On some results in harmonic analysis on the discrete cube

Krzysztof Oleszkiewicz

Institute of Mathematics
University of Warsaw

Geometric Tomography workshop, Banff, 2020

Discrete cube

$[n] := \{1, 2, \ldots, n\}$

Discrete cube (hypercube) $C_n := \{-1, 1\}^n$, equipped with a normalized counting (uniform probability) measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$

Disclaimer: There will be no "cheating" as long as the discrete cube C_n is considered, with $n < \infty$. Many results of the present talk can be extended to the case $n = \infty$ and more general product probability spaces. However, usually technical details become much more delicate then.

Hamming's metric: For $x, y \in C_n$ let

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2}\|x - y\|_1.$$

Expectation: For $f : C_n \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}[f] = 2^{-n} \sum_{x \in C_n} f(x).$$

Discrete cube

$[n] := \{1, 2, \ldots, n\}$

Discrete cube (hypercube) $C_n := \{-1, 1\}^n$, equipped with a normalized counting (uniform probability) measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$

Disclaimer: There will be no "cheating" as long as the discrete cube C_n is considered, with $n < \infty$. Many results of the present talk can be extended to the case $n = \infty$ and more general product probability spaces. However, usually technical details become much more delicate then.

Hamming's metric: For $x, y \in C_n$ let

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2}\|x - y\|_1.$$

Expectation: For $f : C_n \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}[f] = 2^{-n} \sum_{x \in C_n} f(x).$$

Discrete cube

$[n] := \{1, 2, \ldots, n\}$

Discrete cube (hypercube) $C_n := \{-1, 1\}^n$, equipped with a normalized counting (uniform probability) measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$

Disclaimer: There will be no "cheating" as long as the discrete cube C_n is considered, with $n < \infty$. Many results of the present talk can be extended to the case $n = \infty$ and more general product probability spaces. However, usually technical details become much more delicate then.

Hamming's metric: For $x, y \in C_n$ let

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2}\|x - y\|_1.$$

Expectation: For $f : C_n \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}[f] = 2^{-n} \sum_{x \in C_n} f(x).$$

Discrete cube

$[n] := \{1, 2, \ldots, n\}$

Discrete cube (hypercube) $C_n := \{-1, 1\}^n$, equipped with a normalized counting (uniform probability) measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$

Disclaimer: There will be no "cheating" as long as the discrete cube C_n is considered, with $n < \infty$. Many results of the present talk can be extended to the case $n = \infty$ and more general product probability spaces. However, usually technical details become much more delicate then.

Hamming's metric: For $x, y \in C_n$ let

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2}\|x - y\|_1.$$

Expectation: For $f : C_n \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}[f] = 2^{-n} \sum_{x \in C_n} f(x).$$

Discrete cube

$[n] := \{1, 2, \ldots, n\}$

Discrete cube (hypercube) $C_n := \{-1, 1\}^n$, equipped with a normalized counting (uniform probability) measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$

Disclaimer: There will be no "cheating" as long as the discrete cube C_n is considered, with $n < \infty$. Many results of the present talk can be extended to the case $n = \infty$ and more general product probability spaces. However, usually technical details become much more delicate then.

Hamming's metric: For $x, y \in C_n$ let

$$d(x, y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2}\|x - y\|_1.$$

Expectation: For $f : C_n \longrightarrow \mathbb{R}$ we have

$$\mathbb{E}[f] = 2^{-n} \sum_{x \in C_n} f(x).$$

L^2 structure

Scalar product: For $f, g : C_n \longrightarrow \mathbb{R}$ let

$$\langle f, g \rangle = \mathbb{E}[f \cdot g] = 2^{-n} \cdot \sum_{x \in C_n} f(x)g(x).$$

We denote $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p > 0$ and $\|f\|_\infty = \max_{x \in C_n} |f(x)|$.

Note that $\langle f, f \rangle = \|f\|_2^2$.

Hilbert space:

$$\mathcal{H}_n := L^2(C_n, \mathbb{R}); \quad \dim \mathcal{H}_n = 2^n$$

L^2 structure

Scalar product: For $f, g : C_n \longrightarrow \mathbb{R}$ let

$$\langle f, g \rangle = \mathbb{E}[f \cdot g] = 2^{-n} \cdot \sum_{x \in C_n} f(x)g(x).$$

We denote $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p > 0$ and $\|f\|_\infty = \max_{x \in C_n} |f(x)|$.

Note that $\langle f, f \rangle = \|f\|_2^2$.

Hilbert space:

$$\mathcal{H}_n := L^2(C_n, \mathbb{R}); \quad \dim \mathcal{H}_n = 2^n$$

L^2 structure

Scalar product: For $f, g : C_n \longrightarrow \mathbb{R}$ let

$$\langle f, g \rangle = \mathbb{E}[f \cdot g] = 2^{-n} \cdot \sum_{x \in C_n} f(x)g(x).$$

We denote $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p > 0$ and $\|f\|_\infty = \max_{x \in C_n} |f(x)|$.

Note that $\langle f, f \rangle = \|f\|_2^2$.

Hilbert space:

$$\mathcal{H}_n := L^2(C_n, \mathbb{R}); \quad \dim \mathcal{H}_n = 2^n$$

Walsh system

Boolean function: $f : C_n \to \{-1, 1\}$

Motivation:

- theoretical computer science
- social choice theory

Walsh functions: For $x \in \{-1, 1\}^n$ and $S \subseteq [n]$ let

$$w_S(x) = \prod_{i \in S} x_i,$$

$$w_\emptyset \equiv 1$$

$r_i := w_i = w_{\{i\}}$ - i-th coordinate projection π_i ($i \in [n]$)

$r_1, r_2, \ldots, r_n$ - a Rademacher sequence:
independent symmetric ± 1 Bernoulli random variables

Walsh system

Boolean function: $f : C_n \to \{-1, 1\}$

Motivation:

- theoretical computer science
- social choice theory

Walsh functions: For $x \in \{-1, 1\}^n$ and $S \subseteq [n]$ let

$$w_S(x) = \prod_{i \in S} x_i,$$

$$w_\emptyset \equiv 1$$

$r_i := w_i = w_{\{i\}}$ - i-th coordinate projection π_i $(i \in [n])$

$r_1, r_2, \ldots, r_n$ - a Rademacher sequence:
independent symmetric ± 1 Bernoulli random variables

Walsh system

Boolean function: $f : C_n \to \{-1, 1\}$

Motivation:

- theoretical computer science
- social choice theory

Walsh functions: For $x \in \{-1, 1\}^n$ and $S \subseteq [n]$ let

$$w_S(x) = \prod_{i \in S} x_i,$$

$$w_\emptyset \equiv 1$$

$r_i := w_i = w_{\{i\}}$ - i-th coordinate projection π_i $(i \in [n])$

$r_1, r_2, \ldots, r_n$ - a Rademacher sequence:
independent symmetric ± 1 Bernoulli random variables

Walsh system

Boolean function: $f : C_n \to \{-1, 1\}$

Motivation:

- theoretical computer science
- social choice theory

Walsh functions: For $x \in \{-1, 1\}^n$ and $S \subseteq [n]$ let

$$w_S(x) = \prod_{i \in S} x_i,$$

$$w_\emptyset \equiv 1$$

$r_i := w_i = w_{\{i\}}$ - i-th coordinate projection π_i $(i \in [n])$

$r_1, r_2, \ldots, r_n$ - a Rademacher sequence:
independent symmetric ± 1 Bernoulli random variables

Walsh system

Boolean function: $f : C_n \to \{-1, 1\}$

Motivation:

- theoretical computer science
- social choice theory

Walsh functions: For $x \in \{-1, 1\}^n$ and $S \subseteq [n]$ let

$$w_S(x) = \prod_{i \in S} x_i,$$

$$w_\emptyset \equiv 1$$

$r_i := w_i = w_{\{i\}}$ - i-th coordinate projection π_i ($i \in [n]$)

$r_1, r_2, \ldots, r_n$ - a Rademacher sequence:
independent symmetric ± 1 Bernoulli random variables

Orthonormality

$\mathbb{E}[w_S] = 0$ for $S \neq \emptyset$ and $\mathbb{E}[w_\emptyset] = 1$

Indeed, expectation of the product of independent random variables is equal to the product of their expectations (and they are all equal to zero).

Orthonormality: $w_S \cdot w_T = w_{S \Delta T}$ thus

$$\langle w_S, w_T \rangle = \mathbb{E}[w_{S \Delta T}] = \delta_{S,T}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{S,T} = 1$ if $S = T$ and $\delta_{S,T} = 0$ if $S \neq T$ (Kronecker's delta).

Example: $w_{\{1,2\}} \cdot w_{\{2,3\}} = r_1 r_2 \cdot r_2 r_3 = r_1 r_2^2 r_3 = r_1 r_3$.

We have proved that the Walsh system $(w_S)_{S \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of $\mathcal{H}_n$, it spans the whole space and thus is complete.

Orthonormality

$\mathbb{E}[w_S] = 0$ for $S \neq \emptyset$ and $\mathbb{E}[w_\emptyset] = 1$

Indeed, expectation of the product of independent random variables is equal to the product of their expectations (and they are all equal to zero).

Orthonormality: $w_S \cdot w_T = w_{S \Delta T}$ thus

$$\langle w_S, w_T \rangle = \mathbb{E}[w_{S \Delta T}] = \delta_{S,T}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{S,T} = 1$ if $S = T$ and $\delta_{S,T} = 0$ if $S \neq T$ (Kronecker's delta).

Example: $w_{\{1,2\}} \cdot w_{\{2,3\}} = r_1 r_2 \cdot r_2 r_3 = r_1 r_2^2 r_3 = r_1 r_3.$

We have proved that the Walsh system $(w_S)_{S \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of $\mathcal{H}_n$, it spans the whole space and thus is complete.

Orthonormality

$\mathbb{E}[w_S] = 0$ for $S \neq \emptyset$ and $\mathbb{E}[w_\emptyset] = 1$

Indeed, expectation of the product of independent random variables is equal to the product of their expectations (and they are all equal to zero).

Orthonormality: $w_S \cdot w_T = w_{S \Delta T}$ thus

$$\langle w_S, w_T \rangle = \mathbb{E}[w_{S \Delta T}] = \delta_{S,T}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{S,T} = 1$ if $S = T$ and $\delta_{S,T} = 0$ if $S \neq T$ (Kronecker's delta).

Example: $w_{\{1,2\}} \cdot w_{\{2,3\}} = r_1 r_2 \cdot r_2 r_3 = r_1 r_2^2 r_3 = r_1 r_3$.

We have proved that the Walsh system $(w_S)_{S \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of $\mathcal{H}_n$, it spans the whole space and thus is complete.

Orthonormality

$\mathbb{E}[w_S] = 0$ for $S \neq \emptyset$ and $\mathbb{E}[w_\emptyset] = 1$

Indeed, expectation of the product of independent random variables is equal to the product of their expectations (and they are all equal to zero).

Orthonormality: $w_S \cdot w_T = w_{S \Delta T}$ thus

$$\langle w_S, w_T \rangle = \mathbb{E}[w_{S \Delta T}] = \delta_{S,T}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{S,T} = 1$ if $S = T$ and $\delta_{S,T} = 0$ if $S \neq T$ (Kronecker's delta).

Example: $w_{\{1,2\}} \cdot w_{\{2,3\}} = r_1 r_2 \cdot r_2 r_3 = r_1 r_2^2 r_3 = r_1 r_3.$

We have proved that the Walsh system $(w_S)_{S \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of $\mathcal{H}_n$, it spans the whole space and thus is complete.

There is also a straightforward way to see that every function from $\mathcal{H}_n$ is a linear combination of the Walsh functions. Indeed, for any $y \in C_n$ we have

$$1_y(x) = \prod_{i=1}^{n} \frac{1 + x_i y_i}{2} = 2^{-n} \sum_{S \subseteq [n]} w_S(y) w_S(x),$$

where 1_y denotes the indicator (the characteristic function) of $\{y\}$. Hence

$$f(x) = \sum_{y \in C_n} f(y) 1_y(x) = 2^{-n} \sum_{S \subseteq [n]} \Big(\sum_{y \in C_n} f(y) w_S(y) \Big) w_S(x) =$$

$$= \sum_{S \subseteq [n]} \langle f, w_S \rangle \cdot w_S(x).$$

Therefore every $f \in \mathcal{H}_n$ admits one and only one **Walsh-Fourier expansion:**

$$f = \sum \hat{f}(S) w_S.$$

Elementary argument

There is also a straightforward way to see that every function from $\mathcal{H}_n$ is a linear combination of the Walsh functions. Indeed, for any $y \in C_n$ we have

$$1_y(x) = \prod_{i=1}^{n} \frac{1 + x_i y_i}{2} = 2^{-n} \sum_{S \subseteq [n]} w_S(y) w_S(x),$$

where 1_y denotes the indicator (the characteristic function) of $\{y\}$. Hence

$$f(x) = \sum_{y \in C_n} f(y) 1_y(x) = 2^{-n} \sum_{S \subseteq [n]} \Big(\sum_{y \in C_n} f(y) w_S(y) \Big) w_S(x) =$$

$$= \sum_{S \subseteq [n]} \langle f, w_S \rangle \cdot w_S(x).$$

Therefore every $f \in \mathcal{H}_n$ admits one and only one **Walsh-Fourier expansion:**

$$f = \sum \hat{f}(S) w_S.$$

Simple consequences of the orthonormality

As we have seen above (it follows also from the orthonormality of the Walsh system):

$$\hat{f}(S) = \langle f, w_S \rangle = \mathbb{E}[f \cdot w_S].$$

In particular, for every $f \in \mathcal{H}_n$ we have

$$\mathbb{E}[f] = \mathbb{E}[f \cdot \mathbf{1}] = \mathbb{E}[f \cdot w_\emptyset] = \langle f, w_\emptyset \rangle = \hat{f}(\emptyset)$$

and

$$\mathbb{E}[f^2] = \mathbb{E}[f \cdot f] = \langle f, f \rangle = \langle \sum_{S \subseteq [n]} \hat{f}(S) w_S, \sum_{T \subseteq [n]} \hat{f}(T) w_T \rangle =$$

$$= \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T)\langle w_S, w_T \rangle = \sum_{S \subseteq [n]} \hat{f}(S)^2 \text{ (Plancherel)}.$$

Simple consequences of the orthonormality

As we have seen above (it follows also from the orthonormality of the Walsh system):

$$\hat{f}(S) = \langle f, w_S \rangle = \mathbb{E}[f \cdot w_S].$$

In particular, for every $f \in \mathcal{H}_n$ we have

$$\mathbb{E}[f] = \mathbb{E}[f \cdot \mathbf{1}] = \mathbb{E}[f \cdot w_\emptyset] = \langle f, w_\emptyset \rangle = \hat{f}(\emptyset)$$

and

$$\mathbb{E}[f^2] = \mathbb{E}[f \cdot f] = \langle f, f \rangle = \langle \sum_{S \subseteq [n]} \hat{f}(S) w_S, \sum_{T \subseteq [n]} \hat{f}(T) w_T \rangle =$$

$$= \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T)\langle w_S, w_T \rangle = \sum_{S \subseteq [n]} \hat{f}(S)^2 \ \textbf{(Plancherel)}.$$

Frequencies

Let $S \subseteq [n]$. If the cardinality of S is

small, then we deal with a low frequency.

If it is large, then we deal with a high frequency.

Clear analogy to the trigonometric system terminology.

Frequencies

Let $S \subseteq [n]$. If the cardinality of S is

small, then we deal with a low frequency.

If it is large, then we deal with a high frequency.

Clear analogy to the trigonometric system terminology.

Frequencies

Let $S \subseteq [n]$. If the cardinality of S is

small, then we deal with a low frequency.

If it is large, then we deal with a high frequency.

Clear analogy to the trigonometric system terminology.

Bounded degree chaoses and tail spaces

We will call $f : \{-1,1\} \to \mathbb{R}$ a Rademacher chaos of degree not exceeding d, if $\hat{f}(S) = 0$ for all $S \subseteq [n]$ with $|S| > d$.

Rademacher chaoses of degree not exceeding d form a linear subspace of $\mathcal{H}_n$.

The linear subspace of $\mathcal{H}_n$ spanned by $(w_S)_{|S| \geq k}$ is usually denoted by $T_{\geq k}$ and called the k-th tail space.

Bounded degree chaoses and tail spaces

We will call $f : \{-1,1\} \to \mathbb{R}$ a Rademacher chaos of degree not exceeding d, if $\hat{f}(S) = 0$ for all $S \subseteq [n]$ with $|S| > d$.

Rademacher chaoses of degree not exceeding d form a linear subspace of $\mathcal{H}_n$.

The linear subspace of $\mathcal{H}_n$ spanned by $(w_S)_{|S| \geq k}$ is usually denoted by $T_{\geq k}$ and called the k-th tail space.

Bounded degree chaoses and tail spaces

We will call $f : \{-1,1\} \to \mathbb{R}$ a Rademacher chaos of degree not exceeding d, if $\hat{f}(S) = 0$ for all $S \subseteq [n]$ with $|S| > d$.

Rademacher chaoses of degree not exceeding d form a linear subspace of $\mathcal{H}_n$.

The linear subspace of $\mathcal{H}_n$ spanned by $(w_S)_{|S| \geq k}$ is usually denoted by $T_{\geq k}$ and called the k-th tail space.

LCA setting

Remark: Note that $\{-1,1\}$ (with multiplication as a group action) is a locally compact (compact, in fact) abelian group and $C_n = \{-1,1\}^n$ (with coordinatewise multiplication as a group action) shares this property. The case of the Cantor group ($n = \infty$ with the natural product topology) is covered as well. The standard product probability measure on C_n is the Haar measure then and general harmonic analysis on LCA groups tools apply. It is easy to check that, for $n < \infty$, C_n is self-dual: the group of characters on C_n is just the Walsh system and it is isomorphic with C_n itself and the isomorphism is very natural - $S \subseteq [n]$ is identified with $x \in C_n$ such that $S = \{i \in [n] : x_i = -1\}$. Then the mapping $f \mapsto \hat{f}$, which sends a real function on C_n to its Walsh-Fourier coefficients collection, is just the classical Fourier transform (on LCA groups) up to some normalization. The transform applied twice returns the original function, up to a multiplicative factor. However, in what follows we will not take advantage (at least explicitely) of the group structure of C_n.

Jacek Jendrej, K. O., Jakub Onufry Wojtaszczyk
(University of Warsaw, back in 2013...)

Some extensions of the FKN theorem

Theory of Computing 11 (2015), 445–469

The FKN result

Theorem

There exists a universal constant $L > 0$ *with the following property. For* $f : \{-1,1\}^n \to \{-1,1\}$ *let* $\rho = \left(\sum_{A \subseteq [n]:\, |A| \geq 2} |\hat{f}(A)|^2\right)^{1/2}$.

Then there exists some $B \subseteq [n]$ *with* $|B| \leq 1$ *such that*

$$\sum_{A \subseteq [n]:\, A \neq B} |\hat{f}(A)|^2 \leq L \cdot \rho^2,$$

$$|\hat{f}(B)|^2 \geq 1 - L \cdot \rho^2.$$

The $O(\rho^2)$ bound of Friedgut, Kalai, and Naor (2002) was a bit later strengthened to $2\rho^2 + o(\rho^2)$ by Kindler and Safra.

The FKN result

Theorem

There exists a universal constant $L > 0$ with the following property. For $f : \{-1,1\}^n \to \{-1,1\}$ let $\rho = \left(\sum_{A \subseteq [n]:\, |A| \geq 2} |\hat{f}(A)|^2\right)^{1/2}$.

Then there exists some $B \subseteq [n]$ with $|B| \leq 1$ such that

$$\sum_{A \subseteq [n]:\, A \neq B} |\hat{f}(A)|^2 \leq L \cdot \rho^2,$$

$$|\hat{f}(B)|^2 \geq 1 - L \cdot \rho^2.$$

The $O(\rho^2)$ bound of Friedgut, Kalai, and Naor (2002) was a bit later strengthened to $2\rho^2 + o(\rho^2)$ by Kindler and Safra.

Main idea of FKN

Friedgut, Kalai and Naor have shown that if the variance of the absolute value of a sum of weighted Rademacher variables is much smaller than the variance of the sum, then one of the summands dominates the sum.

Geometric digression

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$, $A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

Geometric digression

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$, $A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

Geometric digression

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$, $A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$, $A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

Geometric digression

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$, $A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

Geometric digression

Let $(F, \|\cdot\|)$ be a normed linear space. We will say that $A \subset F$ is 1-separated if for any distinct $x, y \in A$ there is $\|x - y\| \geq 1$.

Question: Let A and B be 1-separated finite non-empty subsets of F. Does their Minkowski sum $A + B$ necessarily contain some 1-separated subset of cardinality $|A| + |B| - 1$?

Example: $F = \mathbb{R}$, $A = \{1, 2, \ldots, a\}$, $B = \{1, 2, \ldots, b\}$,
$A + B = \{2, 3, \ldots, a + b\}$.

Yes, if $|A| \leq 2$ or $|B| \leq 2$ (easy).

Yes, if $(F, \|\cdot\|)$ is Euclidean.

What if $|A| = |B| = 3$?

An improvement of the FKN theorem

Theorem

There exists a universal constant $L > 0$ with the following property. For $f : \{-1,1\}^n \to \{-1,1\}$ let $\rho = \left(\sum_{A\subseteq[n]:\,|A|\geq 2} |\hat{f}(A)|^2\right)^{1/2}$.

Then there exists some $B \subseteq [n]$ with $|B| \leq 1$ such that

$$\sum_{A\subseteq[n]:\,|A|\leq 1, A\neq B} |\hat{f}(A)|^2 \leq L \cdot \rho^4 \ln(2/\rho),$$

$$|\hat{f}(B)|^2 \geq 1 - \rho^2 - L \cdot \rho^4 \ln(2/\rho).$$

The bound $O\left(\rho^4 \ln(2/\rho)\right)$ is of the optimal order (and was independently proved by O'Donnell). For any $2 \leq m \leq n$ consider just

$$f(x) = 1 - \frac{1}{2^{m-1}} \prod_{i=1}^{m} (1 + x_i).$$

An improvement of the FKN theorem

Theorem

There exists a universal constant $L > 0$ with the following property. For $f : \{-1,1\}^n \to \{-1,1\}$ let $\rho = \left(\sum_{A\subseteq[n]:\,|A|\geq 2} |\hat{f}(A)|^2\right)^{1/2}$.

Then there exists some $B \subseteq [n]$ with $|B| \leq 1$ such that

$$\sum_{A\subseteq[n]:\,|A|\leq 1, A\neq B} |\hat{f}(A)|^2 \leq L \cdot \rho^4 \ln(2/\rho),$$

$$|\hat{f}(B)|^2 \geq 1 - \rho^2 - L \cdot \rho^4 \ln(2/\rho).$$

The bound $O\Big(\rho^4 \ln(2/\rho)\Big)$ is of the optimal order (and was independently proved by O'Donnell). For any $2 \leq m \leq n$ consider just

$$f(x) = 1 - \frac{1}{2^{m-1}} \prod_{i=1}^{m} (1 + x_i).$$

Assumptions and Notation (A & N)

$\xi_1, \xi_2, \ldots, \xi_n$ – independent symmetric ± 1 random variables, $\mathbb{E}\xi_i = 0$, $\mathbb{E}\xi_i^2 = 1$.

Hilbert space $L^2 = L^2(\{-1,1\}^n, \mu)$, where

$$\mu = \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)^{\otimes n}$$

is the distribution of the vector $(\xi_1, \xi_2, \ldots, \xi_n)$.

Let $\xi_0 \equiv 1$.

Let $\mathcal{A}$ be a linear (finite dimensional and thus closed) subspace of L^2 consisting of (restrictions to the discrete cube of) all affine real-valued functions.

Assumptions and Notation (A & N)

$\xi_1, \xi_2, \ldots, \xi_n$ – independent symmetric ± 1 random variables, $\mathbb{E}\xi_i = 0$, $\mathbb{E}\xi_i^2 = 1$.

Hilbert space $L^2 = L^2(\{-1,1\}^n, \mu)$, where

$$\mu = \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)^{\otimes n}$$

is the distribution of the vector $(\xi_1, \xi_2, \ldots, \xi_n)$.

Let $\xi_0 \equiv 1$.

Let $\mathcal{A}$ be a linear (finite dimensional and thus closed) subspace of L^2 consisting of (restrictions to the discrete cube of) all affine real-valued functions.

Assumptions and Notation (A & N)

$\xi_1, \xi_2, \ldots, \xi_n$ – independent symmetric ± 1 random variables, $\mathbb{E}\xi_i = 0$, $\mathbb{E}\xi_i^2 = 1$.

Hilbert space $L^2 = L^2(\{-1,1\}^n, \mu)$, where

$$\mu = \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)^{\otimes n}$$

is the distribution of the vector $(\xi_1, \xi_2, \ldots, \xi_n)$.

Let $\xi_0 \equiv 1$.

Let $\mathcal{A}$ be a linear (finite dimensional and thus closed) subspace of L^2 consisting of (restrictions to the discrete cube of) all affine real-valued functions.

Assumptions and Notation (A & N)

$\xi_1, \xi_2, \ldots, \xi_n$ – independent symmetric ± 1 random variables, $\mathbb{E}\xi_i = 0$, $\mathbb{E}\xi_i^2 = 1$.

Hilbert space $L^2 = L^2(\{-1, 1\}^n, \mu)$, where

$$\mu = \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1 \right)^{\otimes n}$$

is the distribution of the vector $(\xi_1, \xi_2, \ldots, \xi_n)$.

Let $\xi_0 \equiv 1$.

Let $\mathcal{A}$ be a linear (finite dimensional and thus closed) subspace of L^2 consisting of (restrictions to the discrete cube of) all affine real-valued functions.

A & N - continued

We define coordinate projection functions $\pi_i : \{-1,1\}^n \longrightarrow \mathbb{R}$ by $\pi_i(x) = x_i$ for $1 \le i \le n$, and $\pi_0 \equiv 1$ (orthonormal system). Let $\mathcal{A}_\pi = \{\pi_0, -\pi_0, \pi_1, -\pi_1, \ldots, \pi_n, -\pi_n\}$.

For a Boolean (i.e. $\{-1,1\}$-valued) function f on $\{-1,1\}^n$ by $f_{\mathcal{A}}$ we will denote its orthogonal projection in L^2 onto $\mathcal{A}$: $f_{\mathcal{A}}(x) = a_0 + a_1x_1 + \ldots + a_nx_n$, i.e. $f_{\mathcal{A}} = \sum_{i=0}^n a_i\pi_i$.

$$\rho := \mathrm{dist}_{L^2}(f, \mathcal{A}), \quad d := \mathrm{dist}_{L^2}(f, \mathcal{A}_\pi)$$

Easy: if f is Boolean, then $\rho \le \|f - 0\|_{L^2} = 1$ and $d \le \sqrt{2}$ (L^2-distance between two Boolean functions cannot exceed $\sqrt{2}$).

Obviously, $\rho \le d$ (since $\mathcal{A}_\pi \subset \mathcal{A}$).

A & N - continued

We define coordinate projection functions $\pi_i : \{-1,1\}^n \longrightarrow \mathbb{R}$ by $\pi_i(x) = x_i$ for $1 \le i \le n$, and $\pi_0 \equiv 1$ (orthonormal system). Let $\mathcal{A}_\pi = \{\pi_0, -\pi_0, \pi_1, -\pi_1, \ldots, \pi_n, -\pi_n\}$.

For a Boolean (i.e. $\{-1,1\}$-valued) function f on $\{-1,1\}^n$ by $f_\mathcal{A}$ we will denote its orthogonal projection in L^2 onto $\mathcal{A}$: $f_\mathcal{A}(x) = a_0 + a_1 x_1 + \ldots + a_n x_n$, i.e. $f_\mathcal{A} = \sum_{i=0}^n a_i \pi_i$.

$$\rho := \mathrm{dist}_{L^2}(f, \mathcal{A}), \quad d := \mathrm{dist}_{L^2}(f, \mathcal{A}_\pi)$$

Easy: if f is Boolean, then $\rho \le \|f - 0\|_{L^2} = 1$ and $d \le \sqrt{2}$ (L^2-distance between two Boolean functions cannot exceed $\sqrt{2}$).

Obviously, $\rho \le d$ (since $\mathcal{A}_\pi \subset \mathcal{A}$).

A & N - continued

We define coordinate projection functions $\pi_i : \{-1,1\}^n \longrightarrow \mathbb{R}$ by $\pi_i(x) = x_i$ for $1 \le i \le n$, and $\pi_0 \equiv 1$ (orthonormal system). Let $\mathcal{A}_\pi = \{\pi_0, -\pi_0, \pi_1, -\pi_1, \ldots, \pi_n, -\pi_n\}$.

For a Boolean (i.e. $\{-1,1\}$-valued) function f on $\{-1,1\}^n$ by $f_\mathcal{A}$ we will denote its orthogonal projection in L^2 onto $\mathcal{A}$: $f_\mathcal{A}(x) = a_0 + a_1 x_1 + \ldots + a_n x_n$, i.e. $f_\mathcal{A} = \sum_{i=0}^n a_i \pi_i$.

$$\rho := \mathrm{dist}_{L^2}(f, \mathcal{A}), \quad d := \mathrm{dist}_{L^2}(f, \mathcal{A}_\pi)$$

Easy: if f is Boolean, then $\rho \le \|f - 0\|_{L^2} = 1$ and $d \le \sqrt{2}$ (L^2-distance between two Boolean functions cannot exceed $\sqrt{2}$).

Obviously, $\rho \le d$ (since $\mathcal{A}_\pi \subset \mathcal{A}$).

A & N - continued

We define coordinate projection functions $\pi_i : \{-1,1\}^n \longrightarrow \mathbb{R}$ by $\pi_i(x) = x_i$ for $1 \le i \le n$, and $\pi_0 \equiv 1$ (orthonormal system). Let $\mathcal{A}_\pi = \{\pi_0, -\pi_0, \pi_1, -\pi_1, \ldots, \pi_n, -\pi_n\}$.

For a Boolean (i.e. $\{-1,1\}$-valued) function f on $\{-1,1\}^n$ by $f_\mathcal{A}$ we will denote its orthogonal projection in L^2 onto $\mathcal{A}$: $f_\mathcal{A}(x) = a_0 + a_1x_1 + \ldots + a_nx_n$, i.e. $f_\mathcal{A} = \sum_{i=0}^n a_i\pi_i$.

$$\rho := \mathrm{dist}_{L^2}(f, \mathcal{A}), \quad d := \mathrm{dist}_{L^2}(f, \mathcal{A}_\pi)$$

Easy: if f is Boolean, then $\rho \le \|f - 0\|_{L^2} = 1$ and $d \le \sqrt{2}$ (L^2-distance between two Boolean functions cannot exceed $\sqrt{2}$).

Obviously, $\rho \le d$ (since $\mathcal{A}_\pi \subset \mathcal{A}$).

A & N - continued

We define coordinate projection functions $\pi_i : \{-1,1\}^n \longrightarrow \mathbb{R}$ by $\pi_i(x) = x_i$ for $1 \le i \le n$, and $\pi_0 \equiv 1$ (orthonormal system). Let $\mathcal{A}_\pi = \{\pi_0, -\pi_0, \pi_1, -\pi_1, \ldots, \pi_n, -\pi_n\}$.

For a Boolean (i.e. $\{-1,1\}$-valued) function f on $\{-1,1\}^n$ by $f_{\mathcal{A}}$ we will denote its orthogonal projection in L^2 onto $\mathcal{A}$: $f_{\mathcal{A}}(x) = a_0 + a_1x_1 + \ldots + a_nx_n$, i.e. $f_{\mathcal{A}} = \sum_{i=0}^n a_i\pi_i$.

$$\rho := \mathrm{dist}_{L^2}(f, \mathcal{A}), \quad d := \mathrm{dist}_{L^2}(f, \mathcal{A}_\pi)$$

Easy: if f is Boolean, then $\rho \le \|f - 0\|_{L^2} = 1$ and $d \le \sqrt{2}$ (L^2-distance between two Boolean functions cannot exceed $\sqrt{2}$).

Obviously, $\rho \le d$ (since $\mathcal{A}_\pi \subset \mathcal{A}$).

Discrete cube

Now let us see how to strengthen the result of Friedgut, Kalai, and Naor. For a function f defined on the discrete cube $\{-1,1\}^n$ we consider its standard Walsh-Fourier expansion $\sum_A \hat{f}(A) w_A$, where $w_A(x) = \prod_{i \in A} x_i$.

Theorem

There exists a universal constant $L > 0$ with the following property. For $f : \{-1,1\}^n \to \{-1,1\}$ let $\rho = \left(\sum_{A \subseteq [n]:\, |A| \geq 2} |\hat{f}(A)|^2 \right)^{1/2}$.

Then there exists some $B \subseteq [n]$ with $|B| \leq 1$ such that

$$\sum_{A \subseteq [n]:\, |A| \leq 1,\, A \neq B} |\hat{f}(A)|^2 \leq L \cdot \rho^4 \ln(2/\rho),$$

$$|\hat{f}(B)|^2 \geq 1 - \rho^2 - L \cdot \rho^4 \ln(2/\rho).$$

Proof of the discrete cube result - auxiliary notation

Proof:
Let $a_i = \langle f, \pi_i \rangle_{L^2} = \hat{f}(\{i\})$ for $i \in [n]$, and $a_0 = \hat{f}(\emptyset)$.

Let $\theta = \left(4 \log_2(2/d) - 1\right)^{-1}$. There is $\theta \in (0,1]$ because $d \leq \sqrt{2}$.

Let $k \in \{0, 1, \dots, n\}$ be such that $d = \|f - \pi_k\|_{L^2}$ (if the point of $\mathcal{A}_\pi$ closest to f is of the form $-\pi_k$ then a similar reasoning works).

Hence $d^2 = \|f\|_{L^2}^2 + \|\pi_k\|_{L^2}^2 - 2\langle f, \pi_k \rangle_{L^2} = 2(1 - a_k)$.
Remember:

$$(1 - a_k)^2 = d^4/4.$$

Proof of the discrete cube result - auxiliary notation

Proof:
Let $a_i = \langle f, \pi_i \rangle_{L^2} = \hat{f}(\{i\})$ for $i \in [n]$, and $a_0 = \hat{f}(\emptyset)$.

Let $\theta = \left(4\log_2(2/d) - 1\right)^{-1}$. There is $\theta \in (0,1]$ because $d \leq \sqrt{2}$.

Let $k \in \{0, 1, \ldots, n\}$ be such that $d = \|f - \pi_k\|_{L^2}$ (if the point of $\mathcal{A}_\pi$ closest to f is of the form $-\pi_k$ then a similar reasoning works).

Hence $d^2 = \|f\|_{L^2}^2 + \|\pi_k\|_{L^2}^2 - 2\langle f, \pi_k \rangle_{L^2} = 2(1 - a_k)$.
Remember:

$$(1 - a_k)^2 = d^4/4.$$

Proof of the discrete cube result - auxiliary notation

Proof:
Let $a_i = \langle f, \pi_i \rangle_{L^2} = \hat{f}(\{i\})$ for $i \in [n]$, and $a_0 = \hat{f}(\emptyset)$.

Let $\theta = \left(4\log_2(2/d) - 1\right)^{-1}$. There is $\theta \in (0,1]$ because $d \le \sqrt{2}$.

Let $k \in \{0, 1, \ldots, n\}$ be such that $d = \|f - \pi_k\|_{L^2}$ (if the point of $\mathcal{A}_\pi$ closest to f is of the form $-\pi_k$ then a similar reasoning works).

Hence $d^2 = \|f\|_{L^2}^2 + \|\pi_k\|_{L^2}^2 - 2\langle f, \pi_k \rangle_{L^2} = 2(1 - a_k)$.
Remember:

$$(1 - a_k)^2 = d^4/4.$$

Proof:

Let $a_i = \langle f, \pi_i \rangle_{L^2} = \hat{f}(\{i\})$ for $i \in [n]$, and $a_0 = \hat{f}(\emptyset)$.

Let $\theta = \left(4\log_2(2/d) - 1\right)^{-1}$. There is $\theta \in (0, 1]$ because $d \le \sqrt{2}$.

Let $k \in \{0, 1, \ldots, n\}$ be such that $d = \|f - \pi_k\|_{L^2}$ (if the point of $\mathcal{A}_\pi$ closest to f is of the form $-\pi_k$ then a similar reasoning works).

Hence $d^2 = \|f\|_{L^2}^2 + \|\pi_k\|_{L^2}^2 - 2\langle f, \pi_k \rangle_{L^2} = 2(1 - a_k)$.
Remember:

$$(1 - a_k)^2 = d^4/4.$$

Since a function $h = f - \pi_k$ is $\{-2, 0, 2\}$-valued we get

$$\mu(h \neq 0) = \mu(\{x \in \{-1,1\}^n : h(x) \neq 0\}) = \frac{1}{4}\|h\|^2_{L^2} = (d/2)^2.$$

Therefore

$$d^4/2 = 4(d/2)^{\frac{4}{1+\theta}} = 4\Big(\mu(h \neq 0)\Big)^{\frac{2}{1+\theta}} = \|h\|^2_{L^{1+\theta}} \overset{B-B}{\geq}$$

($B - B$ is the classical $L^2 - L^{1+\theta}$ Bonami-Beckner inequality)

$$\sum_{A \subseteq [n]} \theta^{|A|} \cdot |\hat{h}(A)|^2 \geq \theta \cdot \sum_{A \subseteq [n]: |A| \leq 1} |\hat{h}(A)|^2 =$$

$$\theta \cdot \Big((1 - a_k)^2 + \sum_{i \in \{0,1,\dots,n\} \setminus \{k\}} a_i^2\Big) = \theta \cdot \Big(\frac{d^4}{4} + \sum_{i \in \{0,1,\dots,n\} \setminus \{k\}} a_i^2\Big),$$

so that

$$\sum_{i \in \{0,1,\dots,n\} \setminus \{k\}} a_i^2 \leq (2\theta^{-1} - 1)d^4/4 \leq 2d^4 \log_2(2/d). \quad (1)$$

Hypercontractivity

Since a function $h = f - \pi_k$ is $\{-2, 0, 2\}$-valued we get

$$\mu(h \neq 0) = \mu(\{x \in \{-1,1\}^n : h(x) \neq 0\}) = \frac{1}{4}\|h\|_{L^2}^2 = (d/2)^2.$$

Therefore

$$d^4/2 = 4(d/2)^{\frac{4}{1+\theta}} = 4\Big(\mu(h \neq 0)\Big)^{\frac{2}{1+\theta}} = \|h\|_{L^{1+\theta}}^2 \overset{B-B}{\geq}$$

($B - B$ is the classical $L^2 - L^{1+\theta}$ Bonami-Beckner inequality)

$$\sum_{A\subseteq[n]} \theta^{|A|} \cdot |\hat{h}(A)|^2 \geq \theta \cdot \sum_{A\subseteq[n]:|A|\leq 1} |\hat{h}(A)|^2 =$$

$$\theta \cdot \Big((1-a_k)^2 + \sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2\Big) = \theta \cdot \Big(\frac{d^4}{4} + \sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2\Big),$$

so that

$$\sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2 \leq (2\theta^{-1}-1)d^4/4 \leq 2d^4\log_2(2/d). \quad (1)$$

Hypercontractivity

Since a function $h = f - \pi_k$ is $\{-2, 0, 2\}$-valued we get

$$\mu(h \neq 0) = \mu(\{x \in \{-1,1\}^n : h(x) \neq 0\}) = \frac{1}{4}\|h\|_{L^2}^2 = (d/2)^2.$$

Therefore

$$d^4/2 = 4(d/2)^{\frac{4}{1+\theta}} = 4\Big(\mu(h \neq 0)\Big)^{\frac{2}{1+\theta}} = \|h\|_{L^{1+\theta}}^2 \overset{B-B}{\geq}$$

($B - B$ is the classical $L^2 - L^{1+\theta}$ Bonami-Beckner inequality)

$$\sum_{A \subseteq [n]} \theta^{|A|} \cdot |\hat{h}(A)|^2 \geq \theta \cdot \sum_{A \subseteq [n] : |A| \leq 1} |\hat{h}(A)|^2 =$$

$$\theta \cdot \Big((1-a_k)^2 + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big) = \theta \cdot \Big(\frac{d^4}{4} + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big),$$

so that

$$\sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2 \leq (2\theta^{-1} - 1)d^4/4 \leq 2d^4 \log_2(2/d). \quad (1)$$

Since a function $h = f - \pi_k$ is $\{-2, 0, 2\}$-valued we get

$$\mu(h \neq 0) = \mu(\{x \in \{-1,1\}^n : h(x) \neq 0\}) = \frac{1}{4}\|h\|_{L^2}^2 = (d/2)^2.$$

Therefore

$$d^4/2 = 4(d/2)^{\frac{4}{1+\theta}} = 4\Big(\mu(h \neq 0)\Big)^{\frac{2}{1+\theta}} = \|h\|_{L^{1+\theta}}^2 \overset{B-B}{\geq}$$

($B - B$ is the classical $L^2 - L^{1+\theta}$ Bonami-Beckner inequality)

$$\sum_{A \subseteq [n]} \theta^{|A|} \cdot |\hat{h}(A)|^2 \geq \theta \cdot \sum_{A \subseteq [n] : |A| \leq 1} |\hat{h}(A)|^2 =$$

$$\theta \cdot \Big((1 - a_k)^2 + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big) = \theta \cdot \Big(\frac{d^4}{4} + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big),$$

so that

$$\sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2 \leq (2\theta^{-1} - 1)d^4/4 \leq 2d^4 \log_2(2/d). \quad (1)$$

Hypercontractivity

Since a function $h = f - \pi_k$ is $\{-2, 0, 2\}$-valued we get

$$\mu(h \neq 0) = \mu(\{x \in \{-1,1\}^n : h(x) \neq 0\}) = \frac{1}{4}\|h\|_{L^2}^2 = (d/2)^2.$$

Therefore

$$d^4/2 = 4(d/2)^{\frac{4}{1+\theta}} = 4\Big(\mu(h \neq 0)\Big)^{\frac{2}{1+\theta}} = \|h\|_{L^{1+\theta}}^2 \overset{B-B}{\geq}$$

($B - B$ is the classical $L^2 - L^{1+\theta}$ Bonami-Beckner inequality)

$$\sum_{A \subseteq [n]} \theta^{|A|} \cdot |\hat{h}(A)|^2 \geq \theta \cdot \sum_{A \subseteq [n] : |A| \leq 1} |\hat{h}(A)|^2 =$$

$$\theta \cdot \Big((1 - a_k)^2 + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big) = \theta \cdot \Big(\frac{d^4}{4} + \sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2\Big),$$

so that

$$\sum_{i \in \{0,1,\ldots,n\} \setminus \{k\}} a_i^2 \leq (2\theta^{-1} - 1)d^4/4 \leq 2d^4 \log_2(2/d). \quad (1)$$

Proof of the discrete cube result - the end

$$\sum_{i=0}^{n} a_i^2 = \left(1-\frac{d^2}{2}\right)^2 + \sum_{i\in\{0,1,\dots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} \left(1-\frac{d^2}{2}\right)^2 + \frac{1}{4}(2\theta^{-1}-1)d^4$$

$$= 1 - d^2 + \frac{1}{2}\theta^{-1}d^4 \leq 1 - d^2 + 2d^4\log_2(2/d),$$

so

$$\rho^2 = \sum_{A\subseteq[n]:|A|\geq 2} |\hat{f}(A)|^2 = 1 - \sum_{i=0}^{n} a_i^2 \geq d^2 - 2d^4\log_2(2/d). \quad (2)$$

We finish the proof by observing that

$$\sum_{i\in\{0,1,\dots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} 2d^4\log_2(2/d) \overset{(2)}{\leq} 2\left(\rho^2+2d^4\log_2(2/d)\right)^2\log_2(2/d)$$

$$\overset{d\geq\rho}{\leq} 2\left(\rho^2+2d^4\log_2(2/\rho)\right)^2\log_2(2/\rho) \overset{d=O(\rho)}{=} 2\rho^4\log_2(2/\rho)+o(\rho^5),$$

uniformly, as $\rho\to 0^+$.

Proof of the discrete cube result - the end

$$\sum_{i=0}^{n} a_i^2 = \left(1-\frac{d^2}{2}\right)^2 + \sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} \left(1-\frac{d^2}{2}\right)^2 + \frac{1}{4}(2\theta^{-1}-1)d^4$$

$$= 1 - d^2 + \frac{1}{2}\theta^{-1}d^4 \leq 1 - d^2 + 2d^4\log_2(2/d),$$

so

$$\rho^2 = \sum_{A\subseteq[n]:|A|\geq 2} |\hat{f}(A)|^2 = 1 - \sum_{i=0}^{n} a_i^2 \geq d^2 - 2d^4\log_2(2/d). \quad (2)$$

We finish the proof by observing that

$$\sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} 2d^4\log_2(2/d) \overset{(2)}{\leq} 2\left(\rho^2+2d^4\log_2(2/d)\right)^2\log_2(2/d)$$

$$\overset{d\geq\rho}{\leq} 2\left(\rho^2+2d^4\log_2(2/\rho)\right)^2\log_2(2/\rho) \overset{d=O(\rho)}{=} 2\rho^4\log_2(2/\rho)+o(\rho^5),$$

uniformly, as $\rho \to 0^+$.

Proof of the discrete cube result - the end

$$\sum_{i=0}^{n} a_i^2 = \left(1-\frac{d^2}{2}\right)^2 + \sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} \left(1-\frac{d^2}{2}\right)^2 + \frac{1}{4}(2\theta^{-1}-1)d^4$$

$$= 1 - d^2 + \frac{1}{2}\theta^{-1}d^4 \leq 1 - d^2 + 2d^4\log_2(2/d),$$

so

$$\rho^2 = \sum_{A\subseteq[n]:|A|\geq 2} |\hat{f}(A)|^2 = 1 - \sum_{i=0}^{n} a_i^2 \geq d^2 - 2d^4\log_2(2/d). \quad (2)$$

We finish the proof by observing that

$$\sum_{i\in\{0,1,\ldots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} 2d^4\log_2(2/d) \overset{(2)}{\leq} 2\left(\rho^2+2d^4\log_2(2/d)\right)^2\log_2(2/d)$$

$$\overset{d\geq\rho}{\leq} 2\left(\rho^2+2d^4\log_2(2/\rho)\right)^2\log_2(2/\rho) \overset{d=O(\rho)}{=} 2\rho^4\log_2(2/\rho)+o(\rho^5),$$

uniformly, as $\rho\to 0^+$.

Proof of the discrete cube result - the end

$$\sum_{i=0}^{n} a_i^2 = \left(1-\frac{d^2}{2}\right)^2 + \sum_{i\in\{0,1,\dots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} \left(1-\frac{d^2}{2}\right)^2 + \frac{1}{4}(2\theta^{-1}-1)d^4$$

$$= 1 - d^2 + \frac{1}{2}\theta^{-1}d^4 \leq 1 - d^2 + 2d^4\log_2(2/d),$$

so

$$\rho^2 = \sum_{A\subseteq[n]:|A|\geq 2} |\hat{f}(A)|^2 = 1 - \sum_{i=0}^{n} a_i^2 \geq d^2 - 2d^4\log_2(2/d). \quad (2)$$

We finish the proof by observing that

$$\sum_{i\in\{0,1,\dots,n\}\setminus\{k\}} a_i^2 \overset{(1)}{\leq} 2d^4\log_2(2/d) \overset{(2)}{\leq} 2\left(\rho^2+2d^4\log_2(2/d)\right)^2\log_2(2/d)$$

$$\overset{d\geq\rho}{\leq} 2\left(\rho^2+2d^4\log_2(2/\rho)\right)^2\log_2(2/\rho) \overset{d=O(\rho)}{=} 2\rho^4\log_2(2/\rho)+o(\rho^5),$$

uniformly, as $\rho\to 0^+$.

Influences

For $f : \{-1,1\}^n \to \mathbb{R}$ and $i \in [n]$, let us define the i-th influence of f by

$$\mathrm{Inf}_i(f) = \sum_{S \subseteq [n]:\, i \in S} \left(\hat{f}(S)\right)^2 = \mathbb{E}[\mathrm{Var}_i(f)].$$

This quantity measures dependence of $f(x)$ on the i-th coordinate of x.

Influences

For $f : \{-1,1\}^n \to \mathbb{R}$ and $i \in [n]$, let us define the i-th influence of f by

$$\mathrm{Inf}_i(f) = \sum_{S \subseteq [n]:\, i \in S} \left(\hat{f}(S)\right)^2 = \mathbb{E}[\mathrm{Var}_i(f)].$$

This quantity measures dependence of $f(x)$ on the i-th coordinate of x.

Kahn, Kalai, and Linial proved that for every mean-zero function $f : \{-1,1\}^n \to \{-1,1\}$ there exists $i \in [n]$ such that $\mathrm{Inf}_i(f) \geq c \cdot \frac{\log n}{n}$, where $c > 0$ is some universal constant.

The assumption that $\mathbb{E}[f] = 0$ can be weakened, but not completely removed (since for $f \equiv 1$ all influences are obviously equal to zero).

The $\frac{\log n}{n}$ order of the bound is optimal: Tribes function.

The KKL Theorem

Kahn, Kalai, and Linial proved that for every mean-zero function $f : \{-1,1\}^n \to \{-1,1\}$ there exists $i \in [n]$ such that $\mathrm{Inf}_i(f) \geq c \cdot \frac{\log n}{n}$, where $c > 0$ is some universal constant.

The assumption that $\mathbb{E}[f] = 0$ can be weakened, but not completely removed (since for $f \equiv 1$ all influences are obviously equal to zero).

The $\frac{\log n}{n}$ order of the bound is optimal: Tribes function.

The KKL Theorem

Kahn, Kalai, and Linial proved that for every mean-zero function $f : \{-1,1\}^n \to \{-1,1\}$ there exists $i \in [n]$ such that $\mathrm{Inf}_i(f) \geq c \cdot \frac{\log n}{n}$, where $c > 0$ is some universal constant.

The assumption that $\mathbb{E}[f] = 0$ can be weakened, but not completely removed (since for $f \equiv 1$ all influences are obviously equal to zero).

The $\frac{\log n}{n}$ order of the bound is optimal: Tribes function.

Discrete partial derivative

For $x = (x_1, x_2, \ldots, x_n)$ and $i \in [n]$, let $\tau_i(x)$ denote the reflection of x with respect to the i-th coordinate:

$$\tau_i(x) = (x_1, x_2, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n).$$

Now we can define a linear partial derivative operator D_i acting on real-valued functions on the discrete cube. For $f : \{-1, 1\}^n \to \mathbb{R}$, we put

$$D_i(f)(x) = \left(f(x) - f(\tau_i(x))\right)/2.$$

We have

$$D_i f = \sum_{S \subseteq [n]:\, i \in S} \hat{f}(S) w_S.$$

and

$$\mathrm{Inf}_i(f) = \|D_i f\|_2^2.$$

Discrete partial derivative

For $x = (x_1, x_2, \ldots, x_n)$ and $i \in [n]$, let $\tau_i(x)$ denote the reflection of x with respect to the i-th coordinate:

$$\tau_i(x) = (x_1, x_2, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n).$$

Now we can define a linear partial derivative operator D_i acting on real-valued functions on the discrete cube. For $f : \{-1, 1\}^n \to \mathbb{R}$, we put

$$D_i(f)(x) = (f(x) - f(\tau_i(x)))/2.$$

We have

$$D_i f = \sum_{S \subseteq [n]:\, i \in S} \hat{f}(S) w_S.$$

and

$$\mathrm{Inf}_i(f) = \|D_i f\|_2^2.$$

Discrete partial derivative

For $x = (x_1, x_2, \ldots, x_n)$ and $i \in [n]$, let $\tau_i(x)$ denote the reflection of x with respect to the i-th coordinate:

$$\tau_i(x) = (x_1, x_2, \ldots, x_{i-1}, -x_i, x_{i+1}, \ldots, x_n).$$

Now we can define a linear partial derivative operator D_i acting on real-valued functions on the discrete cube. For $f : \{-1,1\}^n \to \mathbb{R}$, we put

$$D_i(f)(x) = \left(f(x) - f(\tau_i(x))\right)/2.$$

We have

$$D_i f = \sum_{S \subseteq [n]:\, i \in S} \hat{f}(S) w_s.$$

and

$$\mathrm{Inf}_i(f) = \|D_i f\|_2^2.$$

Second order quantities

For $i, j \in [n]$ with $i \neq j$, let $D_{i,j} = D_i \circ D_j$.

One easily checks that

$$D_{i,j} f = \sum_{S \subseteq [n]:\, i,j \in S} \hat{f}(S) w_S.$$

It is natural to define $\mathrm{Inf}_{i,j} f$ as $\|D_{i,j} f\|_2^2$:

$$\mathrm{Inf}_{i,j} f = \sum_{S \subseteq [n]:\, i,j \in [n]} \left(\hat{f}(S)\right)^2.$$

Second order quantities

For $i, j \in [n]$ with $i \neq j$, let $D_{i,j} = D_i \circ D_j$.

One easily checks that

$$D_{i,j} f = \sum_{S \subseteq [n]:\, i,j \in S} \hat{f}(S) w_S.$$

It is natural to define $\mathrm{Inf}_{i,j} f$ as $\|D_{i,j} f\|_2^2$:

$$\mathrm{Inf}_{i,j} f = \sum_{S \subseteq [n]:\, i,j \in [n]} \left(\hat{f}(S)\right)^2.$$

Second order quantities

For $i, j \in [n]$ with $i \neq j$, let $D_{i,j} = D_i \circ D_j$.

One easily checks that

$$D_{i,j} f = \sum_{S \subseteq [n]\colon i,j \in S} \hat{f}(S) w_S.$$

It is natural to define $\mathrm{Inf}_{i,j} f$ as $\|D_{i,j} f\|_2^2$:

$$\mathrm{Inf}_{i,j} f = \sum_{S \subseteq [n]\colon i,j \in [n]} \left(\hat{f}(S)\right)^2.$$