Probabilistic inequalities

Krzysztof Oleszkiewicz

Warsaw, November 17, 2016

Rademacher random variables

Throughout the lecture, $r_1, r_2, \ldots$ denote independent symmetric ± 1 real random variables. They are called Rademacher random variables (or: symmetric Bernoulli random variables).

An easy and standard way to construct the sequence $r_1, r_2, \ldots, r_n$: consider the discrete cube $\{-1,1\}^n$ equipped with the normalized counting (i.e. uniform probability) measure $\mathbf{P} = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$, so that $\mathbf{P}(A) = \mathrm{card}(A)/2^n$ for every $A \subseteq \{-1,1\}^n$.

Then simply define $r_k : \{-1,1\}^n \to \mathbf{R}$ by $r_k(x) = x_k$, for $1 \leq k \leq n$ and $x = (x_1, x_2, \ldots, x_n)$.

Thus, one can think of them as coordinate functions on the discrete cube. Less formally, but equivalently and more intuitively, one can also treat them as outcomes of n symmetric coin-tossing experiments (with heads $\equiv -1$ and tails $\equiv 1$).

Rademacher random variables

Throughout the lecture, $r_1, r_2, \ldots$ denote independent symmetric ± 1 real random variables. They are called Rademacher random variables (or: symmetric Bernoulli random variables).

An easy and standard way to construct the sequence $r_1, r_2, \ldots, r_n$: consider the discrete cube $\{-1, 1\}^n$ equipped with the normalized counting (i.e. uniform probability) measure $\mathbf{P} = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$, so that $\mathbf{P}(A) = \mathrm{card}(A)/2^n$ for every $A \subseteq \{-1, 1\}^n$.

Then simply define $r_k : \{-1, 1\}^n \to \mathbf{R}$ by $r_k(x) = x_k$, for $1 \leq k \leq n$ and $x = (x_1, x_2, \ldots, x_n)$.

Thus, one can think of them as coordinate functions on the discrete cube. Less formally, but equivalently and more intuitively, one can also treat them as outcomes of n symmetric coin-tossing experiments (with heads $\equiv -1$ and tails $\equiv 1$).

Rademacher random variables

Throughout the lecture, $r_1, r_2, \ldots$ denote independent symmetric ± 1 real random variables. They are called Rademacher random variables (or: symmetric Bernoulli random variables).

An easy and standard way to construct the sequence $r_1, r_2, \ldots, r_n$: consider the discrete cube $\{-1,1\}^n$ equipped with the normalized counting (i.e. uniform probability) measure $\mathbf{P} = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$, so that $\mathbf{P}(A) = \mathrm{card}(A)/2^n$ for every $A \subseteq \{-1,1\}^n$.

Then simply define $r_k : \{-1,1\}^n \to \mathbf{R}$ by $r_k(x) = x_k$, for $1 \leq k \leq n$ and $x = (x_1, x_2, \ldots, x_n)$.

Thus, one can think of them as coordinate functions on the discrete cube. Less formally, but equivalently and more intuitively, one can also treat them as outcomes of n symmetric coin-tossing experiments (with heads $\equiv -1$ and tails $\equiv 1$).

Rademacher sums

$S = a_1 r_1 + a_2 r_2 + \ldots + a_n r_n$, for real coefficients $a_1, a_2, \ldots, a_n$, is called a (weighted) **Rademacher sum**. More intuitively,

$$S = \pm a_1 \pm a_2 \pm \cdots \pm a_n$$

with a random, independent and symmetric choice of signs.

For $p > 0$, we define the p-th absolute moment of S by

$$\mathbf{E}|S|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left| \sum_{k=1}^{n} a_k x_k \right|^p .$$

Often it is more convenient to consider the p-th norm of S,

$$\|S\|_p := (\mathbf{E}|S|^p)^{1/p} .$$

Rademacher sums

$S = a_1 r_1 + a_2 r_2 + \ldots + a_n r_n$, for real coefficients $a_1, a_2, \ldots, a_n$, is called a (weighted) **Rademacher sum**. More intuitively,

$$S = \pm a_1 \pm a_2 \pm \cdots \pm a_n$$

with a random, independent and symmetric choice of signs.

For $p > 0$, we define the p-th absolute moment of S by

$$\mathbf{E}|S|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left| \sum_{k=1}^{n} a_k x_k \right|^p .$$

Often it is more convenient to consider the p-th norm of S,

$$\|S\|_p := \left(\mathbf{E}|S|^p\right)^{1/p} .$$

Rademacher sums

$S = a_1 r_1 + a_2 r_2 + \ldots + a_n r_n$, for real coefficients $a_1, a_2, \ldots, a_n$, is called a (weighted) **Rademacher sum**. More intuitively,

$$S = \pm a_1 \pm a_2 \pm \cdots \pm a_n$$

with a random, independent and symmetric choice of signs.

For $p > 0$, we define the p-th absolute moment of S by

$$\mathbf{E}|S|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left| \sum_{k=1}^{n} a_k x_k \right|^p .$$

Often it is more convenient to consider the p-th norm of S,

$$\|S\|_p := \left(\mathbf{E}|S|^p\right)^{1/p} .$$

Vector-valued Rademacher sums

Given vectors $v_1, v_2, \ldots, v_n$ of a normed linear space $(V, \|\cdot\|)$, one may define a **vector-valued Rademacher sum**, $S = r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = \pm v_1 \pm v_2 \pm \ldots \pm v_n$.

For such a V-valued sum, one studies its p-th absolute moment, $\mathbf{E}\|S\|^p$, for $p > 0$:

$$\mathbf{E}\|S\|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left\| \sum_{k=1}^{n} x_k v_k \right\|^p,$$

with the p-th norm, $\|S\|_p$, defined as $(\mathbf{E}\|S\|^p)^{1/p}$.

$(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$ yields standard (real-valued) Rademacher sums.

Vector-valued Rademacher sums

Given vectors $v_1, v_2, \ldots, v_n$ of a normed linear space $(V, \|\cdot\|)$, one may define a **vector-valued Rademacher sum**, $S = r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = \pm v_1 \pm v_2 \pm \ldots \pm v_n$.

For such a V-valued sum, one studies its p-th absolute moment, $\mathbf{E}\|S\|^p$, for $p > 0$:

$$\mathbf{E}\|S\|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left\| \sum_{k=1}^{n} x_k v_k \right\|^p,$$

with the p-th norm, $\|S\|_p$, defined as $(\mathbf{E}\|S\|^p)^{1/p}$.

$(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$ yields standard (real-valued) Rademacher sums.

Vector-valued Rademacher sums

Given vectors $v_1, v_2, \ldots, v_n$ of a normed linear space $(V, \|\cdot\|)$, one may define a **vector-valued Rademacher sum**, $S = r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = \pm v_1 \pm v_2 \pm \ldots \pm v_n$.

For such a V-valued sum, one studies its p-th absolute moment, $\mathbf{E}\|S\|^p$, for $p > 0$:

$$\mathbf{E}\|S\|^p = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} \left\| \sum_{k=1}^{n} x_k v_k \right\|^p,$$

with the p-th norm, $\|S\|_p$, defined as $(\mathbf{E}\|S\|^p)^{1/p}$.

$(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$ yields standard (real-valued) Rademacher sums.

Aleksandr Khin[±t]chin[±e], 1923

For any $p, q > 0$ there exists a positive constant $C_{p,q}$ such that every real-valued Rademacher sum S satisfies the inequality

$$\|S\|_p \leq C_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $C_{p,q}$?

Easy case: $C_{p,q}^{\mathrm{opt}} = 1$ whenever $p \leq q$.
From now on, we assume that $p > q$.

Important case: $p = 2$ or $q = 2$
(since $\mathbf{E}S^2 = \sum a_k^2$ is easy to control).

Easy: $C_{p,2}^{\mathrm{opt}}$ for p even (then $\mathbf{E}|S|^p = \mathbf{E}S^p$ is a polynomial in a_k's).

Aleksandr Khin[$\pm$t]chin[$\pm$e], 1923

For any $p, q > 0$ there exists a positive constant $C_{p,q}$ such that every real-valued Rademacher sum S satisfies the inequality

$$\|S\|_p \le C_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $C_{p,q}$?

Easy case: $C_{p,q}^{\mathrm{opt}} = 1$ whenever $p \le q$.
From now on, we assume that $p > q$.

Important case: $p = 2$ or $q = 2$
(since $\mathbf{E}S^2 = \sum a_k^2$ is easy to control).

Easy: $C_{p,2}^{\mathrm{opt}}$ for p even (then $\mathbf{E}|S|^p = \mathbf{E}S^p$ is a polynomial in a_k's).

Aleksandr Khin[$\pm$t]chin[$\pm$e], 1923

For any $p, q > 0$ there exists a positive constant $C_{p,q}$ such that every real-valued Rademacher sum S satisfies the inequality

$$\|S\|_p \leq C_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $C_{p,q}$?

Easy case: $C_{p,q}^{\mathrm{opt}} = 1$ whenever $p \leq q$.
From now on, we assume that $p > q$.

Important case: $p = 2$ or $q = 2$
(since $\mathbf{E}S^2 = \sum a_k^2$ is easy to control).

Easy: $C_{p,2}^{\mathrm{opt}}$ for p even (then $\mathbf{E}|S|^p = \mathbf{E}S^p$ is a polynomial in a_k's).

Aleksandr Khin[±t]chin[±e], 1923

For any $p, q > 0$ there exists a positive constant $C_{p,q}$ such that every real-valued Rademacher sum S satisfies the inequality

$$\|S\|_p \leq C_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $C_{p,q}$?

Easy case: $C_{p,q}^{\mathrm{opt}} = 1$ whenever $p \leq q$.
From now on, we assume that $p > q$.

Important case: $p = 2$ or $q = 2$
(since $\mathbf{E}S^2 = \sum a_k^2$ is easy to control).

Easy: $C_{p,2}^{\mathrm{opt}}$ for p even (then $\mathbf{E}|S|^p = \mathbf{E}S^p$ is a polynomial in a_k's).

Aleksandr Khin[±t]chin[±e], 1923

For any $p, q > 0$ there exists a positive constant $C_{p,q}$ such that every real-valued Rademacher sum S satisfies the inequality

$$\|S\|_p \leq C_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $C_{p,q}$?

Easy case: $C_{p,q}^{\mathrm{opt}} = 1$ whenever $p \leq q$.
From now on, we assume that $p > q$.

Important case: $p = 2$ or $q = 2$
(since $\mathbf{E}S^2 = \sum a_k^2$ is easy to control).

Easy: $C_{p,2}^{\mathrm{opt}}$ for p even (then $\mathbf{E}|S|^p = \mathbf{E}S^p$ is a polynomial in a_k's).

Optimal constants in the Khinchin inequality

$C_{p,2}^{\mathrm{opt}}$ for $p > 3$ (**Whittle 1960**),

$C_{2,1}^{\mathrm{opt}} = \sqrt{2}$ (Littlewood's problem, **Szarek 1976**).

Haagerup 1982: $C_{p,2}^{\mathrm{opt}} = \gamma_p/\gamma_2$ for $p > 2$,
$C_{2,q}^{\mathrm{opt}} = \max\left(2^{\frac{1}{q}-\frac{1}{2}}, \gamma_2/\gamma_q\right)$ for $q \in (0,2)$,

where $\gamma_p := \|G\|_p$ with $G \sim \mathcal{N}(0,1)$,
i.e. $\gamma_p = 2^{1/2}\pi^{-\frac{1}{2p}}\left(\Gamma\left(\frac{p+1}{2}\right)\right)^{1/p}$, in particular $\gamma_2 = 1$.

$C_{p,q}^{\mathrm{opt}} = \gamma_p/\gamma_q$ if $p > q$ are both even
(**Nayar & O. 2012**; partial case $q|p$ solved by **Czerwiński 2008**).

$C^{\mathrm{opt}}_{p,2}$ for $p > 3$ (**Whittle 1960**),

$C^{\mathrm{opt}}_{2,1} = \sqrt{2}$ (Littlewood's problem, **Szarek 1976**).

Haagerup 1982: $C^{\mathrm{opt}}_{p,2} = \gamma_p/\gamma_2$ for $p > 2$,

$C^{\mathrm{opt}}_{2,q} = \max\left(2^{\frac{1}{q}-\frac{1}{2}}, \gamma_2/\gamma_q\right)$ for $q \in (0,2)$,

where $\gamma_p := \|G\|_p$ with $G \sim \mathcal{N}(0,1)$,

i.e. $\gamma_p = 2^{1/2}\pi^{-\frac{1}{2p}}\left(\Gamma\left(\frac{p+1}{2}\right)\right)^{1/p}$, in particular $\gamma_2 = 1$.

$C^{\mathrm{opt}}_{p,q} = \gamma_p/\gamma_q$ if $p > q$ are both even
(**Nayar & O. 2012**; partial case $q|p$ solved by **Czerwiński 2008**).

Optimal constants in the Khinchin inequality

$C^{\mathrm{opt}}_{p,2}$ for $p > 3$ (**Whittle 1960**),

$C^{\mathrm{opt}}_{2,1} = \sqrt{2}$ (Littlewood's problem, **Szarek 1976**).

Haagerup 1982: $C^{\mathrm{opt}}_{p,2} = \gamma_p/\gamma_2$ for $p > 2$,
$C^{\mathrm{opt}}_{2,q} = \max\left(2^{\frac{1}{q}-\frac{1}{2}}, \gamma_2/\gamma_q\right)$ for $q \in (0,2)$,
where $\gamma_p := \|G\|_p$ with $G \sim \mathcal{N}(0,1)$,
i.e. $\gamma_p = 2^{1/2}\pi^{-\frac{1}{2p}}\left(\Gamma\left(\frac{p+1}{2}\right)\right)^{1/p}$, in particular $\gamma_2 = 1$.

$C^{\mathrm{opt}}_{p,q} = \gamma_p/\gamma_q$ if $p > q$ are both even
(**Nayar & O. 2012**; partial case $q|p$ solved by **Czerwiński 2008**).

Optimal constants in the Khinchin inequality

$C_{p,2}^{\mathrm{opt}}$ for $p > 3$ (**Whittle 1960**),

$C_{2,1}^{\mathrm{opt}} = \sqrt{2}$ (Littlewood's problem, **Szarek 1976**).

Haagerup 1982: $C_{p,2}^{\mathrm{opt}} = \gamma_p/\gamma_2$ for $p > 2$,

$C_{2,q}^{\mathrm{opt}} = \max\left(2^{\frac{1}{q}-\frac{1}{2}}, \gamma_2/\gamma_q\right)$ for $q \in (0,2)$,

where $\gamma_p := \|G\|_p$ with $G \sim \mathcal{N}(0,1)$,

i.e. $\gamma_p = 2^{1/2}\pi^{-\frac{1}{2p}}\left(\Gamma\left(\frac{p+1}{2}\right)\right)^{1/p}$, in particular $\gamma_2 = 1$.

$C_{p,q}^{\mathrm{opt}} = \gamma_p/\gamma_q$ if $p > q$ are both even
(**Nayar & O. 2012**; partial case $q|p$ solved by **Czerwiński 2008**).

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \leq K_{p,q} \cdot \|S\|_q .$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \leq q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$.
Certainly, $K_{p,q}^{\mathrm{opt}} \geq C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,
thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;

$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Khinchin-Kahane inequality

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \le K_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \le q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$. Certainly, $K_{p,q}^{\mathrm{opt}} \ge C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,
thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;
$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \leq K_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \leq q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$.
Certainly, $K_{p,q}^{\mathrm{opt}} \geq C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,
thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;
$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Khinchin-Kahane inequality

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \leq K_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \leq q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$.
Certainly, $K_{p,q}^{\mathrm{opt}} \geq C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,

thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;

$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Khinchin-Kahane inequality

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \leq K_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \leq q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$.
Certainly, $K_{p,q}^{\mathrm{opt}} \geq C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,
thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;
$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Kahane 1964: For any $p, q > 0$ there exists a positive constant $K_{p,q}$ such that for every vector-valued Rademacher sum S we have

$$\|S\|_p \le K_{p,q} \cdot \|S\|_q.$$

Given p and q, what is the optimal value of the constant $K_{p,q}$?

Obviously, $K_{p,q}^{\mathrm{opt}} = 1 = C_{p,q}^{\mathrm{opt}}$ whenever $p \le q$.

Kwapień's Conjecture: For every $p > q > 0$ there is $K_{p,q}^{\mathrm{opt}} = C_{p,q}^{\mathrm{opt}}$.
Certainly, $K_{p,q}^{\mathrm{opt}} \ge C_{p,q}^{\mathrm{opt}}$.

Known: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$ **(Latała & O. 1994)**,

thus also $K_{p,q}^{\mathrm{opt}} = 2^{\frac{1}{q}-\frac{1}{p}}$ for $q \in (0,1]$ and $p \in [q,2]$;

$K_{4,2}^{\mathrm{opt}} = \sqrt[4]{3}$ **(Kwapień, Latała & O. 1996)**.

Probabilistic inequality as a goal, harmonic analysis as a tool

We will prove that, for any vector-valued Rademacher sum S,

$$\mathbf{E}\|S\|^2 \le 2(\mathbf{E}\|S\|)^2,$$

i.e.

$$\|S\|_2 \le \sqrt{2} \cdot \|S\|_1.$$

Note: If $(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$, $n = 2$, and $v_1 = v_2 = 1$, then $\mathbf{P}(S = 0) = 1/2 = \mathbf{P}(|S| = 2)$, so that $\|S\|_2 = \sqrt{2}$ and $\|S\|_1 = 1$. Thus the constant $\sqrt{2}$ cannot be improved.

The proof that will be presented is an insightful reinterpretation of **Latała & O. 1994**, due to **Kwapień**.

We will prove that, for any vector-valued Rademacher sum S,

$$\mathbf{E}\|S\|^2 \le 2(\mathbf{E}\|S\|)^2,$$

i.e.

$$\|S\|_2 \le \sqrt{2} \cdot \|S\|_1.$$

Note: If $(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$, $n = 2$, and $v_1 = v_2 = 1$, then $\mathbf{P}(S = 0) = 1/2 = \mathbf{P}(|S| = 2)$, so that $\|S\|_2 = \sqrt{2}$ and $\|S\|_1 = 1$. Thus the constant $\sqrt{2}$ cannot be improved.

The proof that will be presented is an insightful reinterpretation of **Latała & O. 1994**, due to **Kwapień**.

Probabilistic inequality as a goal: $K_{2,1}^{\mathrm{opt}} = \sqrt{2}$

We will prove that, for any vector-valued Rademacher sum S,

$$\mathbf{E}\|S\|^2 \le 2(\mathbf{E}\|S\|)^2,$$

i.e.

$$\|S\|_2 \le \sqrt{2} \cdot \|S\|_1 .$$

Note: If $(V, \|\cdot\|) = (\mathbf{R}, |\cdot|)$, $n = 2$, and $v_1 = v_2 = 1$, then $\mathbf{P}(S = 0) = 1/2 = \mathbf{P}(|S| = 2)$, so that $\|S\|_2 = \sqrt{2}$ and $\|S\|_1 = 1$. Thus the constant $\sqrt{2}$ cannot be improved.

The proof that will be presented is an insightful reinterpretation of **Latała & O. 1994**, due to **Kwapień**.

Harmonic analysis on the discrete cube as a tool

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

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

$$w_A(x) = \prod_{k \in A} x_k,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection ($k \in [n]$).

Recall: $r_1, r_2, \ldots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

Harmonic analysis on the discrete cube as a tool

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

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

$$w_A(x) = \prod_{k \in A} x_k ,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection ($k \in [n]$).

Recall: $r_1, r_2, \dots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

Harmonic analysis on the discrete cube as a tool

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

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

$$w_A(x) = \prod_{k \in A} x_k ,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection ($k \in [n]$).

Recall: $r_1, r_2, \ldots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

Harmonic analysis on the discrete cube as a tool

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

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

$$w_A(x) = \prod_{k \in A} x_k,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection ($k \in [n]$).

Recall: $r_1, r_2, \ldots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

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

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

$$w_A(x) = \prod_{k \in A} x_k,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection ($k \in [n]$).

Recall: $r_1, r_2, \ldots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

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

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

$$w_A(x) = \prod_{k \in A} x_k,$$

$$w_\emptyset \equiv 1.$$

Thus, $r_k = w_{\{k\}}$ is just the k-th coordinate projection $(k \in [n])$.

Recall: $r_1, r_2, \dots, r_n$ is a Rademacher sequence
– independent symmetric ± 1 Bernoulli random variables.

Note: $w_A : \{-1, 1\}^n \to \mathbf{R}$ can be expressed as $w_A = \prod_{k \in A} r_k$.

Scalar product: For $f, g : \{-1,1\}^n \longrightarrow \mathbf{R}$, let

$$\langle f, g\rangle = \mathbf{E}[f \cdot g] = 2^{-n} \cdot \sum_{x\in\{-1,1\}^n} f(x)g(x).$$

Note that $\langle f, f\rangle = \mathbf{E}f^2 = 2^{-n} \cdot \sum_{x\in\{-1,1\}^n} f(x)^2$.

L^2-**norm:** $\|f\|_2 := \sqrt{\langle f, f\rangle} = (\mathbf{E}f^2)^{1/2}$.

L^2-structure

Scalar product: For $f, g : \{-1, 1\}^n \longrightarrow \mathbf{R}$, let

$$\langle f, g \rangle = \mathbf{E}[f \cdot g] = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} f(x)g(x).$$

Note that $\langle f, f \rangle = \mathbf{E}f^2 = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} f(x)^2$.

L^2-norm: $\|f\|_2 := \sqrt{\langle f, f \rangle} = (\mathbf{E}f^2)^{1/2}$.

L^2-structure

Scalar product: For $f, g : \{-1,1\}^n \longrightarrow \mathbf{R}$, let

$$\langle f, g \rangle = \mathbf{E}[f \cdot g] = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} f(x)g(x).$$

Note that $\langle f, f \rangle = \mathbf{E}f^2 = 2^{-n} \cdot \sum_{x \in \{-1,1\}^n} f(x)^2$.

L^2-**norm:** $\|f\|_2 := \sqrt{\langle f, f \rangle} = (\mathbf{E}f^2)^{1/2}$.

Orthonormality

$\mathbf{E}[w_A] = 0$ for $A \neq \emptyset$ and $\mathbf{E}[w_\emptyset] = 1$.

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

Another explanation: Let $k \in A$. Then, flipping the sign of the k-th coordinate of x changes the sign of $w_A(x)$. Thus, values of w_A on $\{-1,1\}^n$ can be grouped into pairs adding to zero.

Orthonormality: $w_A \cdot w_B = w_{A\Delta B}$, so that

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

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

We have proved that the Walsh system $(w_A)_{A\subseteq[n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of the space of all real functions on $\{-1,1\}^n$, it spans the whole space.

Orthonormality

$\mathbf{E}[w_A] = 0$ for $A \neq \emptyset$ and $\mathbf{E}[w_\emptyset] = 1$.

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

Another explanation: Let $k \in A$. Then, flipping the sign of the k-th coordinate of x changes the sign of $w_A(x)$. Thus, values of w_A on $\{-1, 1\}^n$ can be grouped into pairs adding to zero.

Orthonormality: $w_A \cdot w_B = w_{A \Delta B}$, so that

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

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

We have proved that the Walsh system $(w_A)_{A \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of the space of all real functions on $\{-1, 1\}^n$, it spans the whole space.

Orthonormality

$\mathbf{E}[w_A] = 0$ for $A \neq \emptyset$ and $\mathbf{E}[w_\emptyset] = 1$.

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

Another explanation: Let $k \in A$. Then, flipping the sign of the k-th coordinate of x changes the sign of $w_A(x)$. Thus, values of w_A on $\{-1,1\}^n$ can be grouped into pairs adding to zero.

Orthonormality: $w_A \cdot w_B = w_{A \Delta B}$, so that

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

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

We have proved that the Walsh system $(w_A)_{A \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of the space of all real functions on $\{-1,1\}^n$, it spans the whole space.

Orthonormality

$\mathbf{E}[w_A] = 0$ for $A \neq \emptyset$ and $\mathbf{E}[w_\emptyset] = 1$.

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

Another explanation: Let $k \in A$. Then, flipping the sign of the k-th coordinate of x changes the sign of $w_A(x)$. Thus, values of w_A on $\{-1, 1\}^n$ can be grouped into pairs adding to zero.

Orthonormality: $w_A \cdot w_B = w_{A \Delta B}$, so that

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

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

We have proved that the Walsh system $(w_A)_{A \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of the space of all real functions on $\{-1, 1\}^n$, it spans the whole space.

Orthonormality

$\mathbf{E}[w_A] = 0$ for $A \neq \emptyset$ and $\mathbf{E}[w_\emptyset] = 1$.

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

Another explanation: Let $k \in A$. Then, flipping the sign of the k-th coordinate of x changes the sign of $w_A(x)$. Thus, values of w_A on $\{-1,1\}^n$ can be grouped into pairs adding to zero.

Orthonormality: $w_A \cdot w_B = w_{A\Delta B}$, so that

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

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

We have proved that the Walsh system $(w_A)_{A \subseteq [n]}$ is orthonormal (and therefore linearly independent). Since it is of cardinality 2^n, which is equal to the linear dimension of the space of all real functions on $\{-1,1\}^n$, it spans the whole space.

Sum over neighbours operator

Vertices x and y of $\{-1,1\}^n$ are called *neighbours* ($x \sim y$), if

$$\mathrm{card}\,\{k \in [n] : x_k \neq y_k\} = 1.$$

We define a linear operator K acting on the space of all real-valued functions on the discrete cube. Namely, for $f : \{-1,1\}^n \to \mathbf{R}$, let $Kf : \{-1,1\}^n \to \mathbf{R}$ be defined by

$$Kf(x) := \sum_{y \in \{-1,1\}^n :\, x \sim y} f(y).$$

Sum over neighbours operator

Vertices x and y of $\{-1,1\}^n$ are called *neighbours* $(x \sim y)$, if

$$\mathrm{card}\,\{k \in [n] : x_k \neq y_k\} = 1.$$

We define a linear operator K acting on the space of all real-valued functions on the discrete cube. Namely, for $f : \{-1,1\}^n \to \mathbf{R}$, let $Kf : \{-1,1\}^n \to \mathbf{R}$ be defined by

$$Kf(x) := \sum_{y \in \{-1,1\}^n :\, x \sim y} f(y).$$

Eigenstructure of the sum of over neighbours operator

Let $A \subseteq [n]$, and let us denote by $|A|$ the cardinality of A. Then

$$Kw_A = (n - 2|A|) \cdot w_A,$$

i.e. the Walsh function w_A is an eigenfunction (eigenvector) of the operator K, with eigenvalue $n - 2|A|$.

Indeed, for $x \in \{-1,1\}^n$, we have

$$(Kw_A)(x) = \sum_{y \in \{-1,1\}^n : \, x \sim y} w_A(y).$$

For $x \sim y$, let $k(x,y)$ be the only $k \in [n]$ such that $x_k \neq y_k$.
If $k(x,y) \in A$, then $w_A(y) = -w_A(x)$.
If $k(x,y) \notin A$, then $w_A(y) = w_A(x)$.
Thus,
$(Kw_A)(x) = |A| \cdot (-w_A(x)) + (n - |A|) \cdot w_A(x) = (n - 2|A|)w_A(x)$.

Eigenstructure of the sum of over neighbours operator

Let $A \subseteq [n]$, and let us denote by $|A|$ the cardinality of A. Then

$$Kw_A = (n - 2|A|) \cdot w_A,$$

i.e. the Walsh function w_A is an eigenfunction (eigenvector) of the operator K, with eigenvalue $n - 2|A|$.

Indeed, for $x \in \{-1, 1\}^n$, we have

$$(Kw_A)(x) = \sum_{y \in \{-1,1\}^n :\, x \sim y} w_A(y).$$

For $x \sim y$, let $k(x, y)$ be the only $k \in [n]$ such that $x_k \neq y_k$.
If $k(x, y) \in A$, then $w_A(y) = -w_A(x)$.
If $k(x, y) \notin A$, then $w_A(y) = w_A(x)$.
Thus,
$(Kw_A)(x) = |A| \cdot (-w_A(x)) + (n - |A|) \cdot w_A(x) = (n - 2|A|)w_A(x).$

Eigenstructure of the sum of over neighbours operator

Let $A \subseteq [n]$, and let us denote by $|A|$ the cardinality of A. Then

$$Kw_A = (n - 2|A|) \cdot w_A,$$

i.e. the Walsh function w_A is an eigenfunction (eigenvector) of the operator K, with eigenvalue $n - 2|A|$.

Indeed, for $x \in \{-1, 1\}^n$, we have

$$(Kw_A)(x) = \sum_{y \in \{-1,1\}^n :\, x \sim y} w_A(y).$$

For $x \sim y$, let $k(x, y)$ be the only $k \in [n]$ such that $x_k \neq y_k$.
If $k(x, y) \in A$, then $w_A(y) = -w_A(x)$.
If $k(x, y) \notin A$, then $w_A(y) = w_A(x)$.
Thus,
$(Kw_A)(x) = |A| \cdot (-w_A(x)) + (n - |A|) \cdot w_A(x) = (n - 2|A|)w_A(x)$.

Triangle inequality

Recall: $(V, \|\cdot\|)$ is a normed linear space and $v_1, v_2, \ldots, v_n \in V$. For $g : \{-1,1\}^n \to \mathbf{R}$ defined by $g(x) = \|\sum_{k=1}^n x_k v_k\|$, we have

$$Kg \geq (n-2) \cdot g.$$

Indeed, by the triangle inequality,

$$\begin{aligned}(Kg)(x) &= \sum_{y\in\{-1,1\}^n:\, x\sim y} \left\|\sum_{k=1}^n y_k v_k\right\| \\ &\geq \left\|\sum_{k=1}^n \left(\sum_{y\in\{-1,1\}^n:\, x\sim y} y_k\right) v_k\right\| \\ &= |n-2| \cdot \left\|\sum_{k=1}^n x_k v_k\right\| = |n-2| g(x),\end{aligned}$$

since $\sum_{y\in\{-1,1\}^n:\, x\sim y} y_k = (n-1)\cdot x_k + 1\cdot(-x_k) = (n-2)x_k$.

Triangle inequality

Recall: $(V, \|\cdot\|)$ is a normed linear space and $v_1, v_2, \ldots, v_n \in V$.
For $g : \{-1,1\}^n \to \mathbf{R}$ defined by $g(x) = \|\sum_{k=1}^n x_k v_k\|$, we have

$$Kg \geq (n-2)\cdot g.$$

Indeed, by the triangle inequality,

$$(Kg)(x) = \sum_{y \in \{-1,1\}^n :\, x \sim y} \left\| \sum_{k=1}^n y_k v_k \right\|$$

$$\geq \left\| \sum_{k=1}^n \left(\sum_{y \in \{-1,1\}^n :\, x \sim y} y_k \right) v_k \right\|$$

$$= |n-2| \cdot \left\| \sum_{k=1}^n x_k v_k \right\| = |n-2| g(x),$$

since $\sum_{y \in \{-1,1\}^n :\, x \sim y} y_k = (n-1)\cdot x_k + 1 \cdot (-x_k) = (n-2) x_k$.

Triangle inequality

Recall: $(V, \|\cdot\|)$ is a normed linear space and $v_1, v_2, \ldots, v_n \in V$.

For $g : \{-1,1\}^n \to \mathbf{R}$ defined by $g(x) = \|\sum_{k=1}^n x_k v_k\|$, we have

$$Kg \geq (n-2) \cdot g.$$

Indeed, by the triangle inequality,

$$(Kg)(x) = \sum_{y \in \{-1,1\}^n :\, x \sim y} \left\| \sum_{k=1}^n y_k v_k \right\|$$

$$\geq \left\| \sum_{k=1}^n \left(\sum_{y \in \{-1,1\}^n :\, x \sim y} y_k \right) v_k \right\|$$

$$= |n-2| \cdot \left\| \sum_{k=1}^n x_k v_k \right\| = |n-2| g(x),$$

since $\sum_{y \in \{-1,1\}^n :\, x \sim y} y_k = (n-1) \cdot x_k + 1 \cdot (-x_k) = (n-2) x_k$.

Two ways to deal with $\langle g, Kg\rangle$

We have proved the pointwise inequality $Kg \geq (n-2)g$.
Since g is nonnegative, we have also $g \cdot Kg \geq (n-2)g^2$, and thus

$$\langle g, Kg\rangle = \mathbf{E}[g \cdot Kg] \geq (n-2)\mathbf{E}g^2.$$

On the other hand, g admits a unique Fourier-Walsh expansion $g = \sum_{A\subseteq[n]} a_A w_A$, with some real coefficients $(a_A)_{A\subseteq[n]}$. Since

$$Kg = \sum_{A\subseteq[n]} a_A \cdot Kw_A = \sum_{A\subseteq[n]} (n-2|A|)a_A w_A,$$

we have

$$\langle g, Kg\rangle = \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} (n-2|B|)a_B w_B \right\rangle$$

$$= \sum_{A\subseteq[n]} \sum_{B\subseteq[n]} (n-2|B|)a_A a_B \langle w_A, w_B\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2.$$

Two ways to deal with $\langle g, Kg\rangle$

We have proved the pointwise inequality $Kg \geq (n-2)g$.
Since g is nonnegative, we have also $g \cdot Kg \geq (n-2)g^2$, and thus

$$\langle g, Kg\rangle = \mathbf{E}[g \cdot Kg] \geq (n-2)\mathbf{E}g^2.$$

On the other hand, g admits a unique Fourier-Walsh expansion $g = \sum_{A\subseteq[n]} a_A w_A$, with some real coefficients $(a_A)_{A\subseteq[n]}$. Since

$$Kg = \sum_{A\subseteq[n]} a_A \cdot Kw_A = \sum_{A\subseteq[n]} (n-2|A|)a_A w_A,$$

we have

$$\langle g, Kg\rangle = \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} (n-2|B|)a_B w_B \right\rangle$$

$$= \sum_{A\subseteq[n]} \sum_{B\subseteq[n]} (n-2|B|)a_A a_B \langle w_A, w_B\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2.$$

We have proved the pointwise inequality $Kg \geq (n-2)g$.
Since g is nonnegative, we have also $g \cdot Kg \geq (n-2)g^2$, and thus

$$\langle g, Kg\rangle = \mathbf{E}[g \cdot Kg] \geq (n-2)\mathbf{E}g^2.$$

On the other hand, g admits a unique Fourier-Walsh expansion $g = \sum_{A\subseteq[n]} a_A w_A$, with some real coefficients $(a_A)_{A\subseteq[n]}$. Since

$$Kg = \sum_{A\subseteq[n]} a_A \cdot Kw_A = \sum_{A\subseteq[n]} (n-2|A|)a_A w_A,$$

we have

$$\langle g, Kg\rangle = \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} (n-2|B|)a_B w_B \right\rangle$$

$$= \sum_{A\subseteq[n]} \sum_{B\subseteq[n]} (n-2|B|)a_A a_B \langle w_A, w_B\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2.$$

Putting things together

Putting together the two approaches, we arrive at

$$(n-2)\mathbf{E}g^2 \le \langle g, Kg\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2$$

$$\le na_\emptyset^2 + (n-2)\cdot\sum_{k=1}^{n} a_{\{k\}}^2 + (n-4)\cdot \sum_{A\subseteq[n]:\,|A|\ge 2} a_A^2$$

$$= 4a_\emptyset^2 + 2\sum_{k=1}^{n} a_{\{k\}}^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$$

$$= 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2,$$

because $a_{\{k\}} = \langle g, w_{\{k\}}\rangle = \langle g, r_k\rangle = \mathbf{E}[g\cdot r_k] = 0$, for $k\in[n]$ (indeed, g is an even function on the discrete cube, due to the symmetry of the norm, and Rademacher functions r_k are odd).

Putting things together

Putting together the two approaches, we arrive at

$$(n-2)\mathbf{E}g^2 \le \langle g, Kg\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2$$

$$\le na_\emptyset^2 + (n-2)\cdot\sum_{k=1}^n a_{\{k\}}^2 + (n-4)\cdot \sum_{A\subseteq[n]:\,|A|\ge 2} a_A^2$$

$$= 4a_\emptyset^2 + 2\sum_{k=1}^n a_{\{k\}}^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$$

$$= 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2,$$

because $a_{\{k\}} = \langle g, w_{\{k\}}\rangle = \langle g, r_k\rangle = \mathbf{E}[g\cdot r_k] = 0$, for $k\in[n]$ (indeed, g is an even function on the discrete cube, due to the symmetry of the norm, and Rademacher functions r_k are odd).

Putting things together

Putting together the two approaches, we arrive at

$$(n-2)\mathbf{E}g^2 \le \langle g, Kg\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2$$

$$\le na_\emptyset^2 + (n-2)\cdot\sum_{k=1}^n a_{\{k\}}^2 + (n-4)\cdot \sum_{A\subseteq[n]:\,|A|\ge 2} a_A^2$$

$$= 4a_\emptyset^2 + 2\sum_{k=1}^n a_{\{k\}}^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$$

$$= 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2,$$

because $a_{\{k\}} = \langle g, w_{\{k\}}\rangle = \langle g, r_k\rangle = \mathbf{E}[g\cdot r_k] = 0$, for $k\in[n]$ (indeed, g is an even function on the discrete cube, due to the symmetry of the norm, and Rademacher functions r_k are odd).

Putting things together

Putting together the two approaches, we arrive at

$$(n-2)\mathbf{E}g^2 \le \langle g, Kg\rangle = \sum_{A\subseteq[n]} (n-2|A|)a_A^2$$

$$\le na_\emptyset^2 + (n-2)\cdot\sum_{k=1}^{n} a_{\{k\}}^2 + (n-4)\cdot \sum_{A\subseteq[n]:\,|A|\ge 2} a_A^2$$

$$= 4a_\emptyset^2 + 2\sum_{k=1}^{n} a_{\{k\}}^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$$

$$= 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2,$$

because $a_{\{k\}} = \langle g, w_{\{k\}}\rangle = \langle g, r_k\rangle = \mathbf{E}[g\cdot r_k] = 0$, for $k\in[n]$ (indeed, g is an even function on the discrete cube, due to the symmetry of the norm, and Rademacher functions r_k are odd).

The end of the proof

We have proved that $(n-2)\mathbf{E}g^2 \le 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$.
Now it suffices to observe that

$$a_\emptyset = \langle g, w_\emptyset\rangle = \langle g, 1\rangle = \mathbf{E}[g\cdot 1] = \mathbf{E}g,$$

while (the Plancherel theorem for the discrete cube setting)

$$\sum_{A\subseteq[n]} a_A^2 = \sum_{A\subseteq[n]}\sum_{B\subseteq[n]} a_A a_B \langle w_A, w_B\rangle$$

$$= \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} a_B w_B \right\rangle = \langle g, g\rangle = \mathbf{E}[g\cdot g] = \mathbf{E}g^2.$$

Thus,

$$(n-2)\mathbf{E}g^2 \le 4(\mathbf{E}g)^2 + (n-4)\mathbf{E}g^2,$$

i.e., after cancellations, $\mathbf{E}\|S\|^2 = \mathbf{E}g^2 \le 2(\mathbf{E}g)^2 = 2(\mathbf{E}\|S\|)^2$, and the proof is finished.

The end of the proof

We have proved that $(n-2)\mathbf{E}g^2 \le 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$. Now it suffices to observe that

$$a_\emptyset = \langle g, w_\emptyset\rangle = \langle g, 1\rangle = \mathbf{E}[g\cdot 1] = \mathbf{E}g,$$

while (the Plancherel theorem for the discrete cube setting)

$$\sum_{A\subseteq[n]} a_A^2 = \sum_{A\subseteq[n]}\sum_{B\subseteq[n]} a_A a_B \langle w_A, w_B\rangle$$

$$= \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} a_B w_B \right\rangle = \langle g, g\rangle = \mathbf{E}[g\cdot g] = \mathbf{E}g^2.$$

Thus,

$$(n-2)\mathbf{E}g^2 \le 4(\mathbf{E}g)^2 + (n-4)\mathbf{E}g^2,$$

i.e., after cancellations, $\mathbf{E}\|S\|^2 = \mathbf{E}g^2 \le 2(\mathbf{E}g)^2 = 2(\mathbf{E}\|S\|)^2$, and the proof is finished.

The end of the proof

We have proved that $(n-2)\mathbf{E}g^2 \le 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$.
Now it suffices to observe that

$$a_\emptyset = \langle g, w_\emptyset\rangle = \langle g, 1\rangle = \mathbf{E}[g\cdot 1] = \mathbf{E}g,$$

while (the Plancherel theorem for the discrete cube setting)

$$\sum_{A\subseteq[n]} a_A^2 = \sum_{A\subseteq[n]}\sum_{B\subseteq[n]} a_A a_B \langle w_A, w_B\rangle$$

$$= \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} a_B w_B \right\rangle = \langle g, g\rangle = \mathbf{E}[g\cdot g] = \mathbf{E}g^2.$$

Thus,

$$(n-2)\mathbf{E}g^2 \le 4(\mathbf{E}g)^2 + (n-4)\mathbf{E}g^2,$$

i.e., after cancellations, $\mathbf{E}\|S\|^2 = \mathbf{E}g^2 \le 2(\mathbf{E}g)^2 = 2(\mathbf{E}\|S\|)^2$,
and the proof is finished.

The end of the proof

We have proved that $(n-2)\mathbf{E}g^2 \le 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$.
Now it suffices to observe that

$$a_\emptyset = \langle g, w_\emptyset\rangle = \langle g, 1\rangle = \mathbf{E}[g\cdot 1] = \mathbf{E}g,$$

while (the Plancherel theorem for the discrete cube setting)

$$\sum_{A\subseteq[n]} a_A^2 = \sum_{A\subseteq[n]}\sum_{B\subseteq[n]} a_A a_B \langle w_A, w_B\rangle$$

$$= \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} a_B w_B \right\rangle = \langle g, g\rangle = \mathbf{E}[g\cdot g] = \mathbf{E}g^2.$$

Thus,

$$(n-2)\mathbf{E}g^2 \le 4(\mathbf{E}g)^2 + (n-4)\mathbf{E}g^2,$$

i.e., after cancellations, $\mathbf{E}\|S\|^2 = \mathbf{E}g^2 \le 2(\mathbf{E}g)^2 = 2(\mathbf{E}\|S\|)^2$, and the proof is finished.

The end of the proof

We have proved that $(n-2)\mathbf{E}g^2 \le 4a_\emptyset^2 + (n-4)\cdot\sum_{A\subseteq[n]} a_A^2$.
Now it suffices to observe that

$$a_\emptyset = \langle g, w_\emptyset\rangle = \langle g, 1\rangle = \mathbf{E}[g\cdot 1] = \mathbf{E}g,$$

while (the Plancherel theorem for the discrete cube setting)

$$\sum_{A\subseteq[n]} a_A^2 = \sum_{A\subseteq[n]}\sum_{B\subseteq[n]} a_A a_B \langle w_A, w_B\rangle$$

$$= \left\langle \sum_{A\subseteq[n]} a_A w_A, \sum_{B\subseteq[n]} a_B w_B \right\rangle = \langle g, g\rangle = \mathbf{E}[g\cdot g] = \mathbf{E}g^2.$$

Thus,

$$(n-2)\mathbf{E}g^2 \le 4(\mathbf{E}g)^2 + (n-4)\mathbf{E}g^2,$$

i.e., after cancellations, $\mathbf{E}\|S\|^2 = \mathbf{E}g^2 \le 2(\mathbf{E}g)^2 = 2(\mathbf{E}\|S\|)^2$,
and the proof is finished.

Probabilistic inequality as a tool

for proving

a theorem in harmonic analysis

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002):
For $f : \{-1,1\}^n \to \{-1,1\}$,

$\{-1,1\}$-valued functions are called *Boolean*

A Boolean function on the discrete cube
models an n-bit-input $\to$ one-bit-output process

FKN Theorem

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002):
For $f : \{-1,1\}^n \to \{-1,1\}$,

$\{-1,1\}$-valued functions are called *Boolean*

A Boolean function on the discrete cube
models an n-bit-input $\longrightarrow$ one-bit-output process

FKN Theorem

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002):
For $f:\{-1,1\}^n \to \{-1,1\}$, let $f=\sum_{A\subseteq[n]} a_A w_A$ be its unique Fourier-Walsh expansion, and let

$$\rho := \sum_{A\subseteq[n]:\,|A|\geq 2} a_A^2.$$

Recall: by the Plancherel theorem, $\sum_{A\subseteq[n]} a_A^2 = \mathbf{E}f^2 = \mathbf{E}1 = 1.$

FKN Theorem

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002):
For $f : \{-1,1\}^n \to \{-1,1\}$, let $f = \sum_{A\subseteq[n]} a_A w_A$ be its unique Fourier-Walsh expansion, and let

$$\rho := \sum_{A\subseteq[n]:\,|A|\geq 2} a_A^2.$$

Recall: by the Plancherel theorem, $\sum_{A\subseteq[n]} a_A^2 = \mathbf{E}f^2 = \mathbf{E}1 = 1$.

FKN Theorem

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002): For $f : \{-1,1\}^n \to \{-1,1\}$, let $f = \sum_{A\subseteq[n]} a_A w_A$ be its unique Fourier-Walsh expansion, and let

$$\rho := \sum_{A\subseteq[n]:\,|A|\geq 2} a_A^2.$$

Then, among functions $1, -1, r_1, -r_1, r_2, -r_2, \ldots, r_n, -r_n$ there is a function g such that

$$\mathbf{P}(f \neq g) \leq C \cdot \rho,$$

where C is a universal (numerical) constant.

Remark: For f as above and for any $g \in \{\pm 1, \pm r_1, \pm r_2, \ldots, \pm r_n\}$, by the Plancherel theorem applied to $f - g$,

$$4 \cdot \mathbf{P}(f \neq g) = \mathbf{E}(f-g)^2 \geq \sum_{A\subseteq[n]:\,|A|\geq 2} a_A^2 = \rho.$$

FKN Theorem

FKN Theorem (Friedgut, Kalai, Naor / Kindler, Safra 2002): For $f : \{-1,1\}^n \to \{-1,1\}$, let $f = \sum_{A \subseteq [n]} a_A w_A$ be its unique Fourier-Walsh expansion, and let

$$\rho := \sum_{A \subseteq [n] \colon |A| \geq 2} a_A^2 .$$

Then, among functions $1, -1, r_1, -r_1, r_2, -r_2, \ldots, r_n, -r_n$ there is a function g such that

$$\mathbf{P}(f \neq g) \leq C \cdot \rho,$$

where C is a universal (numerical) constant.

Remark: For f as above and for any $g \in \{\pm 1, \pm r_1, \pm r_2, \ldots, \pm r_n\}$, by the Plancherel theorem applied to $f - g$,

$$4 \cdot \mathbf{P}(f \neq g) = \mathbf{E}(f-g)^2 \geq \sum_{A \subseteq [n] \colon |A| \geq 2} a_A^2 = \rho.$$

On the FKN Theorem

The FKN Theorem is one of the standard results of the Boolean analysis and it has found applications in theoretical computer science. In particular, it was used in the celebrated Irit Dinur's proof of the PCP Theorem.

PCP stands for Probabilistically Checkable Proof.

On the FKN Theorem

The FKN Theorem is one of the standard results of the Boolean analysis. It is one of the tools used in the celebrated Irit Dinur's proof of the PCP Theorem.

The FKN Theorem becomes an easy exercise if the universal constant C is replaced by a dimension-dependent C_n. However, until very recently, no elementary proof of the FKN Theorem was known, and the value of C obtained from the existing proofs was quite far from being optimal.

A new, simpler approach of **Jendrej, O., and Wojtaszczyk 2015** yields C close to the best possible constant and leads to various extensions of the FKN Theorem.

Here, Wojtaszczyk stands for Jakub Onufry Wojtaszczyk.

The FKN Theorem is one of the standard results of the Boolean analysis. It is one of the tools used in the celebrated Irit Dinur's proof of the PCP Theorem.

The FKN Theorem becomes an easy exercise if the universal constant C is replaced by a dimension-dependent C_n. However, until very recently, no elementary proof of the FKN Theorem was known, and the value of C obtained from the existing proofs was quite far from being optimal.

A new, simpler approach of **Jendrej, O., and Wojtaszczyk 2015** yields C close to the best possible constant and leads to various extensions of the FKN Theorem.

Here, Wojtaszczyk stands for Jakub Onufry Wojtaszczyk.

On the FKN Theorem

The FKN Theorem is one of the standard results of the Boolean analysis. It is one of the tools used in the celebrated Irit Dinur's proof of the PCP Theorem.

The FKN Theorem becomes an easy exercise if the universal constant C is replaced by a dimension-dependent C_n. However, until very recently, no elementary proof of the FKN Theorem was known, and the value of C obtained from the existing proofs was quite far from being optimal.

A new, simpler approach of **Jendrej, O., and Wojtaszczyk 2015** yields C close to the best possible constant and leads to various extensions of the FKN Theorem.

Here, Wojtaszczyk stands for Jakub Onufry Wojtaszczyk.

Preliminary reduction

As in many previous proofs of the FKN Theorem, it is natural to split the Boolean function f into low and high frequencies:

$$f = \sum_{A \subseteq [n]} a_A w_A = \sum_{A \subseteq [n]\colon |A| < 2} a_A w_A + \sum_{A \subseteq [n]\colon |A| \geq 2} a_A w_A,$$

Preliminary reduction

As in many previous proofs of the FKN Theorem, it is natural to split the Boolean function f, expressing it as $f = S + R$, where

$$S = a_\emptyset + a_{\{1\}} r_1 + a_{\{2\}} r_2 + \ldots + a_{\{n\}} r_n, \quad R = \sum_{A \subseteq [n]:\, |A| \geq 2} a_A w_A.$$

By the Plancherel theorem and assumptions of the FKN Theorem,

$$\mathbf{E} R^2 = \sum_{A \subseteq [n]:\, |A| \geq 2} a_A^2 = \rho,$$

so we can control the L^2-norm of the remainder term, $\|R\|_2 = \sqrt{\rho}$.

Clearly, the leading term S is a shifted Rademacher sum.

Preliminary reduction

As in many previous proofs of the FKN Theorem, it is natural to split the Boolean function f, expressing it as $f = S + R$, where

$$S = a_\emptyset + a_{\{1\}} r_1 + a_{\{2\}} r_2 + \ldots + a_{\{n\}} r_n, \quad R = \sum_{A \subseteq [n]\colon |A| \geq 2} a_A w_A.$$

By the Plancherel theorem and assumptions of the FKN Theorem,

$$\mathbf{E} R^2 = \sum_{A \subseteq [n]\colon |A| \geq 2} a_A^2 = \rho,$$

so we can control the L^2-norm of the remainder term, $\|R\|_2 = \sqrt{\rho}$.

Clearly, the leading term S is a shifted Rademacher sum.

Preliminary reduction

As in many previous proofs of the FKN Theorem, it is natural to split the Boolean function f, expressing it as $f = S + R$, where

$$S = a_\emptyset + a_{\{1\}}r_1 + a_{\{2\}}r_2 + \ldots + a_{\{n\}}r_n, \quad R = \sum_{A \subseteq [n]:\, |A| \geq 2} a_A w_A.$$

By the Plancherel theorem and assumptions of the FKN Theorem,

$$\mathbf{E}R^2 = \sum_{A \subseteq [n]:\, |A| \geq 2} a_A^2 = \rho,$$

so we can control the L^2-norm of the remainder term, $\|R\|_2 = \sqrt{\rho}$.

Clearly, the leading term S is a shifted Rademacher sum.

Upper bound on $\mathrm{Var}(|S|)$

Since $f = R + S$ and f is Boolean, we have $1 = |f| = |R + S|$, and thus, by the triangle inequality, $\Big||S| - 1\Big| \le |R|$, i.e. $(|S| - 1)^2 \le R^2$.

This allows us to bound from above the variance of $|S|$:

$$\mathrm{Var}(|S|) = \mathbf{E}|S|^2 - (\mathbf{E}|S|)^2 = \mathbf{E}(|S| - 1)^2 - (\mathbf{E}|S| - 1)^2 \le \mathbf{E}R^2 = \rho.$$

Actually, this will be the only information about S we will need in our proof – that it is a shifted Rademacher sum with $\mathrm{Var}(|S|) \le \rho$.

Since $f = R + S$ and f is Boolean, we have $1 = |f| = |R + S|$, and thus, by the triangle inequality, $\big||S| - 1\big| \leq |R|$, i.e. $(|S| - 1)^2 \leq R^2$.

This allows us to bound from above the variance of $|S|$:

$$\mathrm{Var}(|S|) = \mathbf{E}|S|^2 - (\mathbf{E}|S|)^2 = \mathbf{E}(|S| - 1)^2 - (\mathbf{E}|S| - 1)^2 \leq \mathbf{E}R^2 = \rho.$$

Actually, this will be the only information about S we will need in our proof – that it is a shifted Rademacher sum with $\mathrm{Var}(|S|) \leq \rho$.

Upper bound on $\mathrm{Var}(|S|)$

Since $f = R + S$ and f is Boolean, we have $1 = |f| = |R + S|$, and thus, by the triangle inequality, $\big||S| - 1\big| \le |R|$, i.e. $(|S| - 1)^2 \le R^2$.

This allows us to bound from above the variance of $|S|$:

$$\mathrm{Var}(|S|) = \mathbf{E}|S|^2 - (\mathbf{E}|S|)^2 = \mathbf{E}(|S| - 1)^2 - (\mathbf{E}|S| - 1)^2 \le \mathbf{E}R^2 = \rho.$$

Actually, this will be the only information about S we will need in our proof – that it is a shifted Rademacher sum with $\mathrm{Var}(|S|) \le \rho$.

Dominating coefficient problem

Our task amounts to proving that one of the numbers $|a_\emptyset|$, $|a_{\{1\}}|$, $|a_{\{2\}}|, \ldots, |a_{\{n\}}|$ is *close enough* to 1. Indeed, by the Plancherel theorem,

$$4 \cdot \mathbf{P}(f \neq r_k) = \mathbf{E}(f - r_k)^2 = (a_{\{k\}} - 1)^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 =$$

$$1 - 2a_{\{k\}} + a_{\{k\}}^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 = 1 - 2a_{\{k\}} + \sum_{A \subseteq [n]} a_A^2 = 2(1 - a_{\{k\}}).$$

Recall that $\sum_{A \subseteq [n]} a_A^2 = \mathbf{E}f^2 = \mathbf{E}1 = 1$, since f is Boolean.

Dominating coefficient problem

Our task amounts to proving that one of the numbers $|a_\emptyset|$, $|a_{\{1\}}|$, $|a_{\{2\}}|, \ldots, |a_{\{n\}}|$ is *close enough* to 1. Indeed, by the Plancherel theorem,

$$4 \cdot \mathbf{P}(f \neq r_k) = \mathbf{E}(f - r_k)^2 = (a_{\{k\}} - 1)^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 =$$

$$1 - 2a_{\{k\}} + a_{\{k\}}^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 = 1 - 2a_{\{k\}} + \sum_{A \subseteq [n]} a_A^2 = 2(1 - a_{\{k\}}).$$

Recall that $\sum_{A \subseteq [n]} a_A^2 = \mathbf{E} f^2 = \mathbf{E} 1 = 1$, since f is Boolean.

Dominating coefficient problem

Our task amounts to proving that one of the numbers $|a_\emptyset|$, $|a_{\{1\}}|$, $|a_{\{2\}}|, \ldots, |a_{\{n\}}|$ is *close enough* to 1. Indeed, by the Plancherel theorem,

$$4 \cdot \mathbf{P}(f \neq r_k) = \mathbf{E}(f - r_k)^2 = (a_{\{k\}} - 1)^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 =$$

$$1 - 2a_{\{k\}} + a_{\{k\}}^2 + \sum_{A \subseteq [n]:\, A \neq \{k\}} a_A^2 = 1 - 2a_{\{k\}} + \sum_{A \subseteq [n]} a_A^2 = 2(1 - a_{\{k\}}),$$

i.e. $\mathbf{P}(f \neq r_k) = (1 - a_{\{k\}})/2$.

Similarly, $\mathbf{P}(f \neq -r_k) = (1 + a_{\{k\}})/2$, $\mathbf{P}(f \neq -1) = (1 + a_\emptyset)/2$, and $\mathbf{P}(f \neq 1) = (1 - a_\emptyset)/2$.

Thus, here, *close enough* means *no further than* $2C\rho$ *apart.*

Dominating coefficient problem

Our task amounts to proving that one of the numbers $|a_\emptyset|$, $|a_{\{1\}}|$, $|a_{\{2\}}|, \ldots, |a_{\{n\}}|$ is *close enough* to 1. Indeed, by the Plancherel theorem,

$$4 \cdot \mathbf{P}(f \neq r_k) = \mathbf{E}(f - r_k)^2 = (a_{\{k\}} - 1)^2 + \sum_{A \subseteq [n]\colon A \neq \{k\}} a_A^2 =$$

$$1 - 2a_{\{k\}} + a_{\{k\}}^2 + \sum_{A \subseteq [n]\colon A \neq \{k\}} a_A^2 = 1 - 2a_{\{k\}} + \sum_{A \subseteq [n]} a_A^2 = 2(1 - a_{\{k\}}),$$

i.e. $\mathbf{P}(f \neq r_k) = (1 - a_{\{k\}})/2$.

Similarly, $\mathbf{P}(f \neq -r_k) = (1 + a_{\{k\}})/2$, $\mathbf{P}(f \neq -1) = (1 + a_\emptyset)/2$, and $\mathbf{P}(f \neq 1) = (1 - a_\emptyset)/2$.

Thus, here, *close enough* means *no further than $2C\rho$ apart.*

Dominating coefficient problem

Our task amounts to proving that one of the numbers $|a_\emptyset|$, $|a_{\{1\}}|$, $|a_{\{2\}}|, \ldots, |a_{\{n\}}|$ is *close enough* to 1. Indeed, by the Plancherel theorem,

$$4 \cdot \mathbf{P}(f \neq r_k) = \mathbf{E}(f - r_k)^2 = (a_{\{k\}} - 1)^2 + \sum_{A \subseteq [n]\colon A \neq \{k\}} a_A^2 =$$

$$1 - 2a_{\{k\}} + a_{\{k\}}^2 + \sum_{A \subseteq [n]\colon A \neq \{k\}} a_A^2 = 1 - 2a_{\{k\}} + \sum_{A \subseteq [n]} a_A^2 = 2(1 - a_{\{k\}}),$$

i.e. $\mathbf{P}(f \neq r_k) = (1 - a_{\{k\}})/2$.

Similarly, $\mathbf{P}(f \neq -r_k) = (1 + a_{\{k\}})/2$, $\mathbf{P}(f \neq -1) = (1 + a_\emptyset)/2$, and $\mathbf{P}(f \neq 1) = (1 - a_\emptyset)/2$.

Thus, here, *close enough* means *no further than* $2C\rho$ *apart*.

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

$$\min\left(\mathrm{Var}(X), \mathrm{Var}(Y)\right) \le M \cdot \mathrm{Var}(|X+Y|),$$

where M is a universal (numerical) constant.

Remark: It can be proved with $M = (7+\sqrt{17})/4 \simeq 2.78$.
On the other hand, it is false for $M < 16/7 \simeq 2.29$.

We will apply the lemma to the case of
$X = a_\emptyset r_0 + a_{\{1\}} r_1 + \ldots + a_{\{k-1\}} r_{k-1}$ and
$Y = a_{\{k\}} r_k + \ldots + a_{\{n\}} r_n$. Since $|X+Y|$ has the same distribution as $|S|$, $\mathrm{Var}(|X+Y|) = \mathrm{Var}(|S|) \le \rho$. So, for any choice of $k \in [n]$, we have

$$a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 = \mathrm{Var}(X) \le M\rho$$

or

$$a_{\{k\}}^2 + \ldots + a_{\{n\}}^2 = \mathrm{Var}(Y) \le M\rho.$$

Key Lemma (probabilistic inequality)

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

$$\min\left(\mathrm{Var}(X), \mathrm{Var}(Y)\right) \leq M \cdot \mathrm{Var}(|X+Y|),$$

where M is a universal (numerical) constant.

Remark: It can be proved with $M = (7+\sqrt{17})/4 \simeq 2.78$. On the other hand, it is false for $M < 16/7 \simeq 2.29$.

We will apply the lemma to the case of $X = a_\emptyset r_0 + a_{\{1\}} r_1 + \ldots + a_{\{k-1\}} r_{k-1}$ and $Y = a_{\{k\}} r_k + \ldots + a_{\{n\}} r_n$. Since $|X+Y|$ has the same distribution as $|S|$, $\mathrm{Var}(|X+Y|) = \mathrm{Var}(|S|) \leq \rho$. So, for any choice of $k \in [n]$, we have

$$a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 = \mathrm{Var}(X) \leq M\rho$$

or

$$a_{\{k\}}^2 + \ldots + a_{\{n\}}^2 = \mathrm{Var}(Y) \leq M\rho.$$

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

$$\min\left(\mathrm{Var}(X), \mathrm{Var}(Y)\right) \le M \cdot \mathrm{Var}(|X+Y|),$$

where M is a universal (numerical) constant.

Remark: It can be proved with $M = (7+\sqrt{17})/4 \simeq 2.78$.
On the other hand, it is false for $M < 16/7 \simeq 2.29$.

We will apply the lemma to the case of
$X = a_{\emptyset} r_0 + a_{\{1\}} r_1 + \ldots + a_{\{k-1\}} r_{k-1}$ and
$Y = a_{\{k\}} r_k + \ldots + a_{\{n\}} r_n$. Since $|X+Y|$ has the same distribution as $|S|$, $\mathrm{Var}(|X+Y|) = \mathrm{Var}(|S|) \le \rho$. So, for any choice of $k \in [n]$, we have

$$a_{\emptyset}^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 = \mathrm{Var}(X) \le M\rho$$

or

$$a_{\{k\}}^2 + \ldots + a_{\{n\}}^2 = \mathrm{Var}(Y) \le M\rho.$$

Final trick

Let us consider **the largest** $k \in [n]$ such that $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 \leq M\rho$. Then, obviously, $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k\}}^2 > M\rho$, so that, by Key Lemma, $a_{\{k+1\}}^2 + \ldots + a_{\{n\}}^2 \leq M\rho$.

Let us consider **the largest** $k \in [n]$ such that
$a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 \le M\rho$. Then, obviously,
$a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k\}}^2 > M\rho$, so that, by Key Lemma,
$a_{\{k+1\}}^2 + \ldots + a_{\{n\}}^2 \le M\rho$.

Final trick

Let us consider **the largest** $k \in [n]$ such that $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 \leq M\rho$. Then, obviously, $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k\}}^2 > M\rho$, so that, by Key Lemma, $a_{\{k+1\}}^2 + \ldots + a_{\{n\}}^2 \leq M\rho$.

But we have already proved that

$$\sum_{A \subseteq [n]:\, |A| \geq 2} a_A^2 = \mathbf{E} R^2 \leq \rho.$$

Thus

$$\sum_{A \subseteq [n]:\, |A| \neq \{k\}} a_A^2 \leq M\rho + M\rho + \rho = (2M+1)\rho.$$

Since $\sum_{A \subseteq [n]:\, |A|} a_A^2 = \mathbf{E} f^2 = 1$, we arrive at $a_{\{k\}}^2 \geq 1 - (2M+1)\rho$, so that $1 - |a_{\{k\}}| = O(\rho)$, and the proof is finished.

Final trick

Let us consider **the largest** $k \in [n]$ such that $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 \le M\rho$. Then, obviously, $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k\}}^2 > M\rho$, so that, by Key Lemma, $a_{\{k+1\}}^2 + \ldots + a_{\{n\}}^2 \le M\rho$.

But we have already proved that

$$\sum_{A \subseteq [n]:\, |A| \ge 2} a_A^2 = \mathbf{E} R^2 \le \rho.$$

Thus

$$\sum_{A \subseteq [n]:\, |A| \ne \{k\}} a_A^2 \le M\rho + M\rho + \rho = (2M+1)\rho.$$

Since $\sum_{A \subseteq [n]:\, |A|} a_A^2 = \mathbf{E} f^2 = 1$, we arrive at $a_{\{k\}}^2 \ge 1 - (2M+1)\rho$, so that $1 - |a_{\{k\}}| = O(\rho)$, and the proof is finished.

Final trick

Let us consider **the largest** $k \in [n]$ such that $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k-1\}}^2 \leq M\rho$. Then, obviously, $a_\emptyset^2 + a_{\{1\}}^2 + \ldots + a_{\{k\}}^2 > M\rho$, so that, by Key Lemma, $a_{\{k+1\}}^2 + \ldots + a_{\{n\}}^2 \leq M\rho$.

But we have already proved that

$$\sum_{A \subseteq [n]:\, |A| \geq 2} a_A^2 = \mathbf{E} R^2 \leq \rho.$$

Thus

$$\sum_{A \subseteq [n]:\, |A| \neq \{k\}} a_A^2 \leq M\rho + M\rho + \rho = (2M+1)\rho.$$

Since $\sum_{A \subseteq [n]:\, |A|} a_A^2 = \mathbf{E} f^2 = 1$, we arrive at $a_{\{k\}}^2 \geq 1 - (2M+1)\rho$, so that $1 - |a_{\{k\}}| = O(\rho)$, and the proof is finished.