On mimicking Rademacher sums

and on some extensions of the FKN theorem

Krzysztof Oleszkiewicz

Institute of Mathematics, University of Warsaw

Functional Inequalities in Discrete Spaces with Applications
UC Berkeley's Simons Institute for the Theory of Computing
September 30 – October 4, 2013

Discrete cube - notation

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

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

Expectation of $f : C_n \longrightarrow \mathbb{R}$ is thus given by

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

L^p**-norm:** $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p \geq 1$.

Discrete cube - notation

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

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

Expectation of $f : C_n \longrightarrow \mathbb{R}$ is thus given by

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

L^p**-norm:** $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p \geq 1$.

Discrete cube - notation

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

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

Expectation of $f : C_n \longrightarrow \mathbb{R}$ is thus given by

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

L^p**-norm:** $\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$ for $p \geq 1$.

L^2-structure

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

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

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

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

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

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

$$w_A(x) = \prod_{i \in A} x_i,$$

$$w_\emptyset \equiv 1$$

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

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

Walsh system

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

$$w_A(x) = \prod_{i \in A} x_i,$$

$$w_\emptyset \equiv 1$$

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

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

Walsh system

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

$$w_A(x) = \prod_{i \in A} x_i,$$

$$w_\emptyset \equiv 1$$

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

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

Walsh system

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

$$w_A(x) = \prod_{i \in A} x_i,$$

$$w_\emptyset \equiv 1$$

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

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

Orthonormality

$\mathbb{E}[w_A] = 0$ for $A \neq \emptyset$, 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: $w_A \cdot w_B = w_{A \Delta B}$ thus

$$\langle w_A, w_B \rangle = \mathbb{E}[w_{A \Delta B}] = \delta_{A,B}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{A,B} = 1$ if $A = B$ and $\delta_{A,B} = 0$ if $A \neq B$ (Kronecker's delta).

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

Orthonormality

$\mathbb{E}[w_A] = 0$ for $A \neq \emptyset$, 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: $w_A \cdot w_B = w_{A \Delta B}$ thus

$$\langle w_A, w_B \rangle = \mathbb{E}[w_{A \Delta B}] = \delta_{A,B}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{A,B} = 1$ if $A = B$ and $\delta_{A,B} = 0$ if $A \neq B$ (Kronecker's delta).

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

Orthonormality

$\mathbb{E}[w_A] = 0$ for $A \neq \emptyset$, 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: $w_A \cdot w_B = w_{A \Delta B}$ thus

$$\langle w_A, w_B \rangle = \mathbb{E}[w_{A \Delta B}] = \delta_{A,B}$$

Here Δ denotes a symmetric set difference (XOR) while $\delta_{A,B} = 1$ if $A = B$ and $\delta_{A,B} = 0$ if $A \neq B$ (Kronecker's delta).

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

Walsh-Fourier expansion

The Walsh functions $(w_A)_{A \subseteq [n]}$ form an orthonormal system of cardinality 2^n, which is equal to the linear dimension of $\mathcal{H}_n$. Thus the system is complete and every $f \in \mathcal{H}_n$ admits the unique **Walsh-Fourier expansion:**

$$f = \sum_{A \subseteq [n]} \hat{f}(A) w_A,$$

with coefficients given by

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

Question

Let $N \geq 2$ and let us denote by V_N the linear span of $(w_A)_{A \subseteq [n]: |A| > N}$.

Problem (R. Bogucki, P. Nayar, M. Wojciechowski):
Let $S : \{-1, 1\}^n \to \mathbb{R}$ be defined by $S = r_1 + r_2 + \ldots + r_n$.
Estimate $\mathrm{dist}_{L^1}(S, V_N)$.

There is $\mathrm{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$.

Actually, for $S = \sum_{i=1}^n a_i r_i$ there is

$$\mathrm{dist}_{L^1}(S, V_N) \leq CN \cdot \max_i |a_i|,$$

and even some more precise estimates are available. However, in this presentation we will deal with the problem in its original form.

Question

Let $N \geq 2$ and let us denote by V_N the linear span of $(w_A)_{A\subseteq[n]:|A|>N}$.

Problem (R. Bogucki, P. Nayar, M. Wojciechowski):
Let $S : \{-1,1\}^n \to \mathbb{R}$ be defined by $S = r_1 + r_2 + \ldots + r_n$. Estimate $\operatorname{dist}_{L^1}(S, V_N)$.

There is $\operatorname{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$.

Actually, for $S = \sum_{i=1}^n a_i r_i$ there is

$$\operatorname{dist}_{L^1}(S, V_N) \leq CN \cdot \max_i |a_i|,$$

and even some more precise estimates are available. However, in this presentation we will deal with the problem in its original form.

Question

Let $N \geq 2$ and let us denote by V_N the linear span of $(w_A)_{A \subseteq [n]: |A| > N}$.

Problem (R. Bogucki, P. Nayar, M. Wojciechowski):
Let $S : \{-1,1\}^n \to \mathbb{R}$ be defined by $S = r_1 + r_2 + \ldots + r_n$.
Estimate $\operatorname{dist}_{L^1}(S, V_N)$.

There is $\operatorname{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$.

Actually, for $S = \sum_{i=1}^n a_i r_i$ there is

$$\operatorname{dist}_{L^1}(S, V_N) \leq CN \cdot \max_i |a_i|,$$

and even some more precise estimates are available. However, in this presentation we will deal with the problem in its original form.

Main result

Theorem

There exists a universal $\kappa > 0$ such that for any integers $N \geq 2$ and $n \geq 1$ there is a function $f : \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E}[|f|] \leq \kappa N$ and such that $\hat{f}(\{i\}) = 1$ for $1 \leq i \leq n$, and $\hat{f}(A) = 0$ for all $A \subseteq \{1,2,\ldots,n\}$ of cardinality $0, 2, 3, 4, \ldots, N$.

Moreover, the $O(N)$ bound is of optimal order. For $N = 2$ one can even find f satisfying the above conditions and such that $\mathbb{E}[|f|] = 1$.

Proof of the bound

Proof: For $N = 2$ it suffices to consider the function $f = \frac{1}{2}\prod_{i=1}^{n}(1 + r_i) - \frac{1}{2}\prod_{i=1}^{n}(1 - r_i)$. Obviously, $\mathbb{E}[|f|] = 1$. No better bound can be hoped for since $\mathbb{E}[fr_i] = 1$ implies $\mathbb{E}[|f|] \geq 1$.

For general $N \geq 2$ let us consider a function of Fejér type:

$$\psi_N(x) = \sum_{k=1}^{N} k \sin kx + \sum_{k=1}^{N} (N - k) \sin(N + k)x,$$

or, equivalently, $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$.

For some universal constants $\kappa_1, \kappa_2 > 0$ we have $|\psi_N(x)| \leq \kappa_1 N^2$ on $[-1/N, 1/N]$ and $|\psi_N(x)| \leq \kappa_2/x^2$ for $|x| > 1/N$.

Thus $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)|\,\mathrm{d}x \leq \kappa N$.

Proof of the bound

Proof: For $N = 2$ it suffices to consider the function $f = \frac{1}{2}\prod_{i=1}^{n}(1 + r_i) - \frac{1}{2}\prod_{i=1}^{n}(1 - r_i)$. Obviously, $\mathbb{E}[|f|] = 1$. No better bound can be hoped for since $\mathbb{E}[fr_i] = 1$ implies $\mathbb{E}[|f|] \geq 1$.

For general $N \geq 2$ let us consider a function of Fejér type:

$$\psi_N(x) = \sum_{k=1}^{N} k \sin kx + \sum_{k=1}^{N} (N - k) \sin(N + k)x,$$

or, equivalently, $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$.

For some universal constants $\kappa_1, \kappa_2 > 0$ we have $|\psi_N(x)| \leq \kappa_1 N^2$ on $[-1/N, 1/N]$ and $|\psi_N(x)| \leq \kappa_2/x^2$ for $|x| > 1/N$.

Thus $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)| \,\mathrm{d}x \leq \kappa N$.

Proof of the bound

Proof: For $N = 2$ it suffices to consider the function $f = \frac{1}{2}\prod_{i=1}^{n}(1+r_i) - \frac{1}{2}\prod_{i=1}^{n}(1-r_i)$. Obviously, $\mathbb{E}[|f|] = 1$. No better bound can be hoped for since $\mathbb{E}[fr_i] = 1$ implies $\mathbb{E}[|f|] \geq 1$.

For general $N \geq 2$ let us consider a function of Fejér type:

$$\psi_N(x) = \sum_{k=1}^{N} k \sin kx + \sum_{k=1}^{N} (N-k)\sin(N+k)x,$$

or, equivalently, $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$.

For some universal constants $\kappa_1, \kappa_2 > 0$ we have $|\psi_N(x)| \leq \kappa_1 N^2$ on $[-1/N, 1/N]$ and $|\psi_N(x)| \leq \kappa_2/x^2$ for $|x| > 1/N$.

Thus $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)|\,\mathrm{d}x \leq \kappa N$.

Proof of the bound

Proof: For $N = 2$ it suffices to consider the function $f = \frac{1}{2}\prod_{i=1}^{n}(1 + r_i) - \frac{1}{2}\prod_{i=1}^{n}(1 - r_i)$. Obviously, $\mathbb{E}[|f|] = 1$. No better bound can be hoped for since $\mathbb{E}[fr_i] = 1$ implies $\mathbb{E}[|f|] \geq 1$.

For general $N \geq 2$ let us consider a function of Fejér type:

$$\psi_N(x) = \sum_{k=1}^{N} k \sin kx + \sum_{k=1}^{N} (N - k) \sin(N + k)x,$$

or, equivalently, $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$.

For some universal constants $\kappa_1, \kappa_2 > 0$ we have $|\psi_N(x)| \leq \kappa_1 N^2$ on $[-1/N, 1/N]$ and $|\psi_N(x)| \leq \kappa_2/x^2$ for $|x| > 1/N$.

Thus $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)| \, \mathrm{d}x \leq \kappa N$.

Proof of the bound

Proof: For $N = 2$ it suffices to consider the function $f = \frac{1}{2}\prod_{i=1}^{n}(1+r_i) - \frac{1}{2}\prod_{i=1}^{n}(1-r_i)$. Obviously, $\mathbb{E}[|f|] = 1$. No better bound can be hoped for since $\mathbb{E}[fr_i] = 1$ implies $\mathbb{E}[|f|] \geq 1$.

For general $N \geq 2$ let us consider a function of Fejér type:

$$\psi_N(x) = \sum_{k=1}^{N} k \sin kx + \sum_{k=1}^{N} (N-k) \sin(N+k)x,$$

or, equivalently, $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$.

For some universal constants $\kappa_1, \kappa_2 > 0$ we have $|\psi_N(x)| \leq \kappa_1 N^2$ on $[-1/N, 1/N]$ and $|\psi_N(x)| \leq \kappa_2/x^2$ for $|x| > 1/N$.

Thus $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)| \, \mathrm{d}x \leq \kappa N$.

Proof of the bound - the end

Recall that we have $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$ and $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)|\,\mathrm{d}x \le \kappa N$.

One easily checks that $\int_{-\pi}^{\pi} \psi_N(x)\sin x\,\mathrm{d}x = \pi$, $\int_{-\pi}^{\pi} \psi_N(x)\,\mathrm{d}x = 0$, and $\int_{-\pi}^{\pi} \psi_N(x)\sin^m x\,\mathrm{d}x = 0$ for $2 \le m \le N$ (for even m this is obvious and for odd m one can use simple induction on m, expressing $\sin mx$ as a polynomial in $\sin x$).

Choosing $f(x)$ defined by the formula

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\prod_{i=1}^{n}(1+r_i \sin x)\,\mathrm{d}x = \sum_{A\subseteq[n]} \frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\sin^{|A|} x\,\mathrm{d}x \cdot w_A,$$

we end the proof.

Proof of the bound - the end

Recall that we have $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$ and $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)|\,\mathrm{d}x \le \kappa N$.

One easily checks that $\int_{-\pi}^{\pi} \psi_N(x)\sin x\,\mathrm{d}x = \pi$, $\int_{-\pi}^{\pi} \psi_N(x)\,\mathrm{d}x = 0$, and $\int_{-\pi}^{\pi} \psi_N(x)\sin^m x\,\mathrm{d}x = 0$ for $2 \le m \le N$ (for even m this is obvious and for odd m one can use simple induction on m, expressing $\sin mx$ as a polynomial in $\sin x$).

Choosing $f(x)$ defined by the formula

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\prod_{i=1}^{n}(1+r_i\sin x)\,\mathrm{d}x = \sum_{A\subseteq[n]} \frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\sin^{|A|} x\,\mathrm{d}x \cdot w_A,$$

we end the proof.

Proof of the bound - the end

Recall that we have $\psi_N(x) = \sin Nx \cdot \sin^2(Nx/2)/\sin^2(x/2)$ and $\frac{1}{\pi}\int_{-\pi}^{\pi} |\psi_N(x)|\,\mathrm{d}x \le \kappa N$.

One easily checks that $\int_{-\pi}^{\pi} \psi_N(x)\sin x\,\mathrm{d}x = \pi$, $\int_{-\pi}^{\pi} \psi_N(x)\,\mathrm{d}x = 0$, and $\int_{-\pi}^{\pi} \psi_N(x)\sin^m x\,\mathrm{d}x = 0$ for $2 \le m \le N$ (for even m this is obvious and for odd m one can use simple induction on m, expressing $\sin mx$ as a polynomial in $\sin x$).

Choosing $f(x)$ defined by the formula

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\prod_{i=1}^{n}(1+r_i \sin x)\,\mathrm{d}x = \sum_{A\subseteq[n]} \frac{1}{\pi}\int_{-\pi}^{\pi} \psi_N(x)\sin^{|A|} x\,\mathrm{d}x\cdot w_A,$$

we end the proof.

Optimality

Now we will prove that in general the linear order of the estimate cannot be improved:

There exists a universal constant $\eta > 0$ with the following property. For every $N \geq 2$ and every function $f : \{-1,1\}^{N^2} \to \mathbb{R}$ such that $\hat{f}(\{i\}) = 1$ for $1 \leq i \leq N^2$ and $\hat{f}(A) = 0$ for all $A \subseteq \{1, 2, \ldots, N^2\}$ of cardinality $0, 2, 3, 4, \ldots, N$ there is $\mathbb{E}[|f|] \geq \eta N$.

Proof of optimality

Let

$$W_N(t) = \sum_{k=0}^{[\frac{N-1}{2}]} \frac{(-1)^k t^{2k+1}}{(2k+1)!}, \quad R_N(t) = \sum_{[\frac{N+1}{2}]}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}.$$

Obviously, $W_N(t) + R_N(t) = \sin t$ and for $t \in [-N/6, N/6]$ we have

$$|R_N(t)| \le \sum_{[\frac{N+1}{2}]}^{\infty} (\mathrm{e}/6)^{2k+1} \le 2^{-N},$$

since $m! \ge (m/\mathrm{e})^m$. Hence $|W_N(t)| \le 2$ for $|t| \le N/6$. Thus

$$2\mathbb{E}[|f|] \ge \mathbb{E} f W_N \left(\frac{1}{6N} \sum_{i=1}^{N^2} r_i \right) = \mathbb{E} \left[\left(\sum_{i=1}^{N^2} r_i \right) W_N \left(\frac{1}{6N} \sum_{i=1}^{N^2} r_i \right) \right].$$

Indeed, $\deg W_N \le N$, so that $W_N(\frac{1}{6N} \sum_{i=1}^{N^2} r_i)$ is a Walsh-Fourier chaos of order not exceeding N.

Proof of optimality

Let

$$W_N(t) = \sum_{k=0}^{[\frac{N-1}{2}]} \frac{(-1)^k t^{2k+1}}{(2k+1)!}, \quad R_N(t) = \sum_{[\frac{N+1}{2}]}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}.$$

Obviously, $W_N(t) + R_N(t) = \sin t$ and for $t \in [-N/6, N/6]$ we have

$$|R_N(t)| \leq \sum_{[\frac{N+1}{2}]}^{\infty} (\mathrm{e}/6)^{2k+1} \leq 2^{-N},$$

since $m! \geq (m/\mathrm{e})^m$. Hence $|W_N(t)| \leq 2$ for $|t| \leq N/6$. Thus

$$2\mathbb{E}[|f|] \geq \mathbb{E} f W_N\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^{N^2} r_i\right) W_N\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right)\right].$$

Indeed, $\deg W_N \leq N$, so that $W_N(\frac{1}{6N}\sum_{i=1}^{N^2} r_i)$ is a Walsh-Fourier chaos of order not exceeding N.

Proof of optimality

Let

$$W_N(t) = \sum_{k=0}^{[\frac{N-1}{2}]} \frac{(-1)^k t^{2k+1}}{(2k+1)!}, \quad R_N(t) = \sum_{[\frac{N+1}{2}]}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}.$$

Obviously, $W_N(t) + R_N(t) = \sin t$ and for $t \in [-N/6, N/6]$ we have

$$|R_N(t)| \leq \sum_{[\frac{N+1}{2}]}^{\infty} (\mathrm{e}/6)^{2k+1} \leq 2^{-N},$$

since $m! \geq (m/\mathrm{e})^m$. Hence $|W_N(t)| \leq 2$ for $|t| \leq N/6$. Thus

$$2\mathbb{E}[|f|] \geq \mathbb{E} f W_N \left(\frac{1}{6N} \sum_{i=1}^{N^2} r_i \right) = \mathbb{E} \left[\left(\sum_{i=1}^{N^2} r_i \right) W_N \left(\frac{1}{6N} \sum_{i=1}^{N^2} r_i \right) \right].$$

Indeed, $\deg W_N \leq N$, so that $W_N(\frac{1}{6N} \sum_{i=1}^{N^2} r_i)$ is a Walsh-Fourier chaos of order not exceeding N.

On the other hand,

$$\mathbb{E}\left[\left(\sum_{i=1}^{N^2} r_i\right) W_N\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right)\right] =$$

$$\mathbb{E}\left[\left(\sum_{i=1}^{N^2} r_i\right) \sin\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right)\right] - \mathbb{E}\left[\left(\sum_{i=1}^{N^2} r_i\right) R_N\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right)\right]$$

$$\geq N^2 \mathbb{E} r_1 \sin\left(\frac{1}{6N}\sum_{i=1}^{N^2} r_i\right) - 2^{-N}\cdot N^2 =$$

$$N^2 \sin\left(\frac{1}{6N}\right)\cos^{N^2-1}\left(\frac{1}{6N}\right) - 2^{-N}\cdot N^2 \sim N.$$

The case of $n < N^2$

The preceding proof can be easily modified to cover the case $n \geq N^2$ instead of $n = N^2$.

The lower bound in the case $n < N^2$ follows - indeed, let N' denote the integer part of $\sqrt{n}$, so that $n \geq N'^2$. Then we have $N' \leq N$ and thus $V_N \subseteq V_{N'}$, so that

$$\mathrm{dist}_{L^1}(S, V_N) \geq \mathrm{dist}_{L^1}(S, V_{N'}) \geq \eta N' \simeq \sqrt{n}.$$

Note that we have also a trivial upper bound

$$\mathrm{dist}_{L^1}(S, V_N) \leq \mathbb{E}[|S|] \leq \left(\mathbb{E}[S^2]\right)^{1/2} = \sqrt{n}.$$

This ends the proof of the $\mathrm{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$ estimate.

The case of $n < N^2$

The preceding proof can be easily modified to cover the case $n \geq N^2$ instead of $n = N^2$.

The lower bound in the case $n < N^2$ follows - indeed, let N' denote the integer part of $\sqrt{n}$, so that $n \geq N'^2$. Then we have $N' \leq N$ and thus $V_N \subseteq V_{N'}$, so that

$$\mathrm{dist}_{L^1}(S, V_N) \geq \mathrm{dist}_{L^1}(S, V_{N'}) \geq \eta N' \simeq \sqrt{n}.$$

Note that we have also a trivial upper bound

$$\mathrm{dist}_{L^1}(S, V_N) \leq \mathbb{E}[|S|] \leq \left(\mathbb{E}[S^2]\right)^{1/2} = \sqrt{n}.$$

This ends the proof of the $\mathrm{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$ estimate.

The case of $n < N^2$

The preceding proof can be easily modified to cover the case $n \geq N^2$ instead of $n = N^2$.

The lower bound in the case $n < N^2$ follows - indeed, let N' denote the integer part of $\sqrt{n}$, so that $n \geq N'^2$. Then we have $N' \leq N$ and thus $V_N \subseteq V_{N'}$, so that

$$\mathrm{dist}_{L^1}(S, V_N) \geq \mathrm{dist}_{L^1}(S, V_{N'}) \geq \eta N' \simeq \sqrt{n}.$$

Note that we have also a trivial upper bound

$$\mathrm{dist}_{L^1}(S, V_N) \leq \mathbb{E}[|S|] \leq \left(\mathbb{E}[S^2]\right)^{1/2} = \sqrt{n}.$$

This ends the proof of the $\mathrm{dist}_{L^1}(S, V_N) \simeq \min(N, \sqrt{n})$ estimate.

The classical theorem of Friedgut, Kalai and Naor (2002), re-proved and extended by Kindler and Safra (2002), states that there exists a universal positive constant C such that for any positive integer n and any $f : \{-1,1\}^n \to \{-1,1\}$ we have

$$\mathbb{E}[(f-g)^2] \le C \cdot \sum_{A\subseteq[n]:|A|\ge 2} \left(\hat{f}(A)\right)^2$$

for some g of the form r_k, $-r_k$ (for some $k \in [n]$), 1, or -1.

We will discuss several new extensions to this theorem obtained in a joint paper with J. Jendrej and J. O. Wojtaszczyk, to appear in Theory of Computing soon (hopefully).

On some extensions of the FKN theorem

The classical theorem of Friedgut, Kalai and Naor (2002), re-proved and extended by Kindler and Safra (2002), states that there exists a universal positive constant C such that for any positive integer n and any $f : \{-1,1\}^n \to \{-1,1\}$ we have

$$\mathbb{E}[(f-g)^2] \leq C \cdot \sum_{A \subseteq [n]: |A| \geq 2} \left(\hat{f}(A)\right)^2$$

for some g of the form r_k, $-r_k$ (for some $k \in [n]$), 1, or -1.

We will discuss several new extensions to this theorem obtained in a joint paper with J. Jendrej and J. O. Wojtaszczyk, to appear in Theory of Computing soon (hopefully).

Main trick

Step 1 ([FKN]): Instead of the Boolean f defined on the discrete cube $\{-1,1\}^n$ (or, more generally, some product probability space), consider its orthogonal projection to the linear subspace of affine functions. On the discrete cube it reads as

$$\sum_{A\subseteq[n]:|A|\leq 1} \hat{f}(A)w_A = a_0 + a_1 r_1 + \ldots + a_n r_n.$$

Step 2: Prove an appropriate lemma of the following form. Let X and Y be independent. If their sum $X+Y$ is "concentrated" then at least one of the variables X, Y is *concentrated* (in a different sense).

Example: $\min(\mathrm{Var}(X), \mathrm{Var}(Y)) \leq C \cdot \mathrm{Var}(|X+Y|)$.

Main trick

Step 1 ([FKN]): Instead of the Boolean f defined on the discrete cube $\{-1,1\}^n$ (or, more generally, some product probability space), consider its orthogonal projection to the linear subspace of affine functions. On the discrete cube it reads as

$$\sum_{A\subseteq[n]:|A|\leq 1} \hat{f}(A)w_A = a_0 + a_1 r_1 + \ldots + a_n r_n.$$

Step 2: Prove an appropriate lemma of the following form. Let X and Y be independent. If their sum $X+Y$ is "concentrated" then at least one of the variables X, Y is *concentrated* (in a different sense).

Example: $\min(\mathrm{Var}(X), \mathrm{Var}(Y)) \leq C \cdot \mathrm{Var}(|X+Y|)$.

Main trick

Step 1 ([FKN]): Instead of the Boolean f defined on the discrete cube $\{-1,1\}^n$ (or, more generally, some product probability space), consider its orthogonal projection to the linear subspace of affine functions. On the discrete cube it reads as

$$\sum_{A \subseteq [n]:|A| \leq 1} \hat{f}(A) w_A = a_0 + a_1 r_1 + \ldots + a_n r_n.$$

Step 2: Prove an appropriate lemma of the following form. Let X and Y be independent. If their sum $X+Y$ is "concentrated" then at least one of the variables X, Y is *concentrated* (in a different sense).

Example: $\min(\mathrm{Var}(X), \mathrm{Var}(Y)) \leq C \cdot \mathrm{Var}(|X+Y|)$.

Reduction from n to two summands

Step 3: For a sum S of independent $X_1, X_2, \ldots, X_n$, deduce from Step 2 a lemma of the following form. If S is "concentrated" then there exists $k \in [n]$ such that $S - X_k$ is *concentrated*. Do it in the following way.

For $I \subseteq [n]$ let $S_I = \sum_{i \in I} X_i$. Choose a minimal I such that S_I is **not** *concentrated*. This implies, in particular, that I is nonempty. Choose any $k \in I$. Use Step 2 for $X = S_I$, $Y = S_{[n] \setminus I}$: $S = X + Y$ is "concentrated" and, by the choice of I, the summand X is not *concentrated*, so that $Y = S_{[n] \setminus I}$ must be *concentrated*. Also, $S_{I \setminus \{k\}}$ is *concentrated* because of the minimality of I. Thus $S_{[n] \setminus \{k\}} = S_{[n] \setminus I} + S_{I \setminus \{k\}}$ is *concentrated* as well.

Potential for extensions and modifications: for example "concentrated" may be replaced by "small", and instead of sums one may consider maxima, etc.

Reduction from n to two summands

Step 3: For a sum S of independent $X_1, X_2, \ldots, X_n$, deduce from Step 2 a lemma of the following form. If S is "concentrated" then there exists $k \in [n]$ such that $S - X_k$ is *concentrated*. Do it in the following way.

For $I \subseteq [n]$ let $S_I = \sum_{i \in I} X_i$. Choose a minimal I such that S_I is **not** *concentrated*. This implies, in particular, that I is nonempty. Choose any $k \in I$. Use Step 2 for $X = S_I$, $Y = S_{[n] \setminus I}$: $S = X + Y$ is "concentrated" and, by the choice of I, the summand X is not *concentrated*, so that $Y = S_{[n] \setminus I}$ must be *concentrated*. Also, $S_{I \setminus \{k\}}$ is *concentrated* because of the minimality of I. Thus $S_{[n] \setminus \{k\}} = S_{[n] \setminus I} + S_{I \setminus \{k\}}$ is *concentrated* as well.

Potential for extensions and modifications: for example "concentrated" may be replaced by "small", and instead of sums one may consider maxima, etc.

Reduction from n to two summands

Step 3: For a sum S of independent $X_1, X_2, \ldots, X_n$, deduce from Step 2 a lemma of the following form. If S is "concentrated" then there exists $k \in [n]$ such that $S - X_k$ is *concentrated*. Do it in the following way.

For $I \subseteq [n]$ let $S_I = \sum_{i \in I} X_i$. Choose a minimal I such that S_I is **not** *concentrated*. This implies, in particular, that I is nonempty. Choose any $k \in I$. Use Step 2 for $X = S_I$, $Y = S_{[n]\setminus I}$: $S = X + Y$ is "concentrated" and, by the choice of I, the summand X is not *concentrated*, so that $Y = S_{[n]\setminus I}$ must be *concentrated*. Also, $S_{I\setminus\{k\}}$ is *concentrated* because of the minimality of I. Thus $S_{[n]\setminus\{k\}} = S_{[n]\setminus I} + S_{I\setminus\{k\}}$ is *concentrated* as well.

Potential for extensions and modifications: for example "concentrated" may be replaced by "small", and instead of sums one may consider maxima, etc.

Symmetric case

Let X and Y be independent square-integrable random variables, at least one of them symmetric. Then

$$\min\Big(\mathrm{Var}(X), \mathrm{Var}(Y)\Big) \leq \frac{7+\sqrt{17}}{4} \cdot \mathrm{Var}(|X+Y|).$$

Let $(X_i)_{i=1}^n$ be a sequence of independent symmetric random variables. Then for some $k \in \{1, 2, \ldots, n\}$ we have

$$\mathrm{Var}\Big(\sum_{i \leq n: i \neq k} X_i\Big) \leq C \cdot \inf_{x \in \mathbb{R}} \mathrm{Var}\Big(\Big|x + \sum_{i \leq n} X_i\Big|\Big),$$

where C is a universal constant.

The result holds true with $C = (7+\sqrt{17})/2 \approx 5.56$. A simple example of $n = 3$ and X_1, X_2, X_3 i.i.d. symmetric ± 1 random variables indicates that the constant C cannot be less than $8/3 \approx 2.67$ (it suffices to check it for $x = 0$).

Symmetric case

Let X and Y be independent square-integrable random variables, at least one of them symmetric. Then

$$\min\Big(\mathrm{Var}(X), \mathrm{Var}(Y)\Big) \leq \frac{7+\sqrt{17}}{4} \cdot \mathrm{Var}(|X+Y|).$$

Let $(X_i)_{i=1}^n$ be a sequence of independent symmetric random variables. Then for some $k \in \{1, 2, \ldots, n\}$ we have

$$\mathrm{Var}\Big(\sum_{i \leq n : i \neq k} X_i\Big) \leq C \cdot \inf_{x \in \mathbb{R}} \mathrm{Var}\Big(\Big|x + \sum_{i \leq n} X_i\Big|\Big),$$

where C is a universal constant.

The result holds true with $C = (7+\sqrt{17})/2 \approx 5.56$. A simple example of $n = 3$ and X_1, X_2, X_3 i.i.d. symmetric ± 1 random variables indicates that the constant C cannot be less than $8/3 \approx 2.67$ (it suffices to check it for $x = 0$).

Symmetric case

Let X and Y be independent square-integrable random variables, at least one of them symmetric. Then

$$\min\Big(\mathrm{Var}(X), \mathrm{Var}(Y)\Big) \leq \frac{7+\sqrt{17}}{4} \cdot \mathrm{Var}(|X+Y|).$$

Let $(X_i)_{i=1}^n$ be a sequence of independent symmetric random variables. Then for some $k \in \{1, 2, \ldots, n\}$ we have

$$\mathrm{Var}\Big(\sum_{i \leq n : i \neq k} X_i\Big) \leq C \cdot \inf_{x \in \mathbb{R}} \mathrm{Var}\Big(\Big|x + \sum_{i \leq n} X_i\Big|\Big),$$

where C is a universal constant.

The result holds true with $C = (7+\sqrt{17})/2 \approx 5.56$. A simple example of $n = 3$ and X_1, X_2, X_3 i.i.d. symmetric ± 1 random variables indicates that the constant C cannot be less than $8/3 \approx 2.67$ (it suffices to check it for $x = 0$).

FKN theorem

The result from the preceding slide contains the FKN theorem, with a reasonable constant, as its special case. Actually, by using a less elementary method, specific to the case of the discrete cube, we were able to further improve the bounds in the FKN theorem (independently an analogous strengthening was obtained by Ryan O'Donnell).

Extension

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \leq \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \leq 25\rho$ or $\mathrm{Var}(Y) \leq 25\rho$.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \leq \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that

$$\mathrm{Var}(S - X_k) \leq 50\rho.$$

The $O(\rho)$ order of the bound cannot be improved in general. If we, however, take into account an additional parameter $\mathrm{Var}(S)$ then one can easily strengthen the estimate.

Extension

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \le 25\rho$ or $\mathrm{Var}(Y) \le 25\rho$.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that

$$\mathrm{Var}(S - X_k) \le 50\rho.$$

The $O(\rho)$ order of the bound cannot be improved in general. If we, however, take into account an additional parameter $\mathrm{Var}(S)$ then one can easily strengthen the estimate.

Extension

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \le 25\rho$ or $\mathrm{Var}(Y) \le 25\rho$.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that

$$\mathrm{Var}(S - X_k) \le 50\rho.$$

The $O(\rho)$ order of the bound cannot be improved in general. If we, however, take into account an additional parameter $\mathrm{Var}(S)$ then one can easily strengthen the estimate.

Rubinstein-type variant

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \le C\rho^2/\mathrm{Var}(X+Y)$ or $\mathrm{Var}(Y) \le C\rho^2/\mathrm{Var}(X+Y)$, where C is a universal positive constant.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that $\mathrm{Var}(S - X_k) \le K\rho^2/\mathrm{Var}(S)$, with a universal $K > 0$.

This Rubinstein-type bound immediately implies the estimate from the preceding slide (unsurprisingly, taking into account an additional parameter adds some precision). The new two-variable lemma has essentially the same proof as the old one.

Aviad Rubinstein, *Boolean functions whose Fourier transform is concentrated on pair-wise disjoint subsets of the inputs*, MSc Thesis, Tel-Aviv University, 2012

Rubinstein-type variant

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \le C\rho^2/\mathrm{Var}(X+Y)$ or $\mathrm{Var}(Y) \le C\rho^2/\mathrm{Var}(X+Y)$, where C is a universal positive constant.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^{n} X_i$. Assume $\mathbb{E}(|S|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that $\mathrm{Var}(S - X_k) \le K\rho^2/\mathrm{Var}(S)$, with a universal $K > 0$.

This Rubinstein-type bound immediately implies the estimate from the preceding slide (unsurprisingly, taking into account an additional parameter adds some precision). The new two-variable lemma has essentially the same proof as the old one.

Aviad Rubinstein, *Boolean functions whose Fourier transform is concentrated on pair-wise disjoint subsets of the inputs*, MSc Thesis, Tel-Aviv University, 2012

Rubinstein-type variant

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \leq \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \leq C\rho^2/\mathrm{Var}(X+Y)$ or $\mathrm{Var}(Y) \leq C\rho^2/\mathrm{Var}(X+Y)$, where C is a universal positive constant.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \leq \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that $\mathrm{Var}(S - X_k) \leq K\rho^2/\mathrm{Var}(S)$, with a universal $K > 0$.

This Rubinstein-type bound immediately implies the estimate from the preceding slide (unsurprisingly, taking into account an additional parameter adds some precision). The new two-variable lemma has essentially the same proof as the old one.

Aviad Rubinstein, *Boolean functions whose Fourier transform is concentrated on pair-wise disjoint subsets of the inputs*, MSc Thesis, Tel-Aviv University, 2012

Rubinstein-type variant

Let X and Y be independent square-integrable random variables. Assume $\mathbb{E}(|X+Y|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then $\mathrm{Var}(X) \le C\rho^2/\mathrm{Var}(X+Y)$ or $\mathrm{Var}(Y) \le C\rho^2/\mathrm{Var}(X+Y)$, where C is a universal positive constant.

Let $X_1, X_2, \ldots, X_n$ be independent square-integrable random variables and let $S = \sum_{i=1}^n X_i$. Assume $\mathbb{E}(|S|-1)^2 \le \rho^2$ for some $\rho \in (0,1]$. Then there exists some $k \in [n]$ such that $\mathrm{Var}(S - X_k) \le K\rho^2/\mathrm{Var}(S)$, with a universal $K > 0$.

This Rubinstein-type bound immediately implies the estimate from the preceding slide (unsurprisingly, taking into account an additional parameter adds some precision). The new two-variable lemma has essentially the same proof as the old one.

Aviad Rubinstein, *Boolean functions whose Fourier transform is concentrated on pair-wise disjoint subsets of the inputs*, MSc Thesis, Tel-Aviv University, 2012

Structural theorem

Let $\varepsilon, \delta > 0$. Let A be a finite subset of a separable Banach space V, with $|A| \geq 2$, such that $\|x - y\| \geq 3|A|\varepsilon$ for any distinct $x, y \in A$. Let $\xi_1, \ldots, \xi_n$ be independent V-valued random vectors and $S = \sum_{i=1}^n \xi_i$. Assume that $\mathbb{P}(\mathrm{dist}(S, A) > \varepsilon) \leq \delta$. For $I \subseteq [n]$ let $S_I = \sum_{i \in I} \xi_i$. Then there exists a nonnegative integer $k < |A|$ and $\{i_1, \ldots, i_k\} \subseteq [n]$ such that for some $v \in V$

$$\mathbb{P}\left(\left\|S_{[n]\setminus\{i_1,\ldots,i_k\}} - v\right\| > |A|\varepsilon\right) \leq |A|^2 \delta^{1/|A|}$$

for some $v \in V$.

Moreover, if V is a Hilbert space and $k > 0$ then there exist vectors $v_1, \ldots, v_k \in V$ and nonempty sets $B_1, \ldots, B_k \subseteq A$ with $\sum_{m=1}^k (|B_m| - 1) < |A|$ such that

$$\mathbb{P}\left(\mathrm{dist}(\xi_{i_m}, v_m + B_m) > \varepsilon\right) \leq 2|A|\delta^{1/(|A|-1)}$$

for every $m \in [k]$.

Structural theorem

Let $\varepsilon, \delta > 0$. Let A be a finite subset of a separable Banach space V, with $|A| \geq 2$, such that $\|x - y\| \geq 3|A|\varepsilon$ for any distinct $x, y \in A$. Let $\xi_1, \ldots, \xi_n$ be independent V-valued random vectors and $S = \sum_{i=1}^n \xi_i$. Assume that $\mathbb{P}(\mathrm{dist}(S, A) > \varepsilon) \leq \delta$. For $I \subseteq [n]$ let $S_I = \sum_{i \in I} \xi_i$. Then there exists a nonnegative integer $k < |A|$ and $\{i_1, \ldots, i_k\} \subseteq [n]$ such that for some $v \in V$

$$\mathbb{P}\left(\left\|S_{[n]\setminus\{i_1,\ldots,i_k\}} - v\right\| > |A|\varepsilon\right) \leq |A|^2 \delta^{1/|A|}$$

for some $v \in V$.

Moreover, if V is a Hilbert space and $k > 0$ then there exist vectors $v_1, \ldots, v_k \in V$ and nonempty sets $B_1, \ldots, B_k \subseteq A$ with $\sum_{m=1}^k (|B_m| - 1) < |A|$ such that

$$\mathbb{P}\left(\mathrm{dist}(\xi_{i_m}, v_m + B_m) > \varepsilon\right) \leq 2|A|\delta^{1/(|A|-1)}$$

for every $m \in [k]$.

Geometric problem

For a general Banach space V a slightly weaker result holds true. Its proof goes via essential reduction to a $2|A|$-dimensional linear subspace and use of F. John's theorem therein to deduce the general case from the result in Hilbert spaces.

We will say that a subset of a metric space is Δ-separated if it does not contain a pair of distinct points whose distance is less than Δ.

Question: Let A and B be finite 1-separated subsets of a normed linear space. Does it imply that there exists $C \subseteq A + B$ with $|C| \geq |A| + |B| - 1$ which is also 1-separated?

True in Hilbert spaces.

For a general Banach space V a slightly weaker result holds true. Its proof goes via essential reduction to a $2|A|$-dimensional linear subspace and use of F. John's theorem therein to deduce the general case from the result in Hilbert spaces.

We will say that a subset of a metric space is Δ-separated if it does not contain a pair of distinct points whose distance is less than Δ.

Question: Let A and B be finite 1-separated subsets of a normed linear space. Does it imply that there exists $C \subseteq A + B$ with $|C| \geq |A| + |B| - 1$ which is also 1-separated?

True in Hilbert spaces.

For a general Banach space V a slightly weaker result holds true. Its proof goes via essential reduction to a $2|A|$-dimensional linear subspace and use of F. John's theorem therein to deduce the general case from the result in Hilbert spaces.

We will say that a subset of a metric space is Δ-separated if it does not contain a pair of distinct points whose distance is less than Δ.

Question: Let A and B be finite 1-separated subsets of a normed linear space. Does it imply that there exists $C \subseteq A + B$ with $|C| \geq |A| + |B| - 1$ which is also 1-separated?

True in Hilbert spaces.