x \in P'$ we have

$$2g(n)|\langle x,u\rangle| \le \sum_{x\in C}|\langle x,u\rangle|. \tag{4}$$

Next, when coloring the points of $P'$ with the two colors red and blue, we no longer do it so that the standard discrepancy is small. We need a biased coloring of the points of $P'$. We color each point in $P'$ blue with probability $c$, and we color it red otherwise, independently from point to point. Let $A$ denote the set of red points, and let $B$ denote the set of blue points. We need to show that with positive probability every subset $Q \subset P'$ which is the intersection of $P'$ with a half-space satisfies $||Q\cap B| - c|Q|| \le g(n)$. This is very similar to how one shows an $O(\sqrt{n\log m})$ discrepancy for a system of $m$ sets on a ground set of $n$ elements, except that here we use, of course, Hoeffding's inequality rather than Chernoff's inequality.

First note that we need to consider only sets $Q$ for which $|Q| \ge g(n)$. Now, by Hoeffding's inequality [Ho63], for every $\epsilon > 0$ we have

$$\Pr\{\big||Q\cap B| - c|Q|\big| \ge |Q|\epsilon\} \le 2e^{-2|Q|\epsilon^2}.$$

We take $\epsilon = \sqrt[3]{\frac{\log n}{n}}$. Therefore,

$$\begin{aligned}\Pr\{\big||Q\cap B| - c|Q|\big| \ge g(n)\} &\le \Pr\{\big||Q\cap B| - c|Q|\big| \ge n\epsilon\}\\ &\le 2e^{-2|Q|\epsilon^2} \le 2e^{-20d\log n} < \frac{1}{n^d}.\end{aligned}$$

It follows from the uniform bound and from the fact that there are at most $n^d$ possible sets $Q$ (as the VC dimension of half-spaces in $\mathbb{R}^d$ is $d+1$) that with positive probability we obtain the required coloring. It remains to show that the set $S = B$ satisfies the desired property. The remainder of the proof of Theorem 3 continues along the same lines as the proof of Theorem 2. We leave it to the reader to complete the details.

The following is yet another consequence of Theorem 2. In what follows, $||\cdot||_2$ denotes the $\ell_2^d$-norm.

THEOREM 4. *Let $P$ be a set of $n$ vectors in $\mathbb{R}^d$. There exists a subset $S \subset P$ of cardinality at least $(\frac{1}{2} - o(1))n$ such that if $V$ is any linear subspace of $\mathbb{R}^d$ and $T$ is the orthogonal projection from $\mathbb{R}^d$ onto $V$, then we have*

$$\sum_{x\in S}||T(x)||_2 \le \sum_{x\in P\setminus S}||T(x)||_2.$$

*Proof.* The theorem follows by taking a subset $S$ of $P$ as guaranteed by Theorem 2 and applying Theorem 2 for every unit vector $u$ in $V$.

Indeed, let $D$ be the set of unit vectors in $V$. Then for any vector $x$ in $\mathbb{R}^d$ we have

$$\int_D |\langle x, u\rangle| du = c||T(x)||_2,$$

where $c$ is a constant which depends only on $V$ (in fact, $c = \int_D |\langle x_0, u\rangle| du$, where $x_0$ is any fixed vector in $D$). The result now follows immediately from Theorem 2. ∎

More generally, one can try to extend Theorem 2 to a larger family of functions $g : \mathbb{R}^d \to \mathbb{R}$, rather than just to functions of the form $g_u(x) = |\langle x, u\rangle|$ for all vectors $u \in \mathbb{R}^d$. The proof of this next result easily follows along the lines of the proof of Theorem 2, and the set $S$ can be taken to be the same as in Theorem 2.

PROPOSITION 2. *Let $P$ be a multiset of $n$ arbitrary vectors in $\mathbb{R}^d$. Let $g$ be a nonnegative and increasing function on $\mathbb{R}_+$ satisfying $g(0) = 0$ and*

$$g(a + b) \le g(a) + g(b)$$

*for every $a, b \in \mathbb{R}_+$. Then there exists a subset $S$ of $P$ of size at least $(\frac{1}{2} - o(1))n$ such that for every vector $u \in \mathbb{R}^d$ we have*

$$\sum_{x\in S} g(|\langle x, u\rangle|) \le \sum_{x\in P\setminus S} g(|\langle x, u\rangle|).$$

The above condition holds, for example, for every nonnegative increasing concave function $g$. There exist nonnegative increasing functions for which the conclusion of the above proposition is not valid. For example, consider the function $g(t)$ which is 0 for all $t < 1$ and is equal to 1 for all $t \ge 1$. Then consider a set $P$ of $n$ distinct points (vectors) on the unit circle in $\mathbb{R}^2$, no two of which are antipodal. For any choice of the set $S$ take $v \in S$, and let $u = v$. We have $|\langle x, u\rangle| < 1$ for every $x \in P$ different from $v$. We also have $|\langle v, u\rangle| = |\langle v, v\rangle| = 1$. Hence,

$$\sum_{x\in S} g(|\langle x, u\rangle|) = 1 > 0 = \sum_{x\in P\setminus S} g(|\langle x, u\rangle|).$$

**4. The continuous case.** What happens if we look at the continuous model (that could be considered as taking $n$ to $\infty$ in Theorem 1)? Does the same result then hold without any error? Unfortunately not. We first start with a positive result (Theorem 5), obtained very simply by dividing the integral appropriately, where we show that we can always approximate the desired property. However, we also prove (Theorem 6) that if the finite-dimensional subspace contains the quadratic polynomials, then we can never get rid of the arbitrarily small "error" $\epsilon$ that appears below in this next result. On the other hand, for many 2-dimensional subspaces (Proposition 3) we can and do obtain exactness.

THEOREM 5. *Let $W$ be any finite-dimensional subspace of $C[0, 1]$ equipped with the $L_1$-norm. Then for every $\epsilon > 0$ there exists a subset $K \subset [0, 1]$ of measure at least $\frac{1}{2} - \epsilon$ such that for every $f \in W$ we have*

$$\int_K |f| < \int_{[0,1]\setminus K} |f|.$$

*Moreover, $K$ can be taken to be a finite union of closed intervals.*

*Proof.* Let $B = \{f : f \in W, \|f\|_1 \le 1\}$ denote the unit ball in $W$. As $W$ is a finite-dimensional subspace of $C[0,1]$, the set $B$ is uniformly equicontinuous. Let $\epsilon > 0$ be given. There exists an integer $N$ such that for every $x, y \in [0,1]$ satisfying $|x - y| < \frac{1}{N}$ we have $|f(x) - f(y)| < \epsilon$ for all $f \in B$.

Set $M = [0,1]$ and $M_i = [\frac{i}{N}, \frac{i+1}{N}]$. Let $K_i$ be any closed subinterval of $M_i$ of length $\frac{1}{N}(\frac{1}{2} - \epsilon)$. Let $K = \cup_{i=0}^{N-1} K_i$. Clearly $K$ is a union of $N$ closed intervals and has measure $\frac{1}{2} - \epsilon$. It remains to show that for every $f \in W$ we have $\int_K |f| < \int_{M \setminus K} |f|$.

Without loss of generality we assume that $\int_M |f| = 1$. Fix an integer index $i$ between 0 and $N-1$. For every $x, y \in M_i$ we have $|f(x) - f(y)| < \epsilon$ and therefore

$$\left| \frac{1}{|K_i|} \int_{K_i} |f| - \frac{1}{|M_i|} \int_{M_i} |f| \right| < \epsilon,$$

implying that

$$\int_{K_i} |f| < \left(\frac{1}{2} - \epsilon\right) \int_{M_i} |f| + \frac{\epsilon}{2N}.$$

Summing over $i$ and since $\int_M |f| = 1$, we obtain

$$\int_K |f| < \left(\frac{1}{2} - \epsilon\right) \int_M |f| + \frac{\epsilon}{2} = \frac{1}{2} \int_M |f| - \epsilon + \frac{\epsilon}{2} < \frac{1}{2} \int_M |f|.$$

This implies that $\int_K |f| < \int_{M \setminus K} |f|$. ∎

We now prove that it is in general impossible to rid ourselves of the unwanted $\epsilon$.

THEOREM 6. *If the linear subspace $W \subset C[0,1]$ contains the quadratic polynomials, then there is no measurable set $K$ of measure $\frac{1}{2}$ with the property that*

$$\int_K |f| \le \int_{[0,1] \setminus K} |f|$$

*for every $f \in W$.*

*Proof.* It suffices to prove the above for $W = \operatorname{span}\{1, x, x^2\}$. Suppose to the contrary that such a $K$ exists. For convenience let $M = [0,1]$.

We first claim that for every $f \in W$ we necessarily have

$$\int_K f = \int_{M \setminus K} f.$$

To see this take $f \in W$, and let $g_+ = C + f$, where $C$ is any constant such that $g_+ \ge 0$. As $g_+ \in W$, we have

$$\int_K g_+ = \int_K |g_+| \le \int_{M \setminus K} |g_+| = \int_{M \setminus K} g_+.$$

Since $K$ has measure $\frac{1}{2}$, we also have

$$\int_K C = \int_{M \setminus K} C.$$

Thus

$$\int_K f \le \int_{M\setminus K} f.$$

We now set $g_- = C - f$, where $C$ is any constant such that $g_- \ge 0$. Applying the above reasoning, we obtain

$$-\int_K f \le -\int_{M\setminus K} f.$$

Thus

$$\int_K f = \int_{M\setminus K} f.$$

Consider a Lebesgue point of $K$ different from 0 or 1. That is, let $x_0 \in K \cap (0,1)$ be a point such that $\lim_{\delta\to 0^+} \frac{\lambda(K\cap[x_0-\delta,x_0+\delta])}{2\delta} = 1$, where $\lambda$ stands for the standard Lebesgue measure. Let $\delta > 0$ be sufficiently small so that $\frac{\lambda(K\cap[x_0-\delta,x_0+\delta])}{2\delta} \ge 0.9$. Denote by $J$ the interval $[x_0-\delta, x_0+\delta]$.

Let $g \in W$ be of the form $g(x) = -(x-(x_0-\delta))(x-(x_0+\delta))$. Thus $g \ge 0$ for $x \in J$, and $g < 0$ for $x \in M \setminus J$. From the choice of $g$ and $x_0$ and since $K$ captures at least 90% of $J$, we have

$$\begin{aligned}\int_{(M\setminus K)\cap J} g &\le \int_{[-\frac{\delta}{10},\frac{\delta}{10}]} (-x^2+\delta^2)dx < \frac{\delta^3}{5} < \delta^3 \\ &< \int_{[-\delta,-\frac{\delta}{10}]\cup[\frac{\delta}{10},\delta]} (-x^2+\delta^2)dx \le \int_{K\cap J} g.\end{aligned} \tag{5}$$

As $g \ge 0$ for $x \in J$, we therefore have

$$\int_{K\cap J} |g| > \int_{(M\setminus K)\cap J} |g|. \tag{6}$$

Since $\int_K g = \int_{M\setminus K} g$, it follows from (5) that

$$\int_{(M\setminus K)\cap(M\setminus J)} -g < \int_{K\cap(M\setminus J)} -g.$$

As $g < 0$ for $x \in M \setminus J$, we therefore conclude that

$$\int_{K\cap(M\setminus J)} |g| > \int_{(M\setminus K)\cap(M\setminus J)} |g|. \tag{7}$$

Inequalities (6) and (7) imply that

$$\int_K |g| > \int_{(M\setminus K)} |g|,$$

contradicting our assumption. ∎

The converse to Theorem 6 can hold for many 2-dimensional subspaces, as shown by this next result.

PROPOSITION 3. *Let $W = \operatorname{span}\{1, g\}$, where $g$ is any piecewise monotone function in $C[0,1]$. Then there exists a set $K$, which is a finite union of closed intervals, of measure $\frac{1}{2}$ with the property that*

$$\int_K |f| \le \int_{[0,1]\setminus K} |f|$$

*for all $f \in W$.*

*Proof.* It suffices to prove the proposition under the assumption that $g$ is monotone in $M = [0,1]$. (For the general case we simply divide $M$ into the finite subintervals on which $g$ is monotone and apply the result on each subinterval.) From the Hobby–Rice theorem (see, for example, [Pi76]), there exists an interval $K = [a, b]$ such that

$$\int_K f = \int_{M\setminus K} f \tag{8}$$

for all $f \in W$. As the constant function is in $W$, we have $|K| = \frac{1}{2}$. We will prove that

$$\int_K |f| \le \int_{M\setminus K} |f| \tag{9}$$

for all $f \in W$.

Let $g_1 \in W$ satisfy $g_1(a) = 0$, $g_1(x) \ge 0$ for $x \ge a$, $g_1(x) \le 0$ for $x \le a$, and $g_1$ not identically zero. Let $g_2 \in W$ satisfy $g_2(b) = 0$, $g_2(x) \ge 0$ for $x \le b$, $g_2(x) \le 0$ for $x \ge b$, and $g_2$ not identically zero. Both $g_1$ and $g_2$ exist since $g$ is monotone on $[0,1]$.

Consider $f \in W$. If $f$ has no sign change in $K = [a, b]$, then

$$\int_K |f| = \left|\int_K f\right| = \left|\int_{M\setminus K} f\right| \le \int_{M\setminus K} |f|.$$

If $f$ has a sign change in $K$, then, up to sign, we can write $f = g_1 - \lambda g_2$, where $\lambda > 0$. For all such $f$ it is readily verified that

$$|f| = |g_1 - \lambda g_2| = |g_1| + \lambda |g_2|$$

on $M \setminus K$. Thus

$$\begin{aligned}\int_K |f| \le \int_K |g_1| + \lambda|g_2| = \int_K g_1 + \lambda g_2 &= \int_{M\setminus K} g_1 + \lambda g_2 \\ &\le \int_{M\setminus K} |g_1| + \lambda|g_2| = \int_{M\setminus K} |g_1 - \lambda g_2| = \int_{M\setminus K} |f|.\end{aligned}$$

This proves the proposition. $\square$

We do not know what happens if $g$ is not piecewise monotone.

**Acknowledgment.** We would like to thank Sariel Har-Peled for his helpful comments.

REFERENCES

[Al90] R. ALEXANDER, *Geometric methods in the study of irregularities of distribution*, Combinatorica, 10 (1990), pp. 115–136.

[Ch00] B. Chazelle, *The Discrepancy Method: Randomness and Complexity*, Cambridge University Press, Cambridge, UK, 2000.

[Ho63] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), pp. 13–30.

[Ma95] J. Matousek, *Tight upper bounds for the discrepancy of half-spaces*, Discrete Comput. Geom., 13 (1995), pp. 593–601.

[Ma99] J. Matousek, *Geometric Discrepancy: An Illustrated Guide*, Algorithms Combin. 18, Springer-Verlag, Berlin, 1999.

[Pi76] A. Pinkus, *A simple proof of the Hobby-Rice theorem*, Proc. Amer. Math. Soc., 60 (1976), pp. 82–84.

[Pi89] A. Pinkus, *On $L^1$-Approximation*, Cambridge Tracts in Math. 93, Cambridge University Press, Cambridge, UK, 1989.

[Sp85] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc., 289 (1985), pp. 679–706.